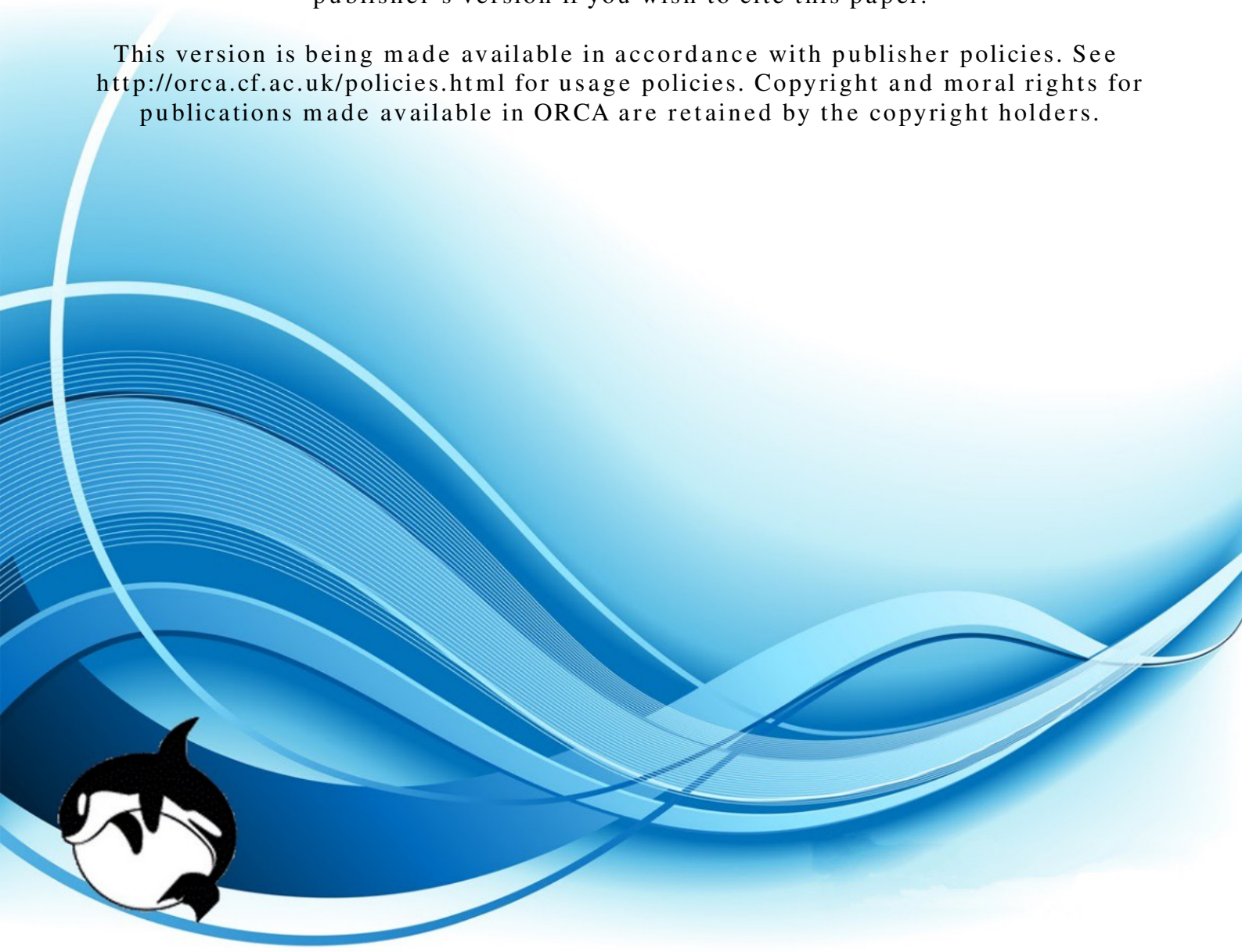


 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

AR mRNA Stability is Increased with AR-antagonist Resistance Via 3' UTR Variants

D. A. Dart[1&3], K. Ashelford[2], W.G. Jiang[1]

[1]Cardiff China Medical Research Collaborative, Cardiff University School of Medicine, Cardiff, CF14 4XN, Wales, UK.

[2]Wales Gene Park, Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, CF14 4XN, Wales, U.K.

[3]Imperial College London, Exhibition Road, London SW7 2AZ, UK.

Running title: Novel AR 3'UTR variants in anti-androgen resistance.

Corresponding author: a.dart@imperial.ac.uk

Keywords: prostate, cancer, bicalutamide, enzalutamide, anti-androgen, therapeutics

Abstract

Advanced prostate cancer is often treated with AR-antagonists which target the androgen receptor (AR) on which the growth of the tumour depends. Prostate cancer often develops AR-antagonist resistance via a plethora of mechanisms, many of which are as yet unknown, but it is thought that AR upregulation or AR ligand binding site mutations, may be responsible. Here we describe the production of cell lines based on LNCaP and VCaP, with acquired resistance to the clinically relevant AR-antagonists, bicalutamide and enzalutamide.

In these resistant cells, we observed, via RNA-seq, that new variants in the 3'UTR of the AR mRNA were detectable and that the levels were increased both with AR-antagonist treatment and with hormonal starvation. Around 20% of AR transcripts showed a 3kb deletion within the 6.7kb 3'UTR sequence. Actinomycin D and luciferase fusion studies indicated that this shorter mRNA variant was inherently more stable in anti-androgen resistant cell lines.

Of additional interest was that the AR UTR variant could be detected in the sera of prostate cancer patients in a cohort of serum samples collected from patients of Gleason grades 6-10, with an increasing level correlated to increasing grade.

We hypothesise that the shorter AR UTR variant is a survival adaptation to low hormone levels and/or AR-antagonist treatment in these cells, where a more stable mRNA may allow higher levels of AR expression under these conditions.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Introduction.

Targeting the androgen receptor (AR) via AR-antagonist therapies is still the mainstay of advanced prostate cancer (PCa) treatments (1,2). AR-antagonists, such as enzalutamide, flutamide and bicalutamide target and competitively bind the AR, to the exclusion of testosterone (or DHT). Additionally AR-antagonists may mediate a misfolding of AR helix 12 and cause the AR to become inactive (3). In either event, the starvation of prostate cancer cells of testosterone leads to a reduction in tumour cell growth. Often patients undergoing AR-antagonist therapy can become resistant to bicalutamide via many pathways - involving AR amplification, AR ligand binding domain (LBD) mutations, cofactor/coregulator changes, and multidrug efflux pumps. Alternatively, in high grade prostate cancers, tumours may synthesise their own androgens from precursor molecules as a mechanism for escape from such therapies (reviewed in (4). Often PCa cells can acquire AR LBD mutations which can change the activity of an antagonist to an agonist, and in some cases patients may acquire PCa that proliferates in the presence of AR-antagonists. Such patients may respond better (initially) when AR-antagonists are withdrawn (5,6). In addition to all these processes, prostate cancer cells have been shown to produce AR splice variants – truncated versions of the AR with missing ligand activation domain which remain constitutively active *e.g.* truncated version known as ARv7 (7).

The LNCaP cell line represents a human PCa model derived from a patient who relapsed with flutamide therapy (8). The LNCaP cell line contains an AR mutation within the LDB (T877A), which allows a greater plethora of ligands to bind and activate the AR, including flutamide. However, the cell line remains sensitive to AR-antagonists such as bicalutamide and enzalutamide, but several publications report that prolonged treatment of the LNCaP cell line with AR-antagonists or prolonged hormonally starved cells can produce relevant cell lines that mimic clinically acquired AR-antagonists resistance.

The VCaP cell line was established in 1997 from a lumbar vertebral bone metastasis from a metastatic lesion from a patient with hormone refractory prostate cancer. The cell line has been reported to have wild type AR but overexpresses it at high levels (9).

Recently, we generated bicalutamide and enzalutamide resistant clones of the LNCaP and VCaP cell lines, as tools for screening novel AR-antagonists compounds (10,11). The VCaP$^{BicR/EnzR}$ resistant clones showed resistance to both bicalutamide and enzalutamide – showing cytostatic effects at high doses. However, the LNCaP^{BicR} cell lines showed a proliferative response to bicalutamide, whilst LNCaP/EnzR cells showed a more cytostatic effect at high doses.

In an attempt to discover the potential mechanism of AR-antagonist resistance and to discover if any novel AR variants were being expressed or if novel LBD mutations were present, we subjected LNCaP^{BicR} cells to RNA-seq. Here we discuss the finding that although no new coding variant or mutations were seen, we did discover an increase in an AR variant, in the 3'UTR region. Here we investigate how the variant conveys extra stability to the AR mRNA transcript and may be in part responsible for increased AR mRNA and protein expression in response to AR-antagonist mediated hormone starvation.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Materials and Methods

Cell culture

LNCaP cells were maintained at 37°C, 5% CO_2 in RPMI medium with 10% foetal bovine serum (First Link Ltd, Brierley Hill, U.K.). PC3, VCaP and Du145 cells were maintained in DMEM medium (Sigma) with 10% foetal bovine serum (First Link UK,). All media was supplemented with 2mM L-glutamine, 100units/ml penicillin, 100mg/ml streptomycin (Sigma). For hormone depletion experiments, 72 hours before androgen exposure, medium was replaced with 'starvation medium' consisting of phenol red-free RPMI (or DMEM) medium, supplemented with 5% charcoal-stripped foetal bovine serum (First Link UK). All cells were obtained from the ATCC cell bank in 2017, and kept in liquid N_2 in aliquots thereafter. All cells were used within 10 passages of the original stock. RNA-seq of the samples also verified these cell lines in terms of oncogenes and documented mutations.

Treatments

All AR-antagonist compounds were obtained from Sigma, and made up in DMSO stock at 10mM and kept at -20C. Actinomycin-D was made up in DMSO at 100μg/ml, and kept at -20C. Working solution was 1μg/ml.

Generation of bicalutamide resistant cells

LNCaP and VCaP cells were grown in increasing doses of bicalutamide or enzalutamide (0-30μM) for up to 6 weeks at which time cells began to show proliferation. Resistant cell lines were expanded from these and resistant clones grew stably in medium containing 20μM AR-antagonist.

MTT Assays

The MTT assay was used as a cell viability assay for the all the cell lines listed using the AR-antagonist compounds. Briefly, cells (5×10^4 cells/ml) were seeded into 96-well plates (200ul/well), allowed to attach and grow for 24hr and subsequently treated with varying concentrations of AR-antagonists (0–100 μM) for 96 hours. Optimal seeding densities were premeasured for linear growth over the 96hours. MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, 5mg/ml in PBS) was added to a final concentration of 0.5mg/ml for 4 h at 37°C. After four hours, purple formazan crystals, formed by mitochondrial reduction of MTT, were solubilized in acidified isopropanol

(200µl/well) and the absorbance was read at 570 nm after 10 min incubation. Percent inhibition of viability was calculated as a % of untreated control and the cytotoxicity / cell growth inhibition was expressed as IC_{50}.

AR 3'UTR Transfection and Luciferase Assays

AR UTR reporters were a kind gift from Dr Paivi Ostling (Science for Life Laboratory, Sweden). Cells were transfected with AR-UTR reporter vectors (12), and a constitutive expression renilla vector (pGL4.75 Promega) alongside using Lipofectamine 3000 (Life Technologies). 24 hours after transfection cells were washed and lysed in reporter lysis buffer (Promega). Lysate was mixed with D-luciferin substrate (Promega), and then with renilla substrate (Promega). Light emission measured using the Promega Glomax multi luminometer. Luciferase activity was normalised to renilla expression.

RNA extraction and RT-PCR

Total RNA samples were prepared using Trizol reagent (Sigma) and converted to cDNA using the GoScript™ Reverse Transcription System (Promega).

Q-PCR

Reactions were performed in triplicate on cDNA samples in 96-well optical plates on an ABI Prism StepOne System (ThermoFisher, U.K.). Reactions consisted of 2 µl cDNA, 7 µl PCR-grade water, 10 µl 2× TaqMan Universal PCR Master Mix (Applied Biosystems), 1 µl Taqman specific assay probes (Applied Biosystems) for PSA, and RPL19 or SYBR green primers for AR, AR-UTR (variants), β-actin and GAPDH. Parameters were: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 sec and 60°C for 1 min. Data was recorded using Sequence Detector Software (SDS version 2.3; PE Applied Biosystems). Levels were normalised to GAPDH, β-actin and RPL19.

RNA-seq analysis

PolyA mRNA was isolated from total RNA using the Dynabeads mRNA DIRECT Kit (Life Technologies, UK) and verified using a Bioanalyser-2100 (Agilent). RNA fragment libraries (150-200bp) were generated using the Ion Total RNA-Seq kit (Life Technologies) and ligated to adapters for cDNA

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

synthesis. cDNA was then amplified using IonXpress RNA-seq barcoded primers (5') and quantified with a Qubit assay (Life Technologies).

cDNA libraries were clonally amplified by emulsion PCR on Ion Sphere Particles (ISP's) using Ion PI template OT2 200 kit (Life technologies, USA) on an Ion OneTouch2 system (Life technologies) as per manufacturer's instructions. The template positive ISP's were recovered and enriched to remove non-template ISP's on Ion One Touch ES (Life technologies). The ISPs were processed using the Ion Proton 200 sequencing kit and loaded onto a P1 chip and sequenced on an Ion Proton (Life technologies) using default parameters (single-end, forward sequencing). Base calling, adaptor trimming, barcode deconvolution and alignment was performed on Torrent Suite version 3.6 (Life technologies) using the STAR RNA-seq aligner plugin. The Partek Genomic Suite 6.6 software was used for data analysis. The RPKM normalization method for RNA-seq 45 was used followed by a 1-way Anova test for gene differential expression (from n=4 samples per group). For transcript analysis, we utilised the RNA STAR aligner (or HISAT2 v2.1.0) followed by the Bowtie and TopHat splice aware mapping software (UCSC Galaxy). Finally, Cufflinks (version 2.2.1) and DESeq2 (version 1.12.4 running on R version 3.3.0) were used to build transcript information and visualised in IGV (Integrative Genomics Viewer) software (Broad Institute, University of California, USA).

For online GEO dataset analysis, we used an SRA download tool (UCSC Galaxy), followed by alignment as listed above, and analysis via StringTie (v1.3.6).

Serum collection

Clinical samples were obtained from the Wales Cancer Bank (WCB) in accordance with their ethics and patient consent, following application and approval (WCB application number 15/009). In total 124 samples of serum were used from patients with confirmed prostate cancer. The median patient age was 65 years and the median follow up period was 4 years. Patient serum was collected at the point of first diagnosis based on pathological Gleason grade from biopsy, prior to any treatment. No TNM data was available. In addition we collected serum samples from normal healthy male volunteers, with an age range of 20-55 years (ethics number SMREG ref:13.69). 50µl of serum was used and RNA was extracted using 150µl of Trizol LS.

Results

AR mRNA transcript and protein expression are increased with acquired AR-antagonist resistance.

We set out to analyse and study the mechanism of action of acquired AR-antagonists resistance in prostate cancer cell models. The baseline sensitivity of LNCaP and VCaP cells to bicalutamide and enzalutamide was measured using the MTT assay for cell viability. Treating cells for 96hours in increasing doses of bicalutamide showed an IC_{50} of 3μM for LNCaP cells and 2.5μM for VCaP cells. Enzalutamide gave IC_{50} values of 2 and 2.3μM respectively. After continual culture and growth in 20μM bicalutamide or enzalutamide for 6 weeks approx. these cells began to grow and proliferate as normal, and cell lines were expanded – henceforth labelled $LNCaP^{BicR/EnzR}$ and $VCaP^{BicR/EnzR}$. A further MTT analysis for the IC_{50} revealed that $LNCaP^{BicR}$ cells now proliferated in bicalutamide reaching maximal growth at approx. 20μM (see figures 1A). Enzalutamide still showed some activity in $LNCaP^{EnzR}$ cells, but the IC_{50} was now 15μM (approx.) (see figure 1A). In VCaP resistant cell lines, both bicalutamide and enzalutamide failed to give an IC_{50} dose, but no dose-responsive proliferation was observed (see figure 1B).

Analysis of AR and prostate specific antigen (PSA) expression in these cell lines confirmed that VCaP and LNCaP were indeed AR^{+ve} when compared to PC3 and Du145 – cell lines known to be androgen receptor negative (see figure 1C). When we compared parental and resistant cell lines we observed that AR expression was increased 2-fold in the resistant cell lines (see figure 1D) – although the levels of AR transcripts was much higher in VCaP cells (x100 fold). We PCR-amplified and sequenced the AR LBD domain from these cell lines and did not observe any AR mutations apart from T877A, carried by the LNCaP cell line. Protein levels for AR also increased in the resistant LNCaP cell lines by 1.3 fold (see figure 1E).

AR transcripts differ in the 3'UTR region in LNCaP with acquired bicalutamide resistance.

We then analysed the changes in gene expression in the LNCaP vs $LNCaP^{BicR}$ cell lines by RNA-seq. When analysing the AR transcripts we did not observe any AR coding region splice variants *e.g.* ARv7, however we did observe that the 3'UTR region of some AR transcripts showed an anomaly. Data is available at NCBI BioProjects accession PRJNA578661 (ncbi.nlm.nih.gov).

In the parental LNCaP cells, three distinct AR transcripts were identified. Two were complete matches of known transcripts (according to RefGene gene model) and one was a potentially novel transcript. However, in the $LNCaP^{BicR}$ cell line five distinct AR transcripts were identified, two of which were complete matches of known transcripts and three were potentially novel transcripts.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Figure 2A showed the analysis schematic from IGV software, showing the Cuffmerge data output. Figure 2B showed a detailed schematic of the 3'UTR region of AR. The novel transcripts had a missing 3kb region in the UTR sequence, from bases 1795-4832 of the UTR sequence. The coordinates are counted from +1 after the stop site in the AR mRNA sequence (Refseq: NM_000044), see supplemental data. The missing region corresponds to ChrX: position 66,945,4780 - 66,948,515 (GRCh37/hg19 assembly). Additionally, we saw another transcript with a slightly larger missing region from bases 1795-4961 (see figure 2B). The novel transcript we will henceforth referred to as $AR^{3.7kb\text{-}UTR}$, as opposed to $AR^{6.7kbUTR}$, for the normal full length AR transcript.

3'UTR deletion correlates with splice sites.

From the IGV analysis, reads mapping to the 3'UTR spanned the missing 3kb region, indicating that a splicing or editing event was responsible. Therefore, we analysed the 3'UTR region of the AR mRNA via various online resources for splice site analysis. Utilising Spliceport (http://spliceport.cbcb.umd.edu/SplicingAnalyser2.html), NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) and RegRNA (http://regrna2.mbc.nctu.edu.tw/detection.html) analysis software we found that the bases 1795 and 4832 (and 4961) were predicted splice donor and acceptor regions respectively with the highest confidence value or free energy score. Figure 2B shows a schematic of the base numbers and deleted region of the 3'UTR of AR. Figure 2C shows the output from SplicePort for both donor and acceptor predicted sites. We did not detect the novel spliced region in genomic DNA from LNCaP cells, again indicating a post transcriptional modification.

Spliced AR 3'UTR variant increases due to hormonal starvation or AR-antagonist treatment.

Utilising primers specific for various regions across the UTR (see figure 3A), we analysed if the shorter UTR transcript could be detected by PCR. Utilising primers C&D for the long UTR variant and B&E for the shorter UTR variant (see figure 3A), we carried out qPCR on the cell lines. The C&D primer set amplify a region present only in the complete full length AR 3'UTR. The B&E primer set only amplify the AR 3'UTR across a small region (200bp) of the shorter spliced variant.

Parental VCaP cells showed a 15-fold increase over LNCaP in AR-UTR expression (see figure 3B). However, interestingly, parental VCaP cells also showed a 4-fold increase in the novel shorter UTR variant than LNCaP cells (see figure 3B). Further, we then analysed the levels of the normal UTR (primers C&D), and UTR variant ($AR^{3.7kb\text{-}UTR}$ - primers B&E) in the resistant cell lines compared to the

parental cell lines. This also showed increased expression of the shorter variant in all resistant cell lines (see figure 3C), compared to their parental cells.

We then went on to study if the increase in levels of the 3'UTR variants in the parental cells changed rapidly following either hormone starvation or AR-antagonist treatment, or whether the AR variant was a product of a longer term adaptation. When LNCaP cells were hormonally starved for 72hours we saw a strong reduction in the levels of the androgen-inducible gene, PSA, as expected (see figure 3D upper panel). Also, as expected, when starved, LNCaP cells upregulated their levels of AR mRNA, therefore, the AR coding region and the normal UTR levels showed an increase of around 6-8 fold (see supplemental figure 1.). However, the shorter UTR variant could be detected at approx. 20 times the level found in parental cells (see figure 3D lower panel). Hormonal starvation for 96 hours also increased the level of AR protein by 2 fold (see figure 3E). When the parental LNCaP cells were treated with 20μM bicalutamide for up to six days, we observed a decrease in PSA, as expected, but an increase in shorter UTR variant, which increased over time (see figure 3F).

The spliced AR 3'UTR variant is more stable in cells with acquired AR-antagonist resistance.

To monitor if the shorter AR UTR variant conferred an advantage to the stability of the AR mRNA strand we treated parental and AR-antagonist resistant cells with actinomycin-D to halt RNA transcription. Total RNA was then extracted from cells at various timepoints. In LNCaP cells actinomycin-D significantly reduced expression of all transcripts over 24 hours. The levels of GAPDH, β-actin and AR (coding region), $AR^{6.7kb\text{-}UTR}$ did not show any significant differences between the cell lines. However, in the resistant cells the residual levels of $AR^{3.7kb\text{-}UTR}$ transcript showed an increased level compared to the parental cell lines at 10 hours (see figure 4A), but was similarly reduced at 24 hours. Similar results were seen in VCaP cells, but to a reduced level (see supplemental figure 2).

We then tested the relative stability of different regions of the AR UTR using luciferase reporter constructs. Using seven AR UTR overlapping constructs (12) of approximately 1kb each, were transfected into LNCaP or $LNCaP^{BicR}$ cells for 24hours, and measured luciferase activity over a constitutively expressed renilla luciferase. Regions 1, 2, 3, 6 and 7 showed increased stability in the $LNCaP^{BicR}$ cells, but regions 4 and 5 which lie within the spliced region of the UTR showed a reduced stability, (see figure 4B).

The 3'UTR variant is present in online GEO datasets including LNCaP-abl cells, immortalised prostate epithelial cells and in patient samples prior to AR-antagonist therapy.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Both the LNCaP and VCaP cell lines showed low levels of the AR UTR variant before the treatment with AR-antagonist, but these cell lines represented cells taken from patients with advanced prostate cancer who had been treated with various endocrine therapies e.g. flutamide (and/or other treatments). We, therefore, analysed several online data sets from the GEO database for the presence of the 3'UTR variant. In all studies analysed, the 3'UTR missing region was seen at some levels, albeit very low. Using online data from Olsen *et al*, GSE71797 (13), in hTert immortalised benign prostate epithelial cells we detected the presence of the transcript in some samples but not all (see supplemental figure 3A). Additionally, using data from Rajan *et al*, GSE48403 (14) we again found the presence of the shorter AR UTR variant. The 3'UTR variant was seen to be elevated in prostate cancer patients post AR-antagonists therapy, but due to low sample numbers (n=6) the data did not reach significance. Data is given in supplemental figure 3B.

The AR UTR was also detected and was increased (p=0.03) in a study by Knuuttila *et al*, (GSE95413) in orthotopically grown castration resistant VCaP xenografts treated with enzalutamide (see supplemental figure 4A). In a study by Shah *et al*, (GSE81796) using C4-2 cells (CRPC) the $AR^{3.7kb\text{-}UTR}$ was present and was increased by enzalutamide treatment (see supplemental figure 4B). And in a study by Coleman et al, (GSE87153), the $AR^{3.7kb\text{-}UTR}$ was seen to be increased in the LNCaP-derived MR49F cell line – grown in castrated mice and enzalutamide treated (see supplemental figure 4C). The variant was present in V16D cell line (LNCaP derived, castration resistant), but was not further elevated with enzalutamide. This missing region in the AR UTR was also observed in a study by Kohli *et al*, (GSE70380), in samples derived from metastatic deposits from a patient with disease progression whilst on androgen deprivation therapy (see supplemental figure 4D).

The LNCaP-abl cell line represents LNCaP cells grown in androgen-depleted medium for over 80 passages, and mimics castrate-resistant prostate cancer. A GEO dataset (GSE114708) utilising the LNCaP-abl cell line (15) showed the presence of the $AR^{3.7kb\text{-}UTR}$ variant, see supplemental figure 4E.

The AR 3'UTR variant is detectable in prostate cancer patient serum.

Since prostate cancer cells and/or prostate cancer cellular DNA/RNA may circulate in the serum of prostate cancer patients, we evaluated whether the novel variant could be detected in their sera. We extracted total RNA from prostate cancer patients' sera, collected from the Wales Cancer Bank (n=124) ranging from Gleason grade 6-10, with associated PSA values. Additionally, serum was collected from apparently healthy male volunteers (n=8). These were determined by clinical pathologists at the Heath Hospital Cardiff, UK. RNA was isolated and qPCR performed, for AR (coding

region), $AR^{6.7kb\text{-}UTR}$ (primers C&D), and for the $AR^{3.7kb\text{-}UTR}$ (primers B&E), along with GAPDH, β-actin and RPL19. Patient data and numbers can be seen in supplemental Table 1.

When analysed according to Gleason grade, we found a positive correlation with PSA as expected, but additionally we saw a positive correlation with increasing grade and the quantity of $AR^{3.7kb\text{-}UTR}$ variant in the serum (see figure 5A&B), and statistically significant differences between Gleason grade 6 and 9 ($p = <0.05$). No correlation was seen between the levels of AR coding region and Gleason grade, however, a positive correlation with grade was also seen for the full length $AR^{6.7kb\text{-}UTR}$, see figures 5C&D.

When analysed for patient follow up and survival, both PSA and $AR^{3.7kb\text{-}UTR}$ was elevated in the deceased group, ($p = <0.05$ & $p = <0.02$) respectively. However, due to very small numbers we treat these results with caution. $AR^{6.7kb\text{-}UTR}$ and AR coding region levels did not correlate with survival.

Discussion

During prostate cancer treatment AR-antagonists are often used to reduce tumour growth and prolong life. Initially, AR-antagonists are effective but frequently these become ineffective and patients relapse with prostate cancer which is resistant to further AR-antagonist treatment. Additionally, AR-antagonists may bind and activate the AR causing a stimulatory effect. In instances such as these, prostate cancer patients may benefit from AR-antagonist withdrawal. We set out to investigate the mechanism behind acquired resistance to AR-antagonist in prostate cancer cells by generating cells resistant to the two most clinically relevant AR-antagonists, namely bicalutamide and enzalutamide.

For LNCaP and VCaP cells, the acquisition of AR-antagonist resistance was established at 20μM, whereas at 30μM no growth or very slow cell growth was observed. However, when these resistant cells were then exposed to a range of drug concentrations 0-100μM we saw different effects dependent on the drug and cell line. VCaP showed resistance to enzalutamide and bicalutamide with no apparent cell killing up to 50-100mM. LNCaP showed similar properties with enzalutamide. However, LNCaP cells showed a proliferation response with bicalutamide, peaking at 20μM. This effect is probably due to the steric fit of bicalutamide in the mutant AR ligand binding domain (T877A) carried by the LNCaP cell line causing a more promiscuous ligand binding. This may also mimic the clinical situation where patients sometimes respond better – albeit briefly – when AR-antagonists are removed.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

When analysing AR mRNA levels in these resistant cells, we found expression to be higher at the mRNA and protein level. The increased expression of AR has been previously reported in cells either starved of hormone or inhibited by AR-antagonists. VCaP cells show a much higher baseline expression of AR and has been reported previously, but even in these cells AR expression was still higher in the resistant clones.

When we analysed the AR by RNA-seq we did not discover any new or additional mutations in any of these cells lines, neither did we find any novel AR coding variants *e.g.* ARv7. However, the RNA-seq and associated Cufflinks analysis did reveal an increase in transcript number and variation. Three isoforms were seen in parental LNCaP cells. The first being AR variant 1 (refseq: NM_000044) and the other being AR variant 2 (refseq: NM_001011546), which are previously identified variants. Variant 2, has an alternate start codon in the 5' region, with an encoded isoform (variant 2) which has a distinct and shorter N-terminus than isoform 1. However, even in these parental cells another isoform was seen with a shorter 5' UTR sequence and a missing 3kb region from the 3'UTR. In the $LNCaP^{BicR}$ cell line five transcripts were seen - the three variants seen previously in the parental cell lines and two more variants which were the AR variant 2 but with a missing 3kb region in the 3'UTR and an alternate variant 1 with a missing slightly larger 3.1kb region.

Since RNAseq reads spanned this area, and were aligned and mapped as genuine reads to the AR gene, we had to surmise that the missing 3 – 3.1kb were caused by 'splicing' or 'stitching' events. Splice site analysis matched up precisely with donor and acceptor sites on several web based platforms. The novel variant region could not be PCR amplified from genomic DNA, also indicating a post transcriptional splicing event.

Although AR coding and AR UTR levels were elevated in AR-antagonists resistant cells, PCR analysis confirmed that the missing UTR variant elevated in increasing amounts, indicating that additional AR transcripts with the missing region were present in these cells. Even though the high levels of $AR^{3.7kb\text{-}UTR}$ were seen in fully resistant cells, we saw that the adaptive response was rapid, being detectable after 24hours of AR-antagonist treatment and after 6 days of hormone starvation.

Additionally, to test whether the AR 'splicing' event was unique to AR pathway inhibition or was a more general stress response, we subjected LNCaP cells to a variety of cellular stresses including hypoxia (mimicked by $CoCl_2$), DNA damaging stress (by doxorubicin, cisplatin and paclitaxel) and environmental stress such as oxygen radicals (by H_2O_2) and extracellular low pH. We observed that both AR variants (3.7 & 6.7kb UTRs) were upregulated in response to environmental stresses *e.g.* oxygen or pH, but were not significantly altered in response to DNA damaging agents (see supplemental figure 5). All stress treatments reduced cellular proliferation as measured by the

expression of the MCM5 DNA replication gene. Oxygen level changes and low pH are significant factors within tumours, and may together with AR antagonism, play a role in prostate cancer. The 'splicing events observed here may part of a more generic stress-induced novel or aberrant splicing pathway.

We hypothesised that the AR variant was a result of an adaptive response to lack of androgen where the AR mRNA levels increased, and as is the case for alternative 'splicing', the mRNA strand itself may have an altered stability. Splicing of the coding region of the AR has been studied extensively, but UTR variants have not (7,16). For example if the shorter UTR variant was more stable, then this would give the AR mRNA a longer half-life, and a higher expression which would give cells a survival advantage. We tested this hypothesis by using actinomycin-D to stop cellular transcription of all transcripts and measured the halflife of the remaining transcripts. 24hour of actinomycin treatment reduced all transcripts measured including GAPDH, β-actin, $AR^{6.7kb\text{-}UTR}$ and $AR^{3.7kb\text{-}UTR}$. However the $AR^{3.7kb\text{-}UTR}$ showed a higher level, and reduced rate of degradation. Additional analysis using AR UTR reporters fused to luciferase showed that two regions (4 & 5) showed a reduced activity / stability in the $LNCaP^{BicR}$ cell line, whereas all other regions showed an increased activity. Regions 4 & 5 were in the missing 3kb region of the novel transcripts. Taken together, we hypothesise that in the adaptation to AR-antagonist resistance the 3kb RNA region missing in $AR^{3.7kb\text{-}UTR}$ may be detrimental for stability of the mRNA strand and may be spliced out to increase stability as a survival advantage for the cancer cells. Although a hypothetical explanation at this stage, we would suggest that this splicing or stitching of the UTR sequence may be a mechanism to avoid or circumvent microRNA or noncoding RNA downregulation of the AR strand. MiR analysis software indicated at least 40 potential binding sites in the 3kb were missing from the $AR^{3.7kb\text{-}UTR}$ mRNA strand. For example, if microRNA upregulation was a part of the adaptive AR-antagonist resistance response, then the shortening of the UTR would lessen the chance of AR itself being targeted by these miRs. Studies by Ostling *et al*, 2011, found several active microRNAs which target the 3'UTR of AR in the deleted region *e.g.* miR-185, miR-371-3p and miR-135. MiR-185 has been shown to downregulate AR levels and is itself downregulated in prostate cancer samples, Chunyan *et al*, 2015; Qu *et al* 2013.

Splicing dysregulation has been hypothesised to be associated with cancer and survival mechanisms *e.g.* in PI3kinase/Akt pathways (17,18). 3'UTR splicing events have been found for other genes and is associated with transcript stability (19-22). Often cells utilise alternative cleavage and polyadenylation sites to reduce UTR length and increase mRNA stability (Mayr *et al*, 2009), but in the enzalutamide resistant cells studied here, although the UTR length is indeed shortened the terminal polyadenylation site is unchanged. This would again imply a novel mRNA 'splicing' and editing event has occurred for the AR.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

The $AR^{3.7kb\text{-}UTR}$ variant was present in the parental LNCaP and VCaP cell lines, but was more prominent in the resistant cells. We then looked at several online GEO dataset to determine if the variant was common or cancer specific. In immortalised (hTERT) prostate epithelial cells and in epithelial cells undergoing mesenchymal transition, the variant was present but at very low levels, however, in cancer samples and cancer samples from AR-antagonist resistant refractory patients we found the levels to be higher, with even higher levels in refractory patients – although not significant due to low numbers (p=0.3). Therefore, the variant may be a naturally occurring transcript present at low levels, which is stabilised or upregulated at times of cellular stress.

Patients' plasma and serum have been used in the other studies to measure circulating AR levels (23,24). In prostate cancer patients' sera, we tested whether the AR variant could be detected, as well as the normal AR UTR and the AR coding region. Additionally, we compared the patient Gleason grade and PSA levels as indicators of disease. All three AR regions were detectable but it was only the UTR regions that correlated with grade across the patients' samples, and showed significant differences between grades 6 and 9 & 10. Additionally, both PSA and $AR^{3.7kb\text{-}UTR}$ showed a significantly higher level in those patients who had died of prostate cancer in the interim follow up. This was a surprising finding as numbers were low, but no significant correlation could be seen between PSA and $AR^{3.7kb\text{-}UTR}$. To our knowledge, these patients were naive to any endocrine treatment protocols, therefore taken together with the online GEO datasets, the appearance of the variant is not caused by AR-antagonist treatments but is upregulated by it. Unfortunately, follow up data on how patients went on to respond to AR-antagonist therapies were unavailable at the time of this study. Previously, the detection of AR-v7 in whole blood did not predict effectiveness of AR-antagonist for castration-resistant prostate cancer (24).

In conclusion, we report that an $AR^{3.7kb\text{-}UTR}$ variant levels increase in prostate cancer cells which have acquired resistance to bicalutamide / enzalutamide or indeed shorter term hormonal starvation. This variant is present at low levels in most cells studied, but the levels are increased in the resistant cells, and the novel variant may be associated with a splicing event. We hypothesise that this variant increases AR mRNA stability, thus may increase protein levels and may provide prostate cancer cells a survival advantage.

Acknowledgements

The authors would like to thank our funders – The Cardiff University – Peking University Cancer Institute and Cancer Research UK. Tissue (serum) samples were obtained from the Wales Cancer

Bank which is funded by the Wales Assembly Government and Cancer Research Wales. Other investigators may have received specimens from the same subjects.

Disclosure of Potential Conflicts of Interest

We declare that there are no conflicts of interests.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

References:

1. Brinkmann AO, Trapman J. Prostate cancer schemes for androgen escape. *Nat Med*. 2000;6(6):628-629.
2. Brinkmann AO. Molecular mechanisms of androgen action--a historical perspective. *Methods Mol Biol*. 2011;776:3-24.
3. Hashimoto Y, Miyachi H. Nuclear receptor antagonists designed based on the helix-folding inhibition hypothesis. *Bioorg Med Chem*. 2005;13(17):5080-5093.
4. Penning TM. Androgen biosynthesis in castration-resistant prostate cancer. *Endocr Relat Cancer*. 2014;21(4):T67-78.
5. Hara T, Miyazaki J, Araki H, Yamaoka M, Kanzaki N, Kusaka M, Miyamoto M. Novel mutations of androgen receptor: a possible mechanism of bicalutamide withdrawal syndrome. *Cancer Res*. 2003;63(1):149-153.
6. Brooke GN, Bevan CL. The role of androgen receptor mutations in prostate cancer progression. *Curr Genomics*. 2009;10(1):18-25.
7. Antonarakis ES, Lu C, Wang H, Luber B, Nakazawa M, Roeser JC, Chen Y, Mohammad TA, Fedor HL, Lotan TL, *et al*. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N Engl J Med*. 2014;371(11):1028-1038.
8. Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy GP. LNCaP model of human prostatic carcinoma. *Cancer Res*. 1983;43(4):1809-1818.
9. Korenchuk S, Lehr JE, MClean L, Lee YG, Whitney S, Vessella R, Lin DL, Pienta KJ. VCaP, a cell-based model system of human prostate cancer. *In Vivo*. 2001;15(2):163-168.
10. Dart DA, Kandil S, Tommasini-Ghelfi S, Serrano de Almeida G, Bevan CL, Jiang W, Westwell AD. Novel Trifluoromethylated Enobosarm Analogues with Potent Anti-androgenic Activity in vitro and Tissue Selectivity in vivo. *Mol Cancer Ther*. 2018.
11. Koushyar S, Economides G, Zaat S, Jiang W, Bevan CL, Dart DA. The prohibitin-repressive interaction with E2F1 is rapidly inhibited by androgen signalling in prostate cancer cells. *Oncogenesis*. 2017;6(5):e333.
12. Östling P, Leivonen SK, Aakula A, Kohonen P, Mäkelä R, Hagman Z, Edsjö A, Kangaspeska S, Edgren H, Nicorici D, *et al*. Systematic analysis of microRNAs targeting the androgen receptor in prostate cancer cells. *Cancer Res*. 2011;71(5):1956-1967.
13. Olsen JR, Azeem W, Hellem MR, Marvyin K, Hua Y, Qu Y, Li L, Lin B, Ke X, Øyan AM, *et al*. Context dependent regulatory patterns of the androgen receptor and androgen receptor target genes. *BMC Cancer*. 2016;16:377.
14. Rajan P, Sudbery IM, Villasevil ME, Mui E, Fleming J, Davis M, Ahmad I, Edwards J, Sansom OJ, Sims D, *et al*. Next-generation sequencing of advanced prostate cancer treated with androgen-deprivation therapy. *Eur Urol*. 2014;66(1):32-39.
15. Bader DA, Hartig SM, Putluri V, Foley C, Hamilton MP, Smith EA, Saha PK, Panigrahi A, Walker C, Zong L, *et al*. Mitochondrial pyruvate import is a metabolic vulnerability in androgen receptor-driven prostate cancer. *Nat Metab*. 2019;1(1):70-85.
16. Ho Y, Dehm SM. Androgen Receptor Rearrangement and Splicing Variants in Resistance to Endocrine Therapies in Prostate Cancer. *Endocrinology*. 2017;158(6):1533-1542.
17. Graham JR, Hendershott MC, Terragni J, Cooper GM. mRNA degradation plays a significant role in the program of gene expression regulated by phosphatidylinositol 3-kinase signaling. *Mol Cell Biol*. 2010;30(22):5295-5305.
18. da Silva MR, Moreira GA, Gonçalves da Silva RA, de Almeida Alves Barbosa É, Pais Siqueira R, Teixera RR, Almeida MR, Silva Júnior A, Fietto JL, Bressan GC. Splicing Regulators and Their Roles in Cancer Biology and Therapy. *Biomed Res Int*. 2015;2015:150514.
19. Jaagura M, Taal K, Koppel I, Tuvikene J, Timmusk T, Tamme R. Rat NEURL1 3'UTR is alternatively spliced and targets mRNA to dendrites. *Neurosci Lett*. 2016;635:71-76.

20. Matoulkova E, Michalova E, Vojtesek B, Hrstka R. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol*. 2012;9(5):563-576.
21. Yokoi S, Udagawa T, Fujioka Y, Honda D, Okado H, Watanabe H, Katsuno M, Ishigaki S, Sobue G. 3'UTR Length-Dependent Control of SynGAP Isoform α2 mRNA by FUS and ELAV-like Proteins Promotes Dendritic Spine Maturation and Cognitive Function. *Cell Rep*. 2017;20(13):3071-3084.
22. Andreassi C, Riccio A. To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol*. 2009;19(9):465-474.
23. Romanel A, Gasi Tandefelt D, Conteduca V, Jayaram A, Casiraghi N, Wetterskog D, Salvi S, Amadori D, Zafeiriou Z, Rescigno P, *et al*. Plasma AR and abiraterone-resistant prostate cancer. *Sci Transl Med*. 2015;7(312):312re310.
24. Takeuchi T, Okuno Y, Hattori-Kato M, Zaitsu M, Mikami K. Detection of AR-V7 mRNA in whole blood may not predict the effectiveness of novel endocrine drugs for castration-resistant prostate cancer. *Res Rep Urol*. 2016;8:21-25.

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Figure legends

Figure 1. AR transcript expression is increased with acquired AR-antagonists resistance.

A, MTT cytotoxicity assays of LNCaP, LNCaP^{BicR} and LNCaP^{EnzR} cells treated with increasing concentrations of either bicalutamide (left hand side) and enzalutamide (right hand side), at 0-100μM for 96hours. **B**, MTT cytotoxicity assays of VCaP, VCaP^{BicR} and VCaP^{EnzR} cells treated with increasing concentrations of either bicalutamide (left hand side) and enzalutamide (right hand side), at 0-100μM for 96hours. **C**, qPCR analysis of AR and PSA expression levels in four prostate cancer cell lines, as indicated. **D**, qPCR analysis of AR expression levels in parental and resistant prostate cancer cell lines as indicated. Data represents the mean and standard error from three independent replicates. Data is normalised to housekeeping genes GAPDH, RPL19 and β-actin. **E**, Western blot for AR levels from LNCaP and LNCaP$^{/Bic/Enz}$ cells, normalised to β-actin. Statistical analysis * (p =0.05) ** (p = 0.01) from a t-test.

Figure 2. AR transcripts differ in the 3'UTR region with acquired AR-antagonists resistance.

A, IGV screenshot of the AR transcripts detected in RNAseq samples from LNCaP and LNCaP^{BicR} cells. Upper panel indicated the genomic location of the AR intron and exon regions as well as the 5' and 3' UTR regions. **B**, Schematic diagram of the full 3' UTR region of AR (6.7kb), with the missing 3kb region indicated as a hatched box, and missing 3.1kb regions (lighter hatched box). The numbers represent the bases in the UTR region at which the missing region are spanned by RNAseq reads. **C**, Analysis results from the Spliceport splice analysis online tool. Left hand side indicates bases scoring highly as splice donor regions and right hand side shows bases scoring highly as splice acceptor regions.

Figure 3. AR 3'UTR variant increases due to hormonal starvation or AR-antagonists treatment.

A, Schematic diagram indicating the binding location of qPCR primer sets used for detecting either a portion of the 6.7kb 3'UTR of AR (upper panel) or the novel shorter AR 3kb transcript (lower panel). Numbers represent base locations in the 3'UTR. **B**, qPCR analysis of the AR UTR utilising the indicated primers sets in parental LNCaP or VCaP cells. **C**, qPCR analysis of AR$^{6.7kb\text{-}UTR}$ variant and AR$^{3.7kb\text{-}UTR}$ variant in parental and resistant LNCaP and VCaP clones. **D**, qPCR analysis of PSA (upper panel) and AR$^{3.7kb\text{-}UTR}$ (lower panel) in LNCaP cells grown in starvation medium for 72hours. **E**, Western blot for AR protein levels in hormonally starved LNCaP cells for 0-96 hours. Normalised to β-actin, with associated densitometry underneath. **F**, qPCR analysis of PSA (upper panel) and AR$^{3.7kb\text{-}UTR}$ variant (lower panel) in LNCaP cells treated with bicalutamide (10μM) for 6 days. Data represents the mean and standard error from three independent replicates. Data is normalised to housekeeping genes GAPDH, RPL19 and β-actin. Statistical analysis * (p =0.05) ** (p = 0.01) from a t-test.

Figure 4. Spliced AR 3'UTR variant is more stable in cells with acquired AR-antagonists resistance.

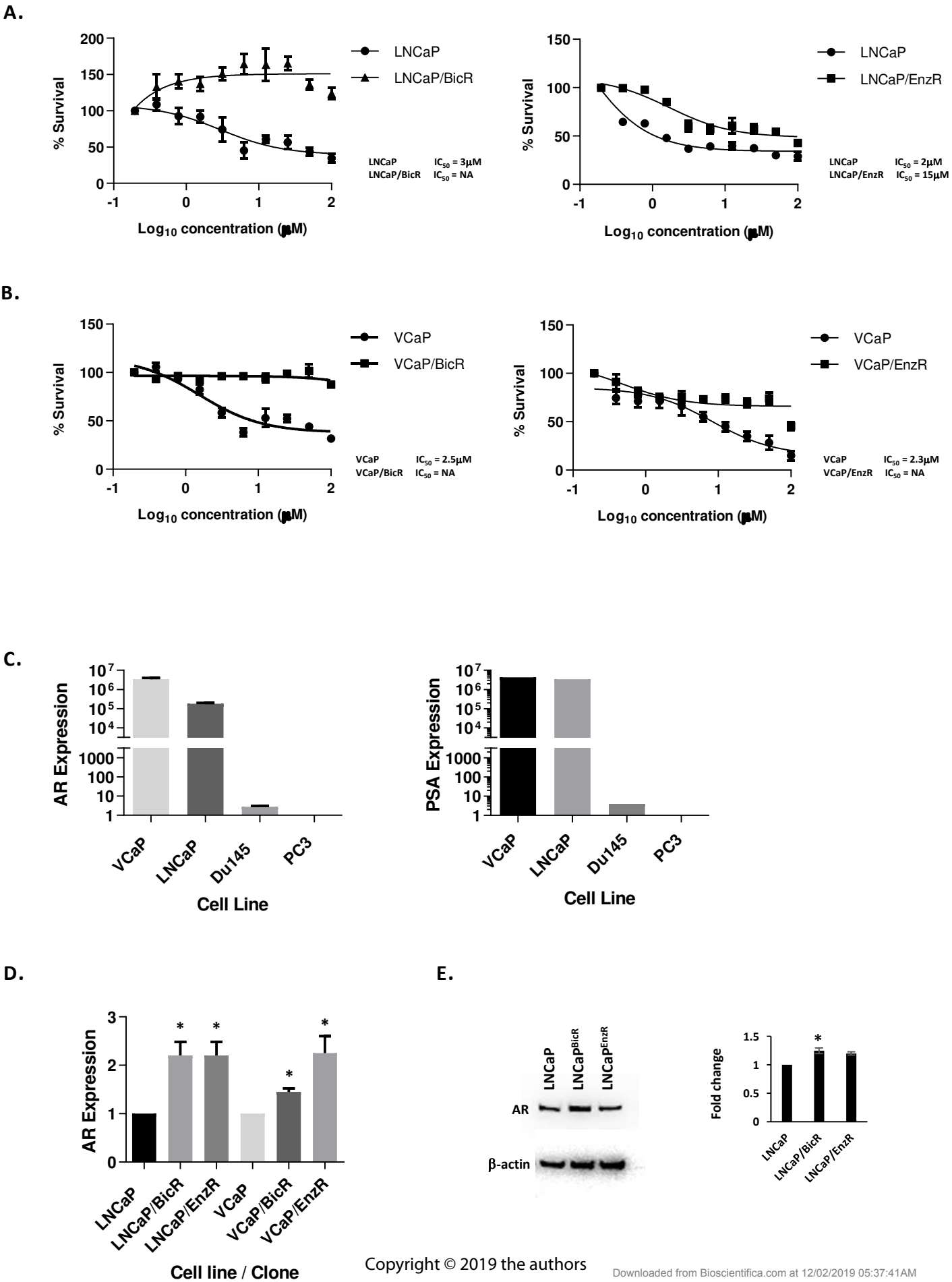
A, qPCR analysis of GAPDH, β-actin, AR$^{3.7kb\text{-}UTR}$ variant and AR$^{6.7kb\text{-}UTR}$ variant from LNCaP or LNCaP^{BicR} cells treated with actinomycin D (1μg/ml) for 0 – 24hours. Data represents the mean and standard error from three independent replicates. Data is normalised to housekeeping genes GAPDH, RPL19 and β-actin, and then expressed as a % of expression at time 0. **B**, Upper panel - schematic diagram

of the location of the Luciferase-AR UTR reporter fusion constructs in the 6.7kb region. Lower panel – luciferase activity of the reporter constructs in LNCaP or LNCaPBicR cells transfected with the firefly luciferase reporters and a constitutive renilla luciferase reporter. Data is expressed as firefly luciferase / renilla luciferase. Statistical analysis * (p =0.05) ** (p = 0.01) from a t-test.

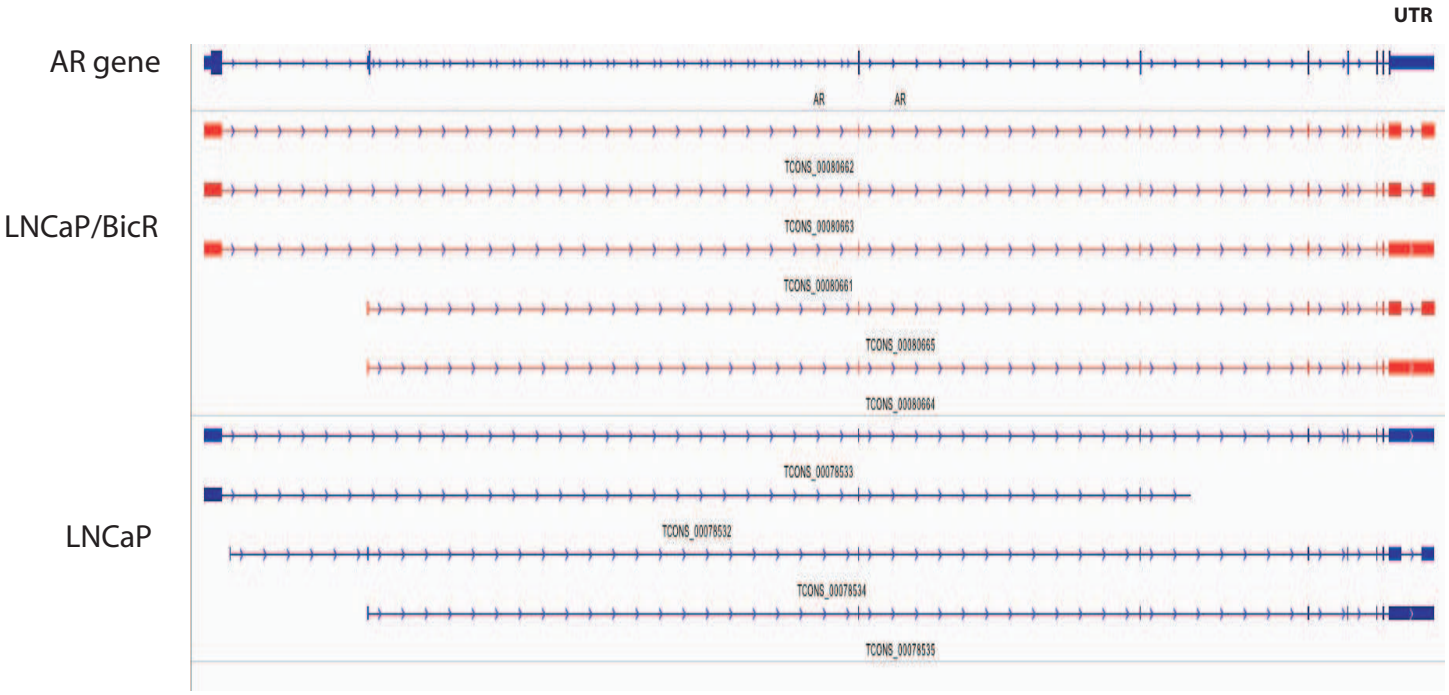
Figure 5. Spliced AR 3'UTR variant is detectable in prostate cancer patient serum.

A, Serum PSA measurements from the prostate cancer patient cohort - data from the Wales Cancer Bank. Data is separated according to Gleason grade 6-10, assigned to each patient at pathological diagnosis. **B**, qPCR analysis of the $AR^{3.7kb\text{-}UTR}$ variant from RNA extracted from patient serum samples. **C**, qPCR analysis of the AR coding region from RNA extracted from patient serum samples. **D**, qPCR analysis of the $AR^{6.7kb\text{-}UTR}$ variant from RNA extracted from patient serum samples. Data represents the mean and standard error from three independent measurements. Data is normalised to housekeeping genes GAPDH, RPL19 and β-actin. Statistical analysis * (p =0.05) ** (p = 0.01) from a 1-way Anova test (Kruskal Wallace test).

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

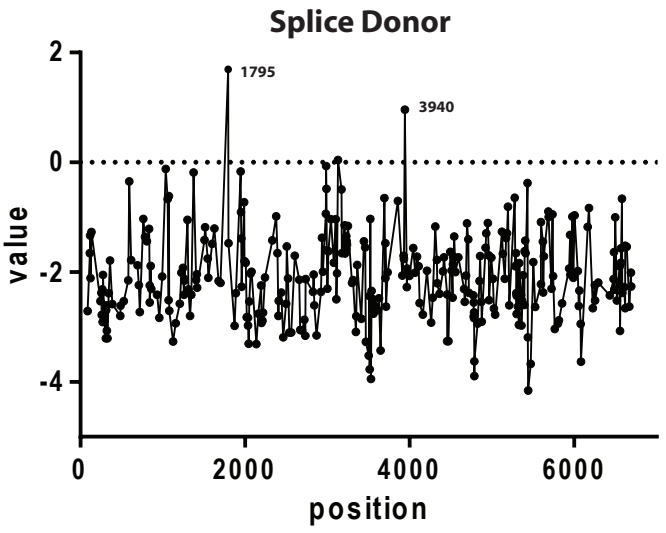


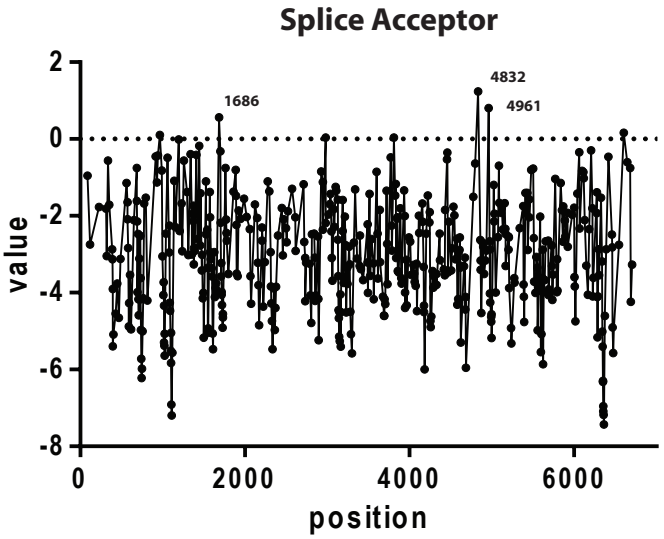
A.



B.

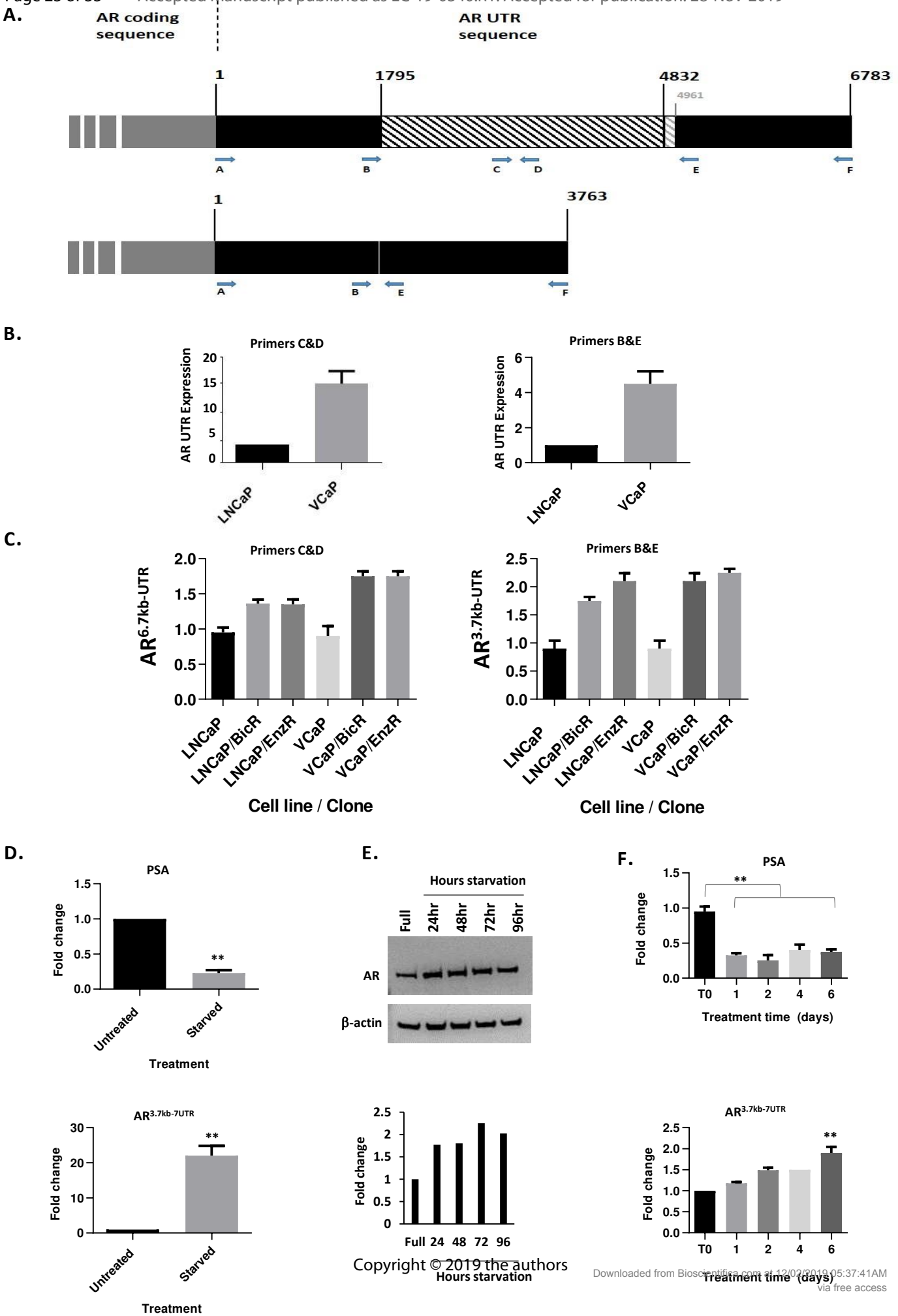
C.





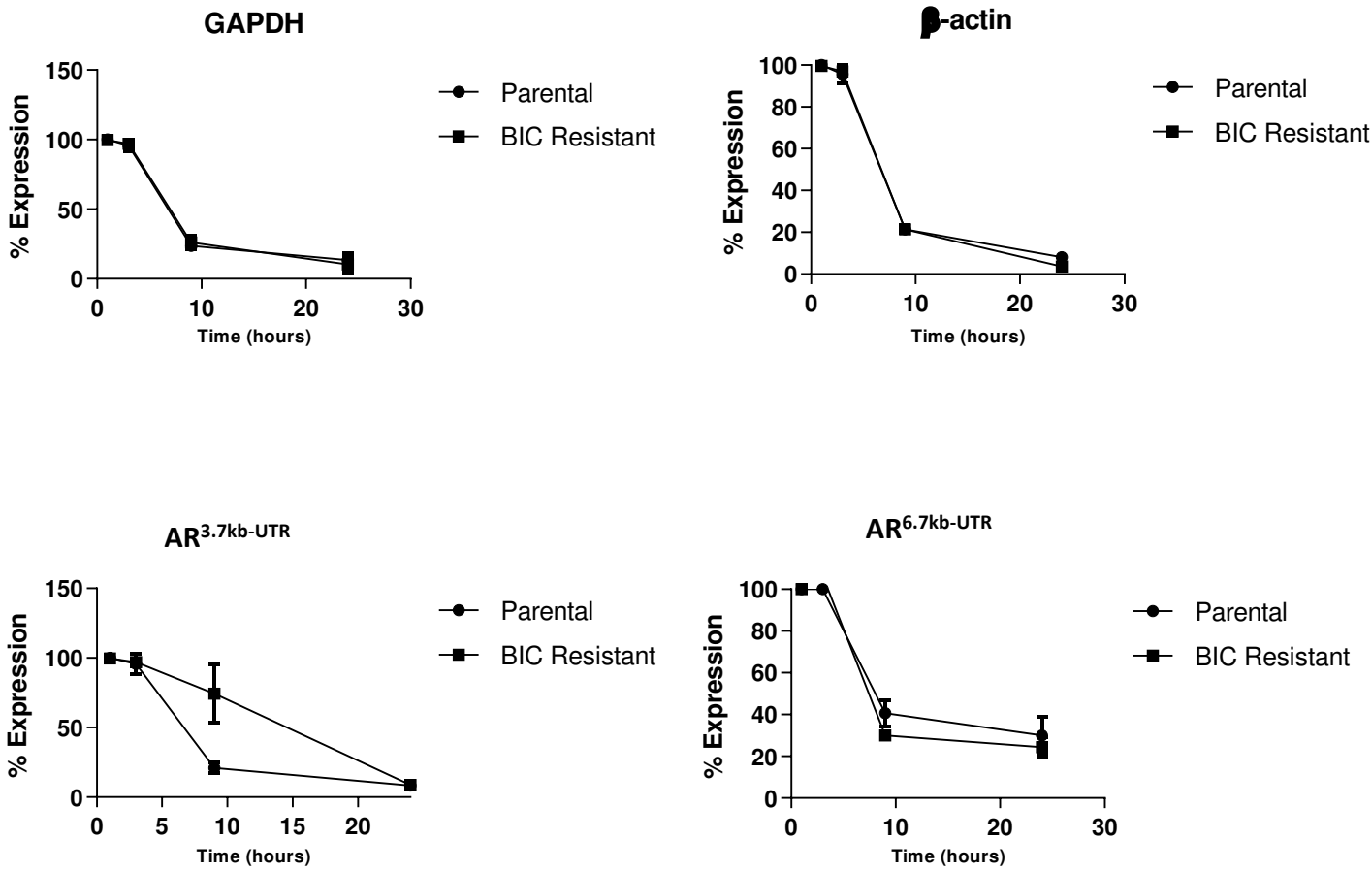
 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Figure 3.

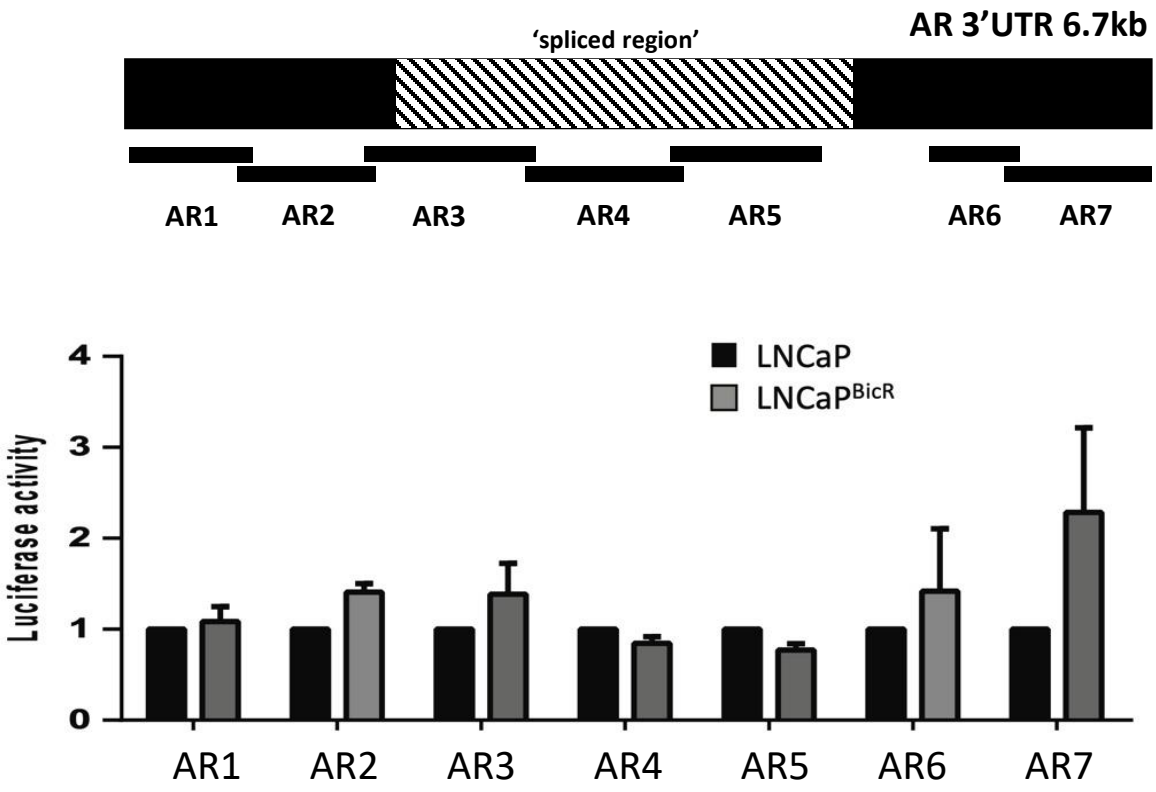


Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

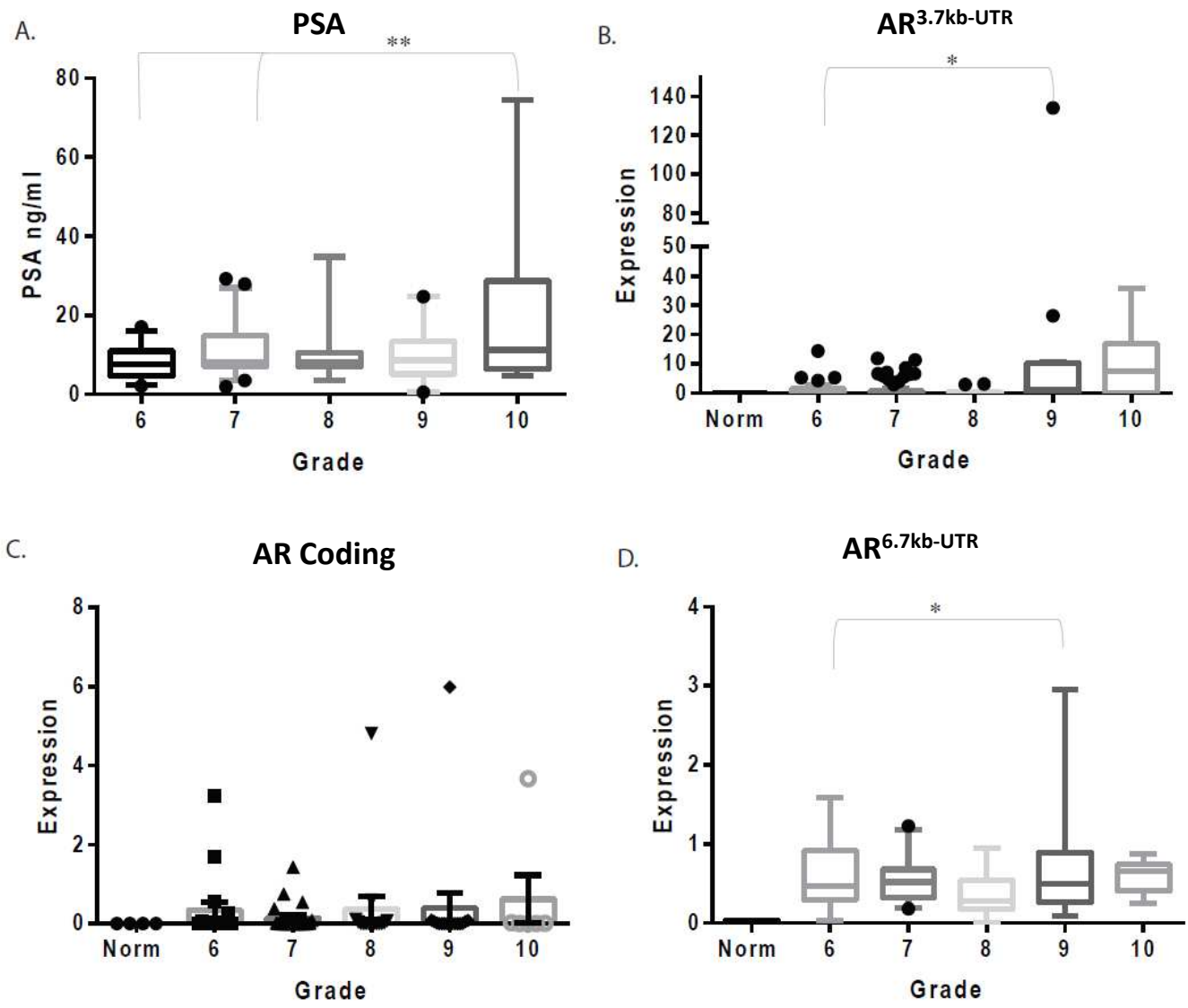
A.



B.

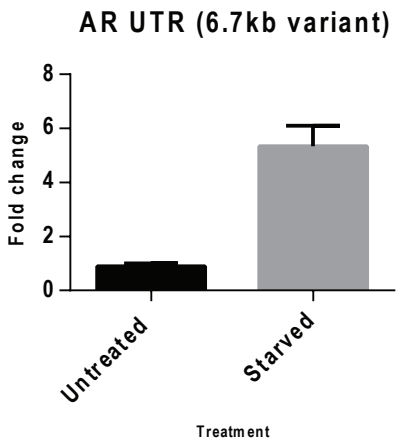


 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

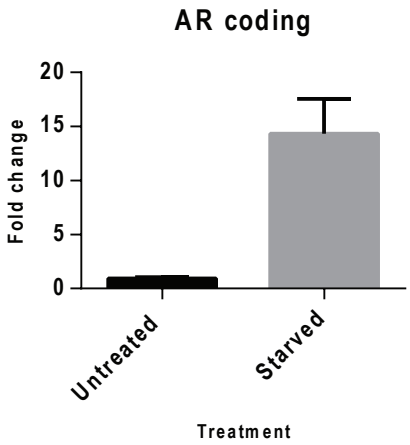


Supplemental figure 1.

A.

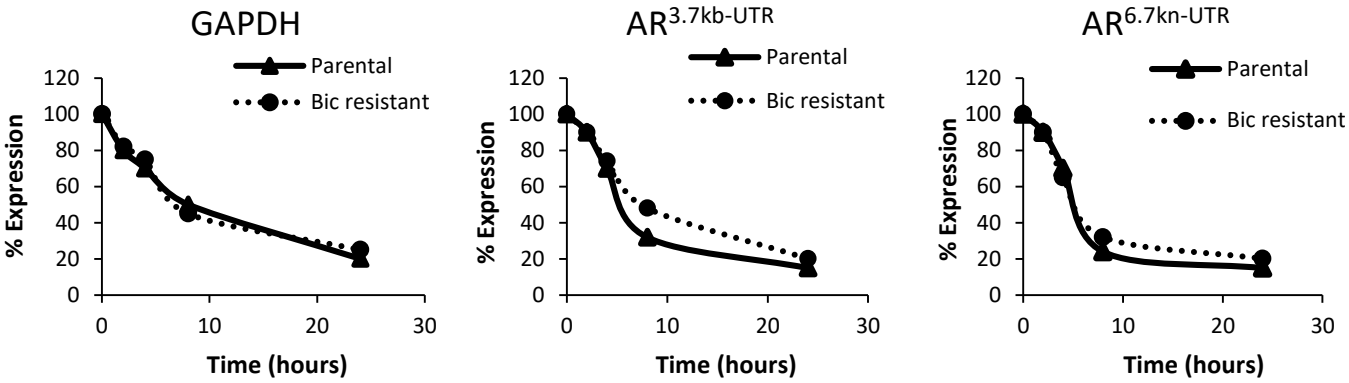


B.



Q-PCR analysis of A, AR UTR and B, AR coding region in LNCaP cells grown in stavation medium for 72 hours.
Data represents the mean and standard error from three independent replicates.
Data is normalised to housekeeping genes GAPDH, RPL19 and b-actin.

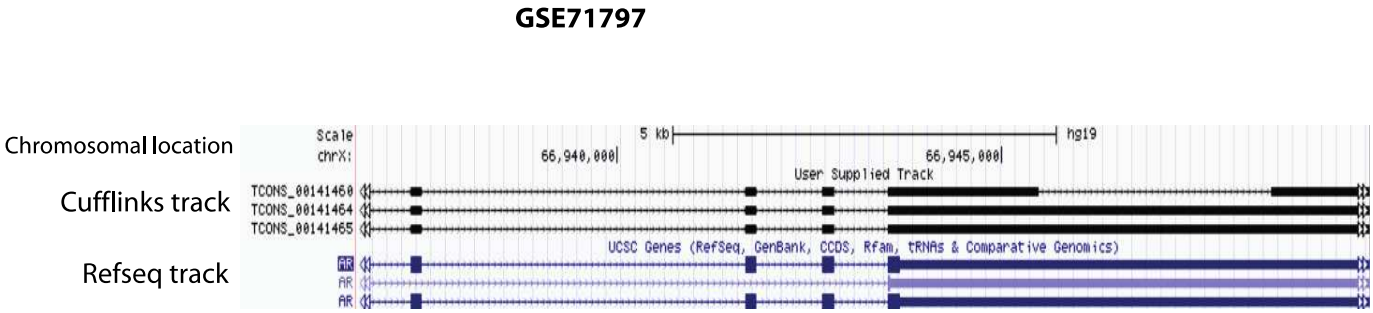
 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019



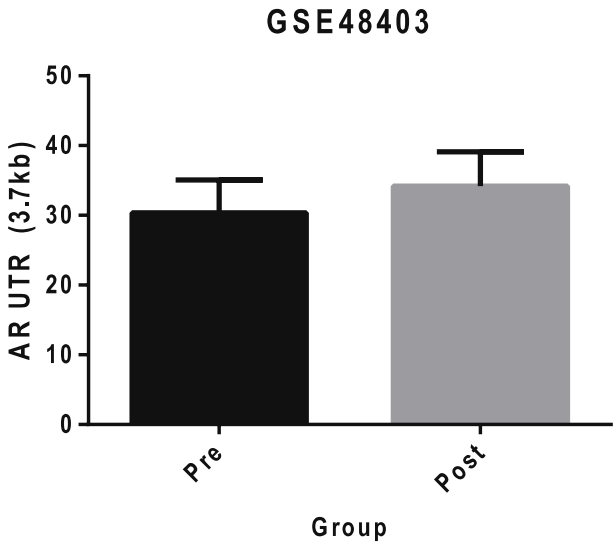
Q-PCR analysis of GAPDH, $AR^{3.7kb\text{-}UTR}$ and $AR^{6.7kb\text{-}UTR}$ from VCaP or $VCaP^{BicR}$ cells treated with actinomycin-D (1ug/ml) for 0-24hours. Data represents the mean and SE from three replicates. Data is normalised to the housekeeping genes GAPDH, RPL19, and β-actin, and then expressed as a % expression at time = 0.

Supplemental figure 3.

A.



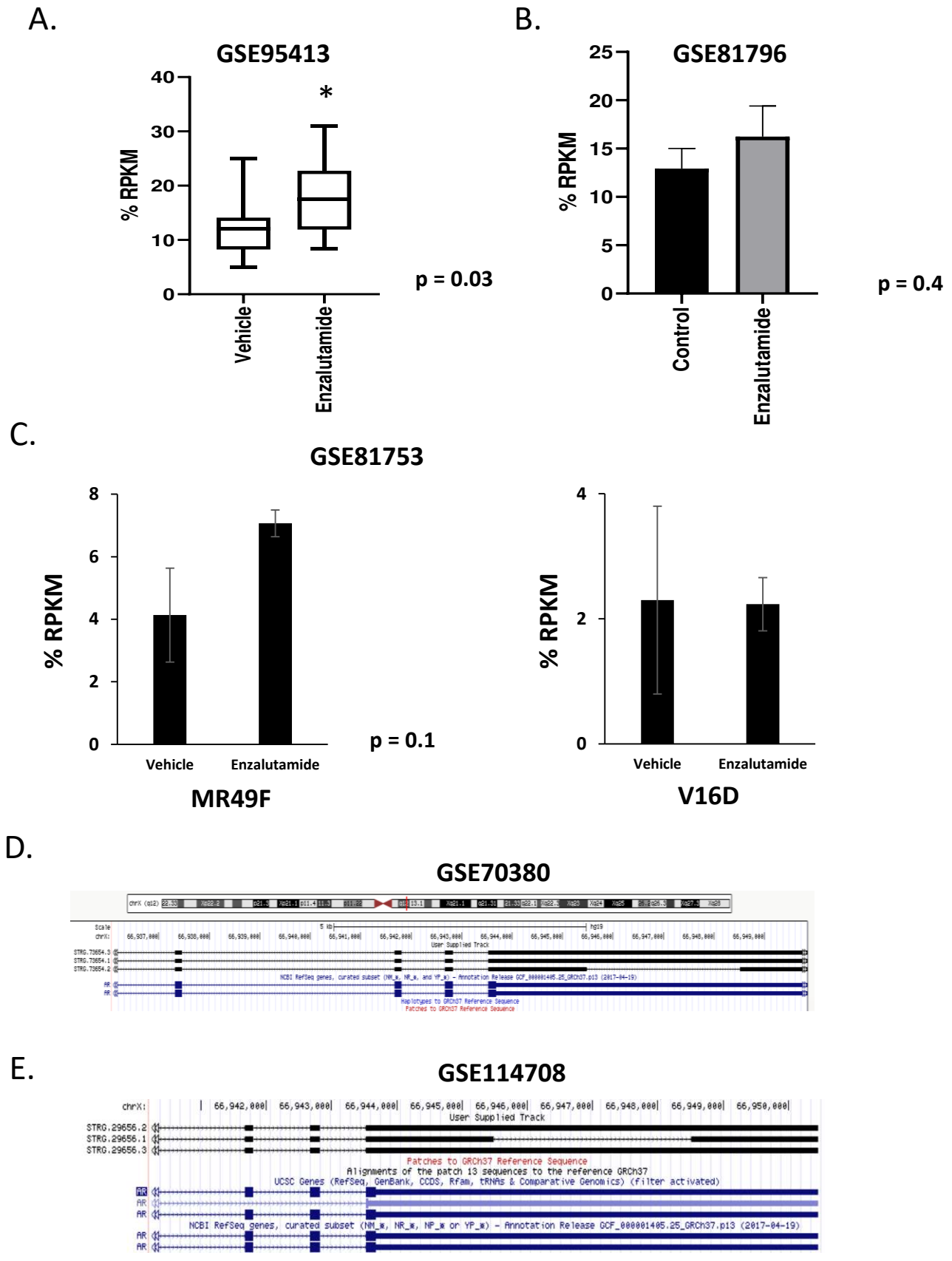
B.



A, Screen shot of the Cuffmerge data analysis on RNA-seq data analysis from GEO data set GSE71797, using UCSC Galaxy software. Data shows that the AR 3'UTR variant is present at low levels in hTERT immortalized prostate basal epithelial cells.
B, Analysis of the number of reads (RPKM) of the AR UTR (3.7kb variant) region from the GEO data set GSE48403, using UCSC Galaxy software. Data represents 6 pair of patients either pre or post androgen deprivation therapy.

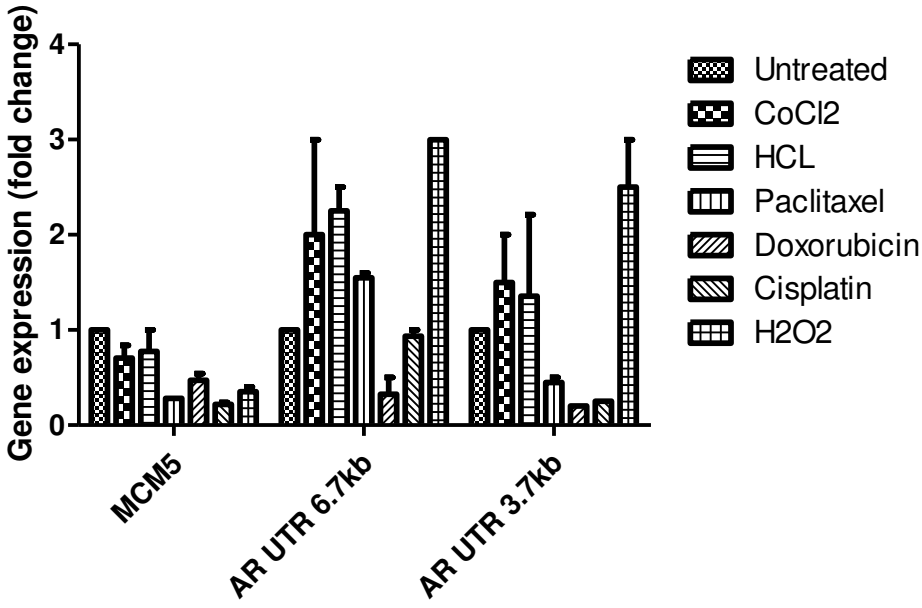
 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Supplemental figure 4.



A, Boxplot showing the analysis of the AR ΔUTR variant (in RPKM) from VCaP xenografts treated with enzalutamide. **B**, Graph showing analysis results of AR ΔUTR (RPKM) from C4-2 cells treated with enzalutamide. **C**, Graph showing analysis results of AR ΔUTR (RPKM) from the MR49F cell line – grown in castrated mice and enzalutamide treated. **D**, Screenshot from Cuffmerge data analysis from metastatic deposits of a patients with disease progression. **E**, Screenshot from Cuffmerge data analysis from LNCaP-abl cells.

Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019



Q-PCR analysis of LNCaP cells treated with various conditions including hypoxia mimic $CoCl_2$ (100μM), HCl (10mM), Paclitaxel (1nM), Doxorubicin (1μM), Cisplatin (3μM), H_2O_2 (0.1mM) for 24 hours. Data is normalised to GAPDH, β-actin and RPL19.

Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

AR mRNA full UTR sequence

agcattggaaaccctatttcccaccccagctcatgccccctttcagatgtcttctgcctgttataactctgcactactcctctgcagtgccttggggaatttcctcta
ttgatgtacagtctgtcatgaacatgttcctgaattctatttgctgggctttttttttctctttctctccttttcttttttcttcttccctccctatctaaccctcccatg
gcaccttcagactttgcttcccattgtggctcctatctgtgttttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcccagtgtcaagttgtgctt
gtttacagcactactctgtgccagccacacaaacgtttacttatcttatgccacgggaagtttagagagctaagattatctggggaaatcaaaacaaaaacaagcaaac
aaaaaaaaaaagcaaaaacaaaacaaaaaataagccaaaaaacctgctagtgttttttcctcaaaaataaataaataaataaataaatacgtacatacatacacacat
acatacaaacatatagaaatccccaaagaggccaatagtgacgagaaggtgaaaattgcaggcccatggggagttactgatttttttcatctcctccctccacgggagac
tttattttctgccaatggctattgccattagagggcagagtgaccccagagctgagttgggcaggggggtggacagagaggagaggacaaggagggcaatggagcatca
gtacctgcccacagccttggtccctgggggctagactgctcaactgtggagcaattcattatactgaaaatgtgcttgttgttgaaaatttgtctgcatgttaatgcct
caccccaaaccctttttctctctcactctctgcctccaacttcagattgactttcaatagtttttctaagacctttgaactgaatgttctcttcagccaaaacttggcg
acttccacagaaaagtctgaccactgagaagaaggagagcagagatttaacccttttgtaaggccccatttggatccaggtctgctttctcatgtgtgagtcagggagga
gctggagccagaggagaagaaaatgatagcttggctgttctcctgcttaggacactgactgaatagttaaactctcactgccactacctttttccccacctttaaaagac
ctgaatgaagttttctgccaaactccgtgaagccacaagcaccttatgtcctcccttcagtgttttgtgggcctgaatttcatcacactgcatttcagccatggtcatc
aagcctgtttgcttcttttgggcatgttcacagattctctgttaagagcccccaccaccaagaaggttagcaggccaacagctctgacatctatctgtagatgccagta
gtcacaaagatttcttaccaactctcagatcgctggagcccttagacaaactggaaagaaggcatcaaagggatcaggcaagctgggcgtcttgccctttgtcccccaga
gatgataccctcccagcaagtggagaagttctcacttccttctttagagcagctaaaggggctacccagatcagggttgaagagaaaactcaattaccagggtgggaag
aatgaaggcactagaaccagaaaccctgcaaatgctcttcttgtcacccagcatatccacctgcagaagtcatgagaagagagaaggaacaaagaggagactctgacta
ctgaattaaaatcttcagcggcaaagcctaaagccagatggacaccatctggtgagtttactcatcatcctcctctgctgctgattctgggctctgacattgcccatac
tcactcagattccccacctttgttgctgcctcttagtcagagggaggccaaaccattgagactttctacagaaccatggcttctttcggaaaggtctggttggtgtggc
tccaatactttgccacccatgaactcagggtgtgccctgggacactggttttatatagtcttttggcacacctgtgttctgttgacttcgttcttcaagcccaagtgca
agggaaaatgtccacctactttctcatcttggcctctgcctccttacttagctcttaatctcatctgttgaactcaagaaatcaagggccagtcatcaagctgcccatt
ttaattgattcactctgtttgttgagaggatagtttctgagtgacatgatatgatccacaagggtttccttccctgatttctgcattgatatttaatagccaaacgaact
tcaaaacagctttaaataacaagggagaggggaacctaagatgagtaatatgccaatccaagactgctggagaaaactaaagctgacaggttccctttttggggtggga
tagacatgttctggttttctttattattacacaatctggctcatgtacaggatcacttttagctgttttaaacagaaaaaaatatccaccactcttttcagttacacta
ggttacattttaataggtcctttacatctgttttggaatgattttcatctttttgtgatacacagattgaattatatcattttcatatctctccttgtaaatactagaag
ctctcctttacatttctctatcaaatttttcatctttatgggtttcccaattgtgactcttgtcttcatgaatatatgtttttcatttgcaaaagccaaaaatcagtga
aacagcagtgtaattaaaagcaacaactggattactccaaatttccaaatgacaaaactagggaaaaatagcctacacaagcctttaggcctactctttctgtgcttgg
gtttgagtgaacaaaggagattttagcttggctctgttctcccatggatgaaaggaggaggatttttttttttctttggccattgatgttctagccaatgtaattgaca
gaagtctcattttgcatgcgctctgctctacaaacagagttggtatggttggtatactgtactcacctgtgagggactggccactcagacccacttagctggtgagcta
gaagatgaggatcactcactggaaaagtcacaaggaccatctccaaacaagttggcagtgctcgatgtggacgaagagtgaggaagagaaaaagaaggagcaccaggga
gaaggctccgtctgtgctgggcagcagacagctgccaggatcacgaactctgtagtcaaagaaaagagtcgtgtggcagtttcagctctcgttcattgggcagctcgcc
taggcccagcctctgagctgacatgggagttgttggattctttgtttcatagctttttctatgccataggcaatattgttgttcttggaaagtttattatttttttaac
tcccttactctgagaaagggatattttgaaggactgtcatatatctttgaaaaaagaaaatctgtaatacatatatttttatgtatgttcactggcactaaaaaatata
gagagcttcattctgtccttgggtagttgctgaggtaattgtccaggttgaaaaataatgtgctgatgctagagtccctctctgtccatactctacttctaaatacat
ataggcatacatagcaagttttatttgacttgtactttaagagaaaatatgtccaccatccacatgatgcacaaatgagctaacattgagcttcaagtagcttctaagt
gtttgtttcattaggcacagcacagatgtggccttttccccccttctctccctgatatctggcagggcataaaggcccaggccacttcctctgccccttcccagccctg
caccaaagctgcatttcaggagactctctccagacagcccagtaactacccgagcatggcccctgcatagccctggaaaaataagaggctgactgtctacgaattatct
tgtgccagttgcccaggtgagagggcactgggccaagggagtggttttcatgtttgacccactacaaggggtcatgggaatcaggaatgccaaagcaccagatcaaatc
caaaacttaaagtcaaaataagccattcagcatgttcagtttcttggaaaaggaagtttctacccctgatgcctttgtaggcagatctgttctcaccattaatctttttt
gaaaatcttttaaagcagtttttaaaaagagagatgaaagcatcacattatataaccaaagattacattgtacctgctaagataccaaaattcataagggcaggggggg
agcaagcattagtgcctctttgataagctgtccaaagacagactaaaggactctgctggtgactgacttataagagctttgtgggttttttttccctaataatatata
tgtttagaagaattgaaaataatttcgggaaaatgggattatgggtccttcactaagtgattttataagcagaactggctttccttttctctagtagttgctgagcaaa
ttgttgaagctccatcattgcatggttggaaatggagctgttcttagccactgtgtttgctagtgcccatgttagcttatctgaagatgtgaaacccttgctgataagg
gagcatttaaagtactagattttgcactagagggacagcaggcagaaatccttatttctgcccactttggatggcacaaaaagttatctgcagttgaaggcagaaagtt
gaaatacattgtaaatgaatatttgtatccatgtttcaaaattgaaatatatatatatatatatatatatatatatatatatatatatatatagtgtgtgtgtgttctgatag
ctttaactttctctgcatctttatatttggttccagatcacacctgatgccatgtacttgtgagagaggatgcagttttgttttggaagctctctcagaacaaacaaga
cacctggattgatcagttaactaaaagttttctcccctattgggtttgacccacaggtcctgtgaaggagcagaggataaaaagagtagaggacatgatacattgtac
tttactagttcaagacagatgaatgtggaaagcataaaaactcaatggaactgactgagatttaccacagggaaggcccaaacttggggccaaaagcctacccaagtga
ttgaccagtggccccctaatgggacctgagctgttggaagaagagaactgttccttggtcttcaccatccttgtgagagaagggcagtttcctgcattggaacctggag
caagcgctctatctttcacacaaattccctcacctgagattgaggtgctcttgttactgggtgtctgtgtgctgtaattctggttttggatatgttctgtaaagatttt
gacaaatgaaaatgtgtttttctctgttaaaacttgtcagagtactagaagttgtatctctgtaggtgcaggtccatttctgcccacaggtagggtgttttttctttgat
taagagattgacacttctgttgcctaggacctcccaactcaaccatttctaggtgaaggcagaaaaatccacattagttactcctcttcagacatttcagctgagataa
caaatcttttggaatttttcacccatagaaagagtggtagatatttgaatttagcaggtggagtttcatagtaaaaacagcttttgactcagctttgatttatcctca
tttgatttggccagaaagtaggtaatatgcattgattggcttctgattccaattcagtatagcaaggtgctaggtttttttcctttccccacctgtctcttagcctgggg
aattaaatgagaagccttagaatgggtggccccttgtgacctgaaacacttcccacataagctacttaacaagattgtcatggagctgcagattccattgcccaccaaag
actagaacacacacatatccatacaccaaaggaaagacaattctgaaatgctgtttctctggtggttccctctctggctgctgcctcacagtatgggaacctgtactct
gcagaggtgacaggccagatttgcattatctcacaaccttagcccttggtgctaactgtcctacagtgaagtgcctgggggttgtcctatcccataagccacttggat
gctgacagcagccaccatcagaatgacccacgcaaaaaaaagaaaaaaaaaattaaaaagtcccctcacaacccagtgacacctttctgctttcctctagactggaaca
ttgattagggagtgcctcagacatgacattcttgtgctgtccttggaattaatctggcagcaggagggagcagactatgtaaacagagataaaaattaattttcaatat
tgaaggaaaaaagaaataagaagagagagagaaagaaagcatcacacaaagattttcttaaaagaaacaattttgcttgaaatctctttagatggggctcatttctcac
ggtggcacttggcctccactgggcagcaggaccagctccaagcgctagtgttctgttctcttttgtaatcttggaatcttttgttgctctaaatacaattaaaaatgg
cagaaacttgtttgttggactacatgtgtgactttgggtctgtctctgcctctgctttcagaaatgtcatccattgtgtaaaatattggcttactggtctgccagctaa
aacttggccacatcccctgttatggctgcaggatcgagttattgttaacaaagagacccaagaaaagctgctaatgtcctcttatcattgttgttaatttgttaaaacc
taaagaaatctaaaatttcaaaaaa

Text indicated the full UTR sequence of AR starting immediately after the STOP codon (not shown). Yellow highlighted text indicates the missing 3kb region found in the AR$^{3.7kn\text{-}UTR}$ variant. Red highlighted text indicates the location of the primers for q-PCR (in the order of A, B, C, D, E, F.

Supplemental Table 1. Numbers and Gleason grades of prostate cancer patient serum samples.

Gleason Grade	**Numbers of samples**
6	28
7	54
8	17
9	19
10	6
Healthy volunteers	8

 Accepted Manuscript published as EC-19-0340.R1. Accepted for publication: 28-Nov-2019

Supplemental table 2

Primer Name	ID or Sequence
b-actin	Hs01060665_g1
RPL19	Hs02338565_gH
GAPDH for	5'-AGCCACATCGCTCAGACAC-3'
GAPDH rev	5'-GCCCAATACGACCAAATCC-3'
PSA	Hs02576345_m1
AR	Hs00171172_m1
A	5'- AGCATTGGAAACCCTATTTC-3'
B	5'-CAGCGGCAAAGCCTAAA-3'
C	5'-GAGAGGGGAACCTAAGAT-3'
D	5'-AACTGAAAAGAGTGGTGGATA-3'
E	5'-CTCTGCTCCTTCACAGGAC-3'
F	5'-GTTTTAACAAATTAACAACAATGATA-3'