Investigating the Role of Constitutively Active Truncated Androgen Receptor Splice Variants in Prostate Cancer

Peter Eoin Lonergan

A thesis submitted to the University of Dublin

for the degree of

Doctor in Medicine

Based on research performed at

Department of Urology

Mayo Clinic

200 First Street SW

Rochester, Minnesota

USA

July 2012

I. Declaration

This thesis is submitted by the undersigned for the degree of Doctor in Medicine at the University of Dublin. It has not been submitted previously as an exercise for a degree at this or any other University. I declare that this thesis is entirely my own work, except where acknowledged in the text. I give permission to the Library to lend or copy this thesis upon request.

Peter E. Lonergan

II. Summary

The androgen receptor (AR) is fundamental for the growth and survival of normal and malignant prostate cells. Therefore, androgen deprivation therapy remains the first-line treatment for disseminated disease, however, relapse and progression to a castration-resistant phenotype for which no durable treatment currently exists, is inevitable. Unequivocal evidence now exists that restored AR activity is fundamental in the progression to castration-resistant prostate cancer (CRPC). Multiple mechanisms by which AR is reactivated under androgen-depleted conditions may be involved in the development of this lethal phenotype. Recent studies have identified alternatively spliced transcripts encoding truncated AR isoforms that lack the ligand-binding domain which is the therapeutic target of androgen deprivation therapy. Many of these truncated AR variants function as constitutively active, ligand-independent transcription factors that can support androgen-independent expression of AR target genes, as well as ligand-independent growth of prostate cancer cells.

In this study, an AR variant gene signature independent of full-length AR consisting of 284 genes was identified in the 22Rv1 cell-based model of CRPC using whole-genome microarray. The clinical relevance of this AR variant gene signature was also determined in a cohort of prostate cancer specimens with the AR variant gene signature sufficient to distinguish between microdissected benign and malignant prostate cancer specimens. Furthermore, a number of putative AR variant regulated pro-oncogenes (TAX1BP1, NRP1, LEF1, LAMP3, IGFBP3) and tumor suppressor genes (SPRY2, RASD1, MBP, BNIP3L), upregulated and downregulated respectively with disease progression and biochemical recurrence were identified. These findings may provide a critical insight into the mechanism of disease progression mediated by truncated AR variants and reveal potential targets for therapeutic intervention in patients who progress despite androgen deprivation therapy.

III. Acknowledgements

I would like to express my sincere thanks to Dr. Donald Tindall for his constant guidance and mentorship over the last two years. I would also like to thank my co-supervisor at Trinity College, Dublin, Professor Marina Lynch for her support during this endeavour.

The opportunity to come to Mayo Clinic would not have been possible without the generous financial support of the US Department of State through a Fulbright Scholarship and also Trinity College, Dublin through the Postgraduate Travelling Scholarship in Medicine.

I would like to thank my colleagues in the laboratory, both past and present, for making it a stimulating and exciting environment to work in: Dr. Lucas Nacusi, Lucy Schmidt, Kevin Regan, Dr. Stephen Boorjian, Dr. Ji Lu, Dr. Kristin Lamont, Dr. Kaustubh Datta, Dr. Zheng Fu, Dr. Diping Wang, Dr. Henry Zhang and Marlene Finley. I am indebted to Dr. Zhifu Sun of the Division of Biomedical Statistics & Informatics who provided the statistical support for this project.

Finally, my family and friends in Dublin and also the many new friends I have made in Rochester deserve special thanks for their support.

IV. Table of Contents

V. List of Figures

VI. List of Tables

VII. List of Abbreviations

The following abbreviations are used:

ADT	Androgen deprivation therapy
AF	Activation function
AIS	Androgen insensitivity syndrome
AR	Androgen receptor
ARE	Androgen response element
BPH	Benign prostatic hyperplasia
BSA	Bovine serum albumin
CAIS	Complete androgen insensitivity syndrome
cDNA	Complimentary deoxyribonucleic acid
CE	Cryptic exon
CRPC	Castration-resistant prostate cancer
CSS	Charcoal stripped serum
CTD	C-terminal domain
DBD	DNA-binding domain
DHEA	Dehydroepiandrosterone
DHT	Dihydrotestosterone
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ER	Estrogen receptor
FBS	Fetal bovine serum
GO	Gene ontology
GP	Gleason pattern
GS	Gleason score

HA	Hemagglutinin
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase
HSD	Hydroxysteroid dehydrogenase
HSP	Heat shock protein
LBD	Ligand-binding domain
LCM	Laser capture microdissection
LHRH	Luteinizing hormone-releasing hormone
LN	Lymph node
MOPS	3-(N-morpholino) propanesulfonic acid
mRNA	Messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
NLS	Nuclear localization signal
NTD	N-terminal domain
PAGE	Polyacrylamide gel electrophoresis
PAIS	Partial androgen insensitivity syndrome
PBS	Phosphate buffered saline
PCa	Prostate cancer
PMSF	Phenylmethylsulphonyl fluoride
PSA	Prostate specific antigen
RACE	Rapid amplification of cDNA ends
RT-PCR	Reverse transcription polymerase chain reaction
SCID	Severe combined immunodeficiency
SDS	Sodium dodecyl sulphate
shRNA	Short hairpin ribonucleic acid
siRNA	Small interfering ribonucleic acid
SLASR	Selective linear amplification of sense RNA
TAU	Transactivation unit
TBS	Tris buffered saline
TBS-T	Tris buffered saline-tween

Small organs have often engaged the attention of great men

Dr. Charles Horace Mayo (1865-1939)

Chapter 1

Introduction

1.1 Prostate cancer incidence and progression

Prostate cancer (PCa) is the most frequently diagnosed non-cutaneous malignancy in men and the second leading cause of male cancer-related mortality in the United States with an estimated 241,740 new diagnoses and 28,170 mortalities annually (Siegel *et al.*, 2012). In Ireland, PCa is also the most frequently diagnosed non-cutaneous malignancy in men at 29.7%, however, it is the third leading cause of male cancer-related mortality at 11.8% of all cancer-related deaths in men (Cancer in Ireland 2011: Annual report of the National Cancer Registry).

Clinically localized PCa is managed primarily through surgery or radiation therapy (Klein *et al.*, 2009). For patients who recur systemically following definitive treatment or who present with locally advanced or metastatic disease, the mainstay of treatment is androgen-deprivation therapy (ADT) typically with a luteinizing-hormone-releasing hormone (LHRH) agonist (Loblaw *et al.*, 2007). Progression-free and overall survival figures of patients with metastatic disease with various methods of ADT have ranged from 12-20 months and 24-36 months, respectively (Leuprolide Study Group, 1984; Crawford *et al.*, 1989; Denis *et al.*, 1993; Eisenberger *et al.*, 1998). Progression to this lethal disease phenotype, heralded by rising serum prostate-specific antigen (PSA), increasing tumor size, new metastatic spread and disease-related symptoms is referred to as castration-resistant prostate cancer (CRPC; Scher *et al.*, 2008).

1.2 Overview of AR structure and function

The human AR gene is a nuclear transcription factor and a member of the steroid hormone receptor superfamily of genes. It is located on the X chromosome (q11-12) and consists of 8 exons. It codes for a protein of 920 amino acids with a mass of 110 kDa. The structural organization of the AR gene, mature spliced mRNA and protein domains are similar to the family members estrogen receptor α (ERα), ERβ and progesterone receptor (Chang *et al.*, 1988; Lubahn *et al.*, 1989). The AR consists of four structurally and functionally distinct domains: a poorly conserved N-terminal domain (NTD), a highly conserved DNA-binding domain (DBD) and a moderately conserved ligand-binding domain (LBD). A short amino acid sequence called the

'hinge region' separates the LBD from the DBD and also contains part of a bipartite ligand-dependent nuclear localization signal (NLS) for AR nuclear transport (Figure 1.1; Gelmann, 2002; Heinlein and Chang, 2004).

1.2.1 AR Exon 1

AR exon 1 encodes the entire NTD, which represents almost 60% of the AR protein but is variable in length due to polymorphic (CAG)n and (GGN)n repeat units encoding polyglutamine and polyglycine tracts respectively (Ferro *et al.*, 2002; Ding *et al.*, 2004, 2005). The AR NTD is a potent transcriptional activator and can activate transcription independently of androgenic stimulus in LBD deletion mutants (Jenster *et al.*, 1991; Simental *et al.*, 1991). This activity has been mapped to two primary transactivation domains termed transactivation unit-1 (TAU-1) and TAU-5. These two domains have been shown to be critical for full AR transcriptional activity (Callewaert *et al.*, 2006). TAU-1 activity has been more precisely mapped to a discrete ^{178}LKDIL^{182} motif (Chamberlain *et al.*, 1996; Callewaert *et al.*, 2006). However, TAU-5 is responsible for the majority of constitutive transcriptional activity within the NTD, mediated through the core sequence ^{435}WHTLF^{439}, accounting for approximately 50% aberrant AR activity in CRPC cells (Dehm *et al.*, 2007).

1.2.2 AR Exons 2 – 3

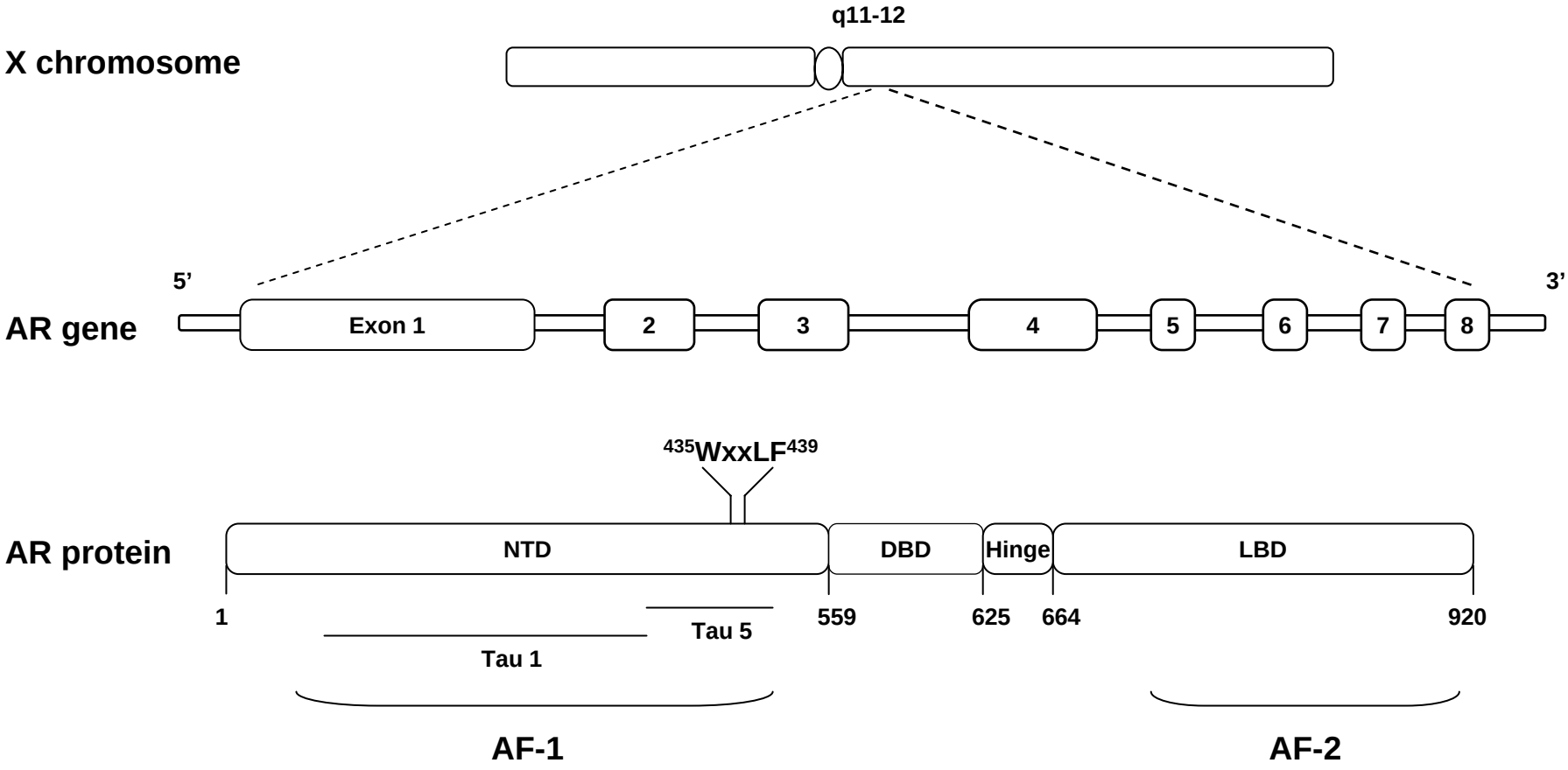
AR exon 2 encodes the first zinc finger in the AR DBD, which is the DNA recognition helix that makes contact with the major groove residues in an androgen-response element (ARE) half-site (Shaffer *et al.*, 2004). The second zinc finger is encoded by AR exon 3 which is the dimerization interface that mediates AR homodimer formation (Shaffer *et al.*, 2004).

1.2.3 AR Exons 4 – 8

AR exons 4 – 8 encode a short 50 amino acid flexible hinge and 11 α-helices folded into an α-helical sandwich to form the well-characterized AR C-terminal domain (CTD), which contains the AR LBD and transcriptional activation function-2 (AF-2) coregulator binding surface (Matias *et al.*, 2000; Sack *et al.*, 2001; Hur *et al.*,

Figure 1.1 Androgen receptor structure

Schematic representation of the full-length AR gene locus with indications of specific motifs and domains (depictions are not to scale).


X chromosome
q11-12
AR gene
5’
3’
Exon 1
2
3
4
5
6
7
8
AR protein
435WxxLF439
NTD
DBD
Hinge
LBD
1
559
625
664
920
Tau 5
Tau 1
AF-1
AF-2

2004; Estebanez-Perpina *et al.*, 2005; Bain *et al.*, 2007). In the absence of ligand, the AR protein is localized to the cytoplasm, where it associates with heat shock protein (HSP)-90, other molecular chaperones and high-molecular-weight immunophilins by virtue of interaction with the CTD (Pratt and Toft, 1997; Smith and Toft, 2008). Androgen binding induces a conformational change in the AR, which exposes the bipartite NLS in the hinge region (Zhou *et al.*, 1994), thus allowing direct interaction with importin-α and subsequent nuclear localization through the nuclear pore complex (Cutress *et al.*, 2008). In the nucleus, AR binds as a dimer to AREs in the promoter and enhancer elements of target genes, the best-characterized of which are PSA, TMPRSS2, NKX3.1 and hK2 (Dehm and Tindall, 2006). Transcriptional activation of these target genes is a complex, multistep process that requires ordered-stepwise recruitment of a plethora of coregulatory proteins (Heinlein and Chang, 2004; Heemers and Tindall, 2007). Although the AF-2 domain is able to recruit well-characterized coactivators such as SRC-1, SRC-2/TIF-2 and SRC-3/AIB1, its contribution to transcriptional activity is relatively weak, compared to the AF-1 domain within the NTD (He *et al.*, 2004).

1.3 Targeting AR in prostate cancer

Observations from as early as the 1780s noted that prostatic epithelium undergoes atrophy following castration (Hunter, 1840), however, it was the pioneering work of Huggins and Hodges in the 1940s which demonstrated that benign prostatic epithelium and prostate adenocarcinoma were biochemically analogous and respond in a similar fashion to androgen deprivation (Huggins and Hodges, 1941; Huggins *et al.*, 1941; Huggins, 1942). This principle of androgen deprivation is the standard-of-care for the clinical management of disseminated PCa. Currently, suppression of testicular androgen synthesis (which accounts for ~90% of serum testosterone) can be effectively achieved with bilateral orchiectomy or medical castration with LHRH agonists (Loblaw *et al.*, 2007). This regime can be combined with direct AR antagonists such as bicalutamide, flutamide or nilutamide, however, whether combined androgen blockade offers benefits beyond castration monotherapy remains controversial (Moul, 2009). This therapeutic intervention

leverages the extraordinary dependence of PCa cells on AR signaling and typically results in suppression of detectable serum PSA, disease-related symptoms and tumor regression. At the cellular level, both cell cycle arrest and cell death are observed upon inhibition of AR activity (Agus *et al.*, 1999), thus underscoring the impact of AR as a valid therapeutic target in advanced disease. However, despite initially favorable results, disease progression to CRPC is inevitable, for which only a limited set of noncurative therapeutic options are available. Although multiple means exist through which CRPC cells evade AR-directed therapies, robust clinical and laboratory investigations support the hypothesis that the CRPC phenotype arises as a result of failure to suppress recurrent AR activity.

1.3.1 Recurrent AR activity in castration-resistant prostate cancer (CRPC)

Radiological evidence of disease progression is almost invariably heralded by a preceding rise in serum PSA, indicating that AR has been aberrantly reactivated despite ADT or combined androgen blockade. Investigating the means by which AR activity is restored in CRPC remains an area of intense investigation and to date four principal mechanisms have been identified in clinical specimens and validated in the laboratory, as playing key roles in disease progression.

1.3.1.1 AR amplification/overexpression

Multiple immunohistochemical studies have shown that AR protein is expressed at high levels (comparable to levels in untreated tumors) in most cases of CRPC, however, the expression may be heterogeneous with a fraction of cells in some tumors being AR low or negative (Sadi *et al.*, 1991; van der Kwast *et al.*, 1991; Ruizeveld de Winter *et al.*, 1994; Mohler *et al.*, 2004). High levels of AR expression are also independently predictive of increased risk of death from PCa (Donovan *et al.*, 2010). Consistent with immunohistochemical data, AR mRNA is also highly expressed in CRPC, with levels being several-fold higher than in primary untreated tumors (Taplin *et al.*, 1995; Holzbeierlein *et al.*, 2004; Stanbrough *et al.*, 2006). One mechanism for increased AR mRNA expression is AR gene amplification, which occurs in about one-third of CRPC cases (Visakorpi *et al.*, 1995; Koivisto *et al.*, 1997).

In addition, there is continued expression of AR regulated genes (such as PSA and also TMPRSS2:ERG fusion genes), indicating that AR transcriptional activity becomes reactivated in CRPC (Mohler *et al.*, 2004; Holzbeierlein *et al.*, 2004; Stanbrough *et al.*, 2006). A landmark study by Chen and colleagues demonstrated, using xenograft models of PCa, that resurgent AR activity after castration of the host is sufficient to confer the CRPC growth phenotype *in vivo* (Chen *et al.*, 2004).

1.3.1.2 AR mutations

A number of AR mutations, primarily within the LBD have been identified in CRPC that can enhance AR activation by weak adrenal androgens and other steroid hormones such as progesterone, estradiol and cortisol (Culig *et al.*, 1993; Taplin *et al.*, 1995). The use of AR antagonists can further select for mutations that convert these therapeutic agents into agonists (Taplin *et al.*, 1999). The most well studied of these mutations at codon 741 (W741C and W741L) identified in LNCaP cells after long term culture with bicalutamide, has also been identified in CRPC patients (Taplin *et al.*, 2003; Hara *et al.*, 2003; Yoshida *et al.*, 2005). However, the overall frequency of AR mutations in patients treated initially with ADT is quite low and this is unlikely to be a major mechanism for progression to CRPC (Taplin *et al.*, 2003). Nonetheless, a recent assessment of AR mutations in circulating tumor cells revealed AR mutations in 20 out of 35 patients with CRPC, indicating that the frequency of alteration could be even higher in cells with metastatic potential (Jiang *et al.*, 2010).

1.3.1.3 Intracrine androgen production

A further mechanism that may contribute to AR activation following castration is the upregulation of enzymes that convert weak adrenal androgens to testosterone to sufficient levels, thereby restoring AR activity (Stanbrough *et al.*, 2006; Mostaghel *et al.*, 2007; Montgomery *et al.*, 2008). It is now well appreciated that adrenal-derived dehydroepiandrosterone (DHEA) can be converted to testosterone and dihydrotestosterone (DHT) through the action of tumor-derived 3β-hydroxysteroid dehydrogenase (3β-HSD) type 1 and type 2 (HSD3B1, HSD3B2), type 5 17β-HSD (AKR1C3) and SRD5A2 (Lin *et al.*, 1997; El-Alfy *et al.*, 1999; Dufort *et al.*,

1999; Penning *et al.*, 2000). Significantly, HSD3B2, AKR1C3, SRD5A1, AKR1C2 and AKR1C1 have been shown to be altered in CRPC compared to primary tumors, lending support to the hypothesis that intracrine androgen synthesis may play a role in CRPC (Ji *et al.*, 2003; Rizner *et al.*, 2003; Stanbrough *et al.*, 2006; Thomas *et al.*, 2008). Therefore, it may be expected that targeting the intracrine androgen synthesis pathway may be of clinical benefit and the use of CYP17,20-lyase inhibitors such as abiraterone have shown particular promise in CRPC (de Bono *et al.*, 2011).

1.3.1.4 AR cofactor alteration and post-translational modifications

Post-translational modifications of AR have been suggested to contribute to enhanced AR function in CRPC, including serine/threonine phosphorylation (Ward and Wiegel, 2009), tyrosine phosphorylation (Guo *et al.*, 2006; Mahajan *et al.*, 2007), acetylation (Popov *et al.*, 2007), ubiquitylation (Dirac and Bernards, 2010; Xu *et al.*, 2009) and sumolyation (Kaikkonen *et al.*, 2009) of AR. However, while many of these events are known to occur downstream of growth factor pathways, it remains uncertain whether therapeutically targeting these growth factor receptors or downstream kinases are of clinical benefit. Finally, an extensive body of literature (Heemers and Tindall, 2007) suggests that AR cofactors may be deregulated in CRPC, including upregulation of coactivators (Agoulnik and Wiegel, 2008), loss of corepressors (Burd *et al.*, 2006) or AR mutations that alter the way that cofactors are perceived (Brooke and Bevan, 2009). However, the simultaneous involvement of multiple coregulators and their overlapping interaction, suggest that additional studies are required to determine the full contribution of coregulators in aberrant AR activity in CRPC.

1.4 Truncated AR variants in benign and non-prostate tissues

For more than a decade, investigators have observed that, in addition to the well-studied 110 kDa AR protein, some lower molecular weight protein bands are immunoreactive with an antibody raised against an epitope mapping to the AR NTD in some AR-expressing cell lines. The origin of these short AR isoforms was quite controversial and at least four potential mechanisms underlying their generation

were proposed: (1) alternative translation start codons; (2) proteolytic cleavage; (3) premature stop codon resulting from mutation; and (4) alternative transcription start site.

1.4.1 AR-A

In 1993, Zoppi and colleagues identified a mutation in the AR NTD carried by a 46,XY phenotypic female (subject R776) with the syndrome of complete testicular feminization (Zoppi *et al.*, 1993). This mutation, termed R776, introduces a premature stop codon in place of the normal AR codon for amino acid 60. COS cells transfected with cDNA encoding the R776 mutation produces an 87 kDa form of AR that lacks the normal NTD. Although this short AR isoform can activate an androgen-responsive reporter gene coexpressed in CV1 cells, the level of activation is markedly reduced when compared to the effect of full-length AR in this system (Zoppi *et al.*, 1993). It was proposed that the 87 kDa form (termed AR-A) is due to translation initiation of AR protein at the internal methionine 188 residue of the full-length AR (Wilson and McPhaul, 1994). They also suggested that AR-A and full-length AR may differ in their ability to activate target genes and are regulated differently in various cell types (Wilson and McPhaul, 1994).

1.4.2 AR45

AR45 is a naturally occurring truncated AR variant originally identified by 5′ rapid amplification of cDNA ends (RACE) with RNA isolated from human placental tissue (Ahrens-Fath *et al.*, 2005). The AR45 transcript was found to originate from alternative splicing of a previously unreported exon located ~22.1 kb downstream of AR intron 1. This cryptic exon, termed exon 1B can be spliced in place of AR exon 1 to yield a 45 kDa isoform containing the DBD, CTD and a novel seven amino acid sequence at the NH_2-terminus in place of the wild-type AR NTD. Analysis of AR45 by RT-PCR revealed expression in a number of human tissues with the highest levels in heart and skeletal muscle tissue. However, these RT-PCR analyses were not quantitative, so it is unclear how AR45 expression levels compare with full-length AR in these tissues. It is also unclear whether AR45 transcript is translated into

endogenous protein in these tissues; however immunoblot analysis of LNCaP lysates demonstrated the presence of a ~45 kDa isoform that was immunoreactive with an antibody specific for the AR CTD. AR45 overexpressed in CV-1 cells was shown to bind androgen, localize to the nucleus, interact with the full-length AR NTD and inhibit full-length AR activity in a ligand- and DBD-dependent manner. Ectopically expressed AR45 in LNCaP cells also inhibited proliferation. These observations indicate that AR45 is a negative regulator of AR signaling. However, when the coactivator TIF2 or the oncogene β-catenin was overexpressed, AR45 stimulated androgen-dependent promoters in the presence of androgens (Ahrens-Fath *et al.*, 2005). Interestingly, a recent study has implicated AR45 in the increased risk of the drug-induced cardiac arrhythmia, long QT syndrome in women. The human ether-a-go-go-related gene (HERG) potassium channel is largely responsible for determining QT interval and androgen-mediated stabilization of HERG protein was observed in cells transfected with AR45 but not full-length AR (Wu *et al.*, 2008).

1.4.3 AR splicing in androgen insensitivity syndrome (AIS)

Androgen insensitivity syndrome (AIS) is an X-linked condition first described in 1974 in 46,XY individuals as a result of AR genetic alterations characterized by end-organ resistance to androgens leading to defects in male sexual differentiation (Wilson *et al.*, 1974). These AR alterations result in AR proteins with impaired function that prevents normal androgen signaling and development of internal and external male phenotypic characteristics (Brinkmann, 2001). The most severe AIS phenotype is complete AIS (CAIS), in which individuals have female external genitalia, a short, blind ending vagina, the absence of Wolffian duct derived structures such as the epididymides, vas deferens and seminal vesicles, the absence of a prostate, development of gynecomastia and the absence of pubic or axillary hair (Quigley *et al.*, 1995). Partial AIS (PAIS) includes a wide range of phenotypes from individuals with a predominantly female appearance to persons with ambiguous genitalia or individuals with a predominantly male phenotype (Quigley *et al.*, 1995). In general, the severity of AIS is proportional to the level of impaired AR activity caused by specific alterations in the AR gene. Several types of AR mutation have

been identified in individuals with AIS, including: (1) single point mutations resulting in amino acid substitutions or premature stop codons; (2) nucleotide insertion/deletion leading to frame shirt and premature termination and (3) complete or partial gene deletion (Gottlieb *et al.*, 2004). Interestingly, a number of alternatively spliced AR transcripts have been identified in AIS leading to AR proteins with impaired function. AR gene defects have been implicated with these alternatively spliced variants, often within intronic splice donor or splice acceptor sites.

1.4.3.1 AR splicing disruption in complete AIS

A number of alternatively spliced AR variants have been described in patients with CAIS. A G→T point mutation within the splice donor site of intron 4 has been described in a patient with CAIS (Ris-Stalpers *et al.*, 1990). This mutation was shown to lead to the use of an alternative cryptic splice donor in exon 4, resulting in an in-frame 123 bp deletion from AR transcript and a concomitant 41 amino acid internal deletion of the AR protein. This variant displayed no androgen-binding activity and no transcriptional activity on an androgen-responsive promoter gene. Similar mutations in splice donor sites of introns 2 and 7 have been reported in two other individuals with CAIS, leading to skipping of exons 2 and 7 from AR transcripts respectively (Lim *et al.*, 1997; Hellwinkel *et al.*, 1999). The exon 2-deleted AR protein was prematurely truncated as a result of a premature stop codon encoded by an out-of-frame exon 3 (Lim *et al.*, 1997). The exon 7-deleted AR protein was ~98 kDa (Hellwinkel *et al.*, 1999). In both these studies, the AR variants identified were unable to bind androgens.

1.4.3.2 AR splicing disruption in partial AIS

Alternatively spliced AR variants have also been identified in individuals with PAIS. This suggests that cells with altered AR splicing can retain some residual AR activity. This may result from AR variants maintaining some degree of transcriptional activity or a residual level of splicing that permits synthesis of the full-length AR transcript and protein. A >6 kb in-frame deletion has been reported in AR intron 2 in a patient with PAIS, which results in skipping of exon 3 and synthesis

of an AR variant protein with a deletion of 39 amino acids from the DBD (Ris-Stalpers *et al.*, 1994). This AR variant exhibited normal androgen binding, but no transcriptional activity on an androgen-responsive promoter gene. A T→A mutation 11 bp upstream of exon 3 has also been described and has been shown to be associated with two alternatively spliced AR transcripts (Bruggenwirth *et al.*, 1997). The first alternatively spliced AR transcript arose through skipping of exon 3 and the other resulted from a cryptic splice acceptor site 71 bp upstream of exon 3. The AR protein encoded by this alternatively spliced acceptor site contained an additional 23 amino acid in-frame insert in the DBD between the two zinc fingers (Bruggenwirth *et al.*, 1997). Functionally, this receptor was shown to bind androgens, but no DNA-binding activity was observed in a band shift assay. In another PAIS individual, a G→T mutation in the splice donor site of intron 6 resulted in expression of a larger AR transcript species (Sammarco *et al.*, 2000). An in-frame stop codon 79 bp downstream from the donor splice site in intron 6, resulted in a smaller, truncated AR protein. Impaired androgen binding was observed in this AR variant (Sammarco *et al.*, 2000). Finally, an additional splicing mutation has been identified as a C→T mutation in codon 888 of exon 8 with replacement of the final 33 amino acids of the full-length AR protein with a novel eight amino acid sequence as a result of a cryptic splice donor site (Hellwinkel *et al.*, 2001). Functionally, this AR variant exhibited impaired androgen-binding capacity and no transcriptional activity on an androgen-responsive promoter gene.

1.5 Truncated AR variants in malignant prostatic tissues

It was recently discovered that AR alternative splicing can give rise to CTD truncated AR protein isoforms that lack the AR LBD in PCa cells. These AR isoforms are lacking the functional domain of the receptor that would be predicated to mediate responses to traditional ADT, as well as new AR-targeted therapies such as abiraterone (de Bono *et al.*, 2011; Mostaghel *et al.*, 2011), enzalutamide (formerly MDV3100; Tran *et al.*, 2009; Scher *et al.*, 2010) and emerging next-generation anti-androgens such as ARN-509 (Clegg *et al.*, 2012).

1.5.1 Identification of AR variants in the CWR22 model

Seminal observations leading to the discovery of alternatively spliced truncated AR variants were initially reported in the 22Rv1 cell line. 22Rv1 cells are derived from the CWR22 transplantable human prostate tumor model (Wainstein *et al.*, 1994). This tumor exhibits androgen-dependent growth and PSA secretion in male nude athymic mice and regresses in response to androgen withdrawal as defined by tumor shrinkage and a decline in serum PSA levels (Nagabhushan *et al.*, 1996). Importantly, it simulates the clinical course of PCa in that relapsed growth of the tumor (designated CWR22R) occurs after several months of androgen withdrawal and is preceded by rising PSA levels. The CWR22Rv1 (also termed 22Rv1) cell line was established from the relapsed CWR22R tumor serially propagated in mice (Sramkoski *et al.*, 1999) and is characterized by androgen-independent proliferation, AR expression and androgen-responsiveness. It serves as a model possessing an intermediate phenotype between that of hormone-sensitive, AR-positive cells lines (e.g. LNCaP) and androgen-independent, AR-negative cells lines (e.g. DU-145 and PC-3).

1.5.1.1 ARΔLBD

In 2002, Tepper and colleagues made the critical observation that 22Rv1 cells express two discrete AR protein isoforms of 110-114 kDa and 75-80 kDa (Tepper *et al.*, 2002). Using an antibody mapping approach, it was demonstrated that the smaller 75-80 kDa AR species contained an intact transactivation domain and DBD but lacked the LBD (referred to as ARΔLBD). Further biochemical characterization, determined that the ARΔLBD isoform was constitutively nuclear and could bind DNA independent of androgens. Co-immunoprecipitation experiments indicated that the ARΔLBD isoform did not interact with full-length AR, suggesting this species functioned as an independent transcription factor (Tepper *et al.*, 2002). In addition, characterization of the 22Rv1 full-length AR transcript using RT-PCR mapping with overlapping primer pairs revealed an insertion of an in-frame 117 bp sequence occurring between the 3'-end of exon 3 and the 5'-end of exon 4, which was identified as AR exon 3. This represents a tandem duplication, referred to as exon 3'

and encodes an additional 39 amino acids and 4-5 kDa of protein mass, thus accounting for the observed increase in size of 22Rv1 full-length AR. Exons 2 and 3 each encode zinc fingers comprising the DBD, therefore, the 114 kDa 22Rv1 full-length AR species contains an additional zinc finger (Tepper *et al.*, 2002). Interestingly, these mutations were not detected in the original androgen-dependent CWR22 xenograft, indicating that this change occurred during the progression to androgen-independence.

Several recent studies have indicated that alternative splicing may be an important contributor to the synthesis of truncated AR species lacking the LBD in 22Rv1 and other cells. This concept was first suggested by Dehm and colleagues in 2008 based on the observation of differential siRNA targeting of the various AR isoforms in 22Rv1 cells (Dehm *et al.*, 2008). siRNA targeted to AR exons 7/8 abolishes expression of full-length AR in 22Rv1 cells but had no effect on the smaller 75-80 kDa isoform. Conversely, siRNA targeted to AR exon 1 knocked down expression of both AR isoforms in 22Rv1 cells. Functionally, androgen-dependent expression of AR target genes and cell proliferation was shown to be attributable to full-length AR, whereas androgen-independent expression of AR target genes and constitutive, androgen-independent cell proliferation was shown to be supported by the smaller 75-80 kDa isoforms. These important observations strongly suggested that different mRNA species encoded the different AR protein species in 22Rv1 cells and provided the foundation for 3′ RACE and other approaches to identify their origin (Dehm *et al.*, 2008; Hu *et al.*, 2009; Guo *et al.*, 2009; Marcias *et al.*, 2010; Sun *et al.*, 2010; Watson *et al.*, 2010; Hu *et al.*, 2011; Yang *et al.*, 2011). This led to the identification of a series of alternatively spliced AR transcripts expressed in 22Rv1 cells that result from cryptic exons located within introns 2 and 3.

1.5.1.2 AR 1/2/2b and AR 1/2/3/2b

AR exon 2b was indentified following 5′ RACE experiments with RNA derived from 22Rv1 cells (Dehm *et al.*, 2008). AR exon 2b was shown to splice downstream of either AR exon 2 or AR exon 3, resulting in truncated AR proteins containing the AR NTD, one or both zinc fingers of the DBD and a short 11 amino

acid COOH-terminus extension encoded by AR exon 2b (Figure 1.2; Table 1.1). Functionally, AR 1/2/2b and AR 1/2/3/2b exhibited constitutive transcriptional activity on promoter-reporter gene assays. AR 1/2/2b transcript was also shown to be expressed in VCaP cells as well as the LuCaP 23.1 and 35 xenograft models of PCa progression.

1.5.1.3 AR-V1 to AR-V7

A comprehensive bioinformatic analysis of the ~170 kb AR intron sequences against the National Center for Biotechnology Information (NCBI) human expressed sequence tag database was performed by Hu and colleagues to identify "intronic" genomic fragments (Hu *et al.*, 2009). Using this approach, three cryptic exons were identified in AR intron 3 (termed CE1, CE2 and CE3) and one in AR intron 2 (termed CE4). In order to determine the exon-intron splicing junctions, RT-PCR was used to amplify and sequence transcripts containing AR exon 2 and the putative cryptic exons in 22Rv1 cells. Sequencing of the amplicons defined the 5′ junctions of CE1, CE2 and CE3 and the 5′ and 3′ junctions of CE4 and was used to construct seven AR variant transcripts, named AR-V1 to AR-V7. The genomic coordinates of CE4 are identical to the novel exon AR 2b identified by Dehm and colleagues (Dehm *et al.*, 2008). Significantly, all seven AR variants harbor premature stop codons downstream of AR exon 2, generating LBD-truncated AR proteins.

Two alternatively spliced AR isoforms identified in this study were further characterized. AR-V1 (composed of AR exons 1/2/3/CE1) was detected by quantitative RT-PCR at greater levels in RNA isolated from CRPC specimens compared with hormone naïve primary PCa specimens (Hu *et al.*, 2009). AR-V7 (composed of AR exons 1/2/3/CE3) transcript was detected in various clinical specimens with the highest levels in CRPC tissues. Interestingly, a subset of hormone-naïve primary PCa specimens expressed AR-V1 and AR-V7 transcript at levels comparable to those CRPC specimens. Elevated AR-V7 mRNA expression in these 66 patients was associated with a poor clinical outcome as defined by biochemical recurrence following prostatectomy. In this same cohort, higher levels of

Figure 1.2 Alternatively spliced AR isoforms identified in prostate cancer

A. Schematic representation of the AR gene locus with locations of alternatively spliced, cryptic exons in gray (depictions are not to scale). Nomenclature of novel exons is based on the relative position within the AR locus.

B. Alternative splicing of cryptic exons in the AR locus or exon skipping gives rise to C-terminal truncated AR transcript isoforms. Additional alternatively-spliced mRNA isoforms which have not been cloned or characterized are not illustrated.

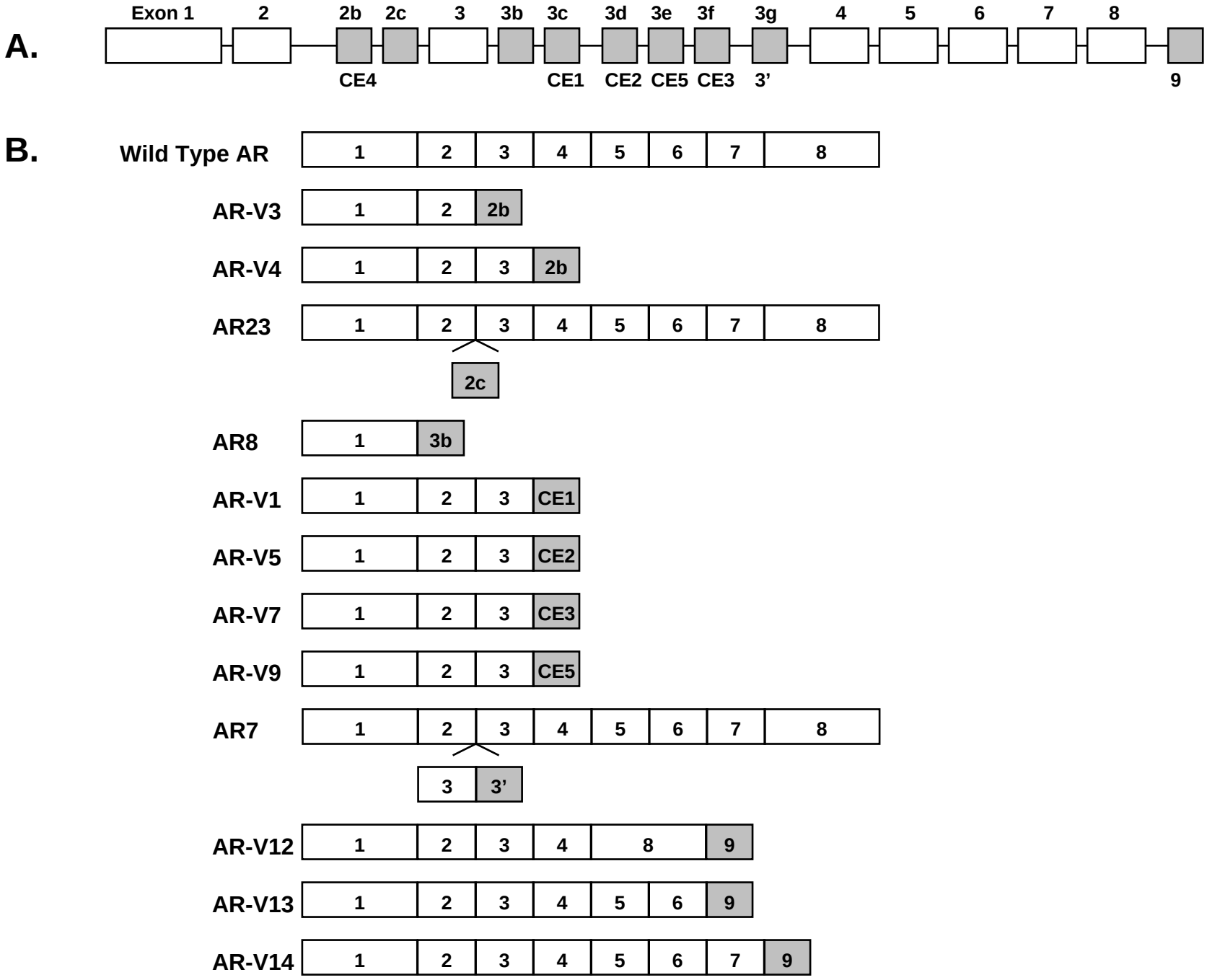

A.
Exon 1
2
2b
2c
3
3b
3c
3d
3e
3f
3g
4
5
6
7
8
CE4
CE1
CE2
CE5
CE3
3'
9
B.
Wild Type AR
1
2
3
4
5
6
7
8
AR-V3
2b
AR-V4
AR23
2c
AR8
3b
AR-V1
CE1
AR-V5
CE2
AR-V7
CE3
AR-V9
CE5
AR7
3'
AR-V12
AR-V13
AR-V14

Table 1.1 Human AR variants

Variant	Alternate names	Intron retained/ novel exon	Splicing junction	Exons excluded	Unique COOH terminus peptide	Localization	Activity	References
AR23		2c	2/2c		23 inserted in DBD	Cytosolic	Conditional	Jagla *et al.*, 2007
AR8		3b/3'	1/3'	2 to 8	33	Cytosolic		Yang *et al.*, 2011
AR-V1	AR4	3c/CE1	3/CE1	4 to 8	19	Cytosolic	Conditional	Hu *et al.*, 2009; Guo *et al.*, 2009
AR-V2		3c/CE2	3/3/CE1	4 to 8	19			Hu *et al.*, 2009
AR-V3	AR 1/2/2b, AR6	2b/CE4	2/CE4	3 to 8	11		Constitutive	Dehm *et al.*, 2009; Hu *et al.*, 2009; Guo *et al.*, 2009
AR-V4	AR1/2/3/2b, AR5, ARV6	2b/CE4	3/CE4	4 to 8	11	Nuclear	Constitutive	Dehm *et al.*, 2009; Hu *et al.*, 2009; Guo *et al.*, 2009; Marcias *et al.*, 2010
AR-V5		3d/CE2	3/CE2	4 to 8	1			Hu *et al.*, 2009
AR-V6		3d/CE2	3/CE2	4 to 8	6			Hu *et al.*, 2009
AR-V7	AR3	3f/CE3	3/CE3	4 to 8	16	Cytosolic/ Nuclear	Constitutive	Hu *et al.*, 2009; Guo *et al.*, 2009
ARV7		3g/3'	3/3/3'	4 to 8	22	Nuclear	Constitutive	Marcias *et al.*, 2010
AR-V8		3	3/intron 3	4 to 8	10			Hu *et al.*, 2011
AR-V9		3e/CE5	3/CE5	4 to 8	23	Cytosolic	Conditional	Watson *et al.*, 2010; Hu *et al.*, 2011
AR-V10		3	3/intron 3	4 to 8	39			Watson *et al.*, 2010
AR-V11		3	3/intron 3	4 to 8	20			Watson *et al.*, 2010
ARV567es	AR-V12	9	4/8/9	5 to 7	10	Nuclear	Constitutive	Sun *et al.*, 2010; Hu *et al.*, 2011
AR-V13		9	6/9	7 to 8	3		Inactive	Hu *et al.*, 2011
AR-V14		9	7/9	8	7			Hu *et al.*, 2011

full-length AR or AR-V1 transcript were not associated with biochemical recurrence (Hu *et al.*, 2009).

1.5.1.4 AR3 to AR5

In a separate study, Guo and colleagues employed the CWR-R1 cell model to identify novel AR splicing variants (Guo *et al.*, 2009). The CWR-R1 cell line is derived from recurrent CWR22 xenograft tumors harvested from male nude mice 140–160 days after castration (Gregory *et al.*, 2001). Expression of AR using an antibody recognizing the AR NTD revealed the full-length AR at 110 kDa and also an 80 kDa immunoreactive band. Using a panel of shRNAs targeting distinct regions of the AR gene suggested that both AR isoforms may be translated from more than one transcript. Using 3' RACE with a primer corresponding to a target sequence in AR exon 1, more than 20 splicing variants were identified in CWR-R1 cells with three variants (designated AR3, AR4 and AR5) predicted to encode a protein ~80 kDa (Guo *et al.*, 2009). The splicing patterns for AR3 and AR4 transcripts were identical to AR-V7 (AR exons 1/2/3/CE3) and AR-V1 (AR exons 1/2/3/CE1) identified by Hu and colleagues respectively in 22Rv1 cells (Hu *et al.*, 2009). Similarly, the splicing configuration of AR5 was identical to AR 1/2/3/2b and AR-V4 (AR exons 1/2/3/CE4) both identified independently in 22Rv1 cells (Dehm *et al.*, 2008; Hu *et al.*, 2009).

Of particular significance in this study, the authors developed an antibody against the unique 16 amino acid sequence at the COOH-terminus of AR3 and demonstrated protein expression of this isoform was increased in CRPC versus hormone-naïve PCa in a human PCa tissue microarray (Guo *et al.*, 2009). Also, cytoplasmic expression of this isoform could predict biochemical recurrence following prostatectomy (Guo *et al.*, 2009). Interestingly, the AR3 isoform was also expressed at the transcript and protein level in benign prostate tissue, indicating that there may also be a role for this splicing event in normal cells. Functionally, AR3 was demonstrated to function as a constitutively active, ligand-independent transcription factor and could induce a CRPC growth phenotype in LNCaP cells grown *in vitro* and as xenografts *in vivo* (Guo *et al.*, 2009). Overall, this body of work demonstrates that an AR variant derived from contiguous splicing of AR exons 1/2/3/CE3, termed

AR3 or AR-V7, is a clinically validated AR isoform that could play an important role in supporting the CRPC phenotype.

1.5.1.5 ARV6 and ARV7

Additional alternatively spliced AR isoforms have also been cloned from 22Rv1 cells using a yeast functional assay to identify AR cDNAs with constitutive or ligand-induced transcriptional activity (Marcias *et al.*, 2010). This approach led to the identification of an additional cryptic exon in AR intron 3 termed 3′. Sequencing of the isolated cDNA, which was named ARV7 in this study, demonstrated that exon 3′ spliced downstream of the tandem-duplicated exon 3 in 22Rv1 cells. This 98 bp intronic insert resulted in a novel stop codon, giving rise to a truncated isoform containing the entire AR NTD, DBD, an extra zinc finger encoded by the extra copy of exon 3 and a novel 22 amino acid COOH-terminus extension encoded by exon 3′. The authors also describe an additional AR splicing isoform in 22Rv1 cells, termed AR6, which exhibits an intronic 280 bp insert that is identical to the 2b exon first reported by Dehm and colleagues (Dehm *et al.*, 2008). However, this intronic insert is downstream of the first AR exon 3, followed by the complete remaining AR exons 4 to 8 and also exhibits a trinucleotide (TAA) deletion. The polyadenylation signal described by Dehm and colleagues is present in the intronic sequence but has no influence on mRNA transcription termination. However, ARV6 mRNA translation results in a truncated AR isoform from an intronic novel stop codon further downstream (Marcias *et al.*, 2010). The reason for this discrepancy is not clear, but could be due to the fact that 22Rv1 cells harbor two genomic copies of AR exon 2b (Li *et al.*, 2011).

1.5.1.6 AR8

AR8 is the most recent alternatively spliced AR isoform to be identified in CWR-R1 cells using 3′ RACE (Yang *et al.*, 2011). This novel AR isoform is composed of AR exons 1, 3 and 3b (identical to CE3). Due to usage of an alternative splice acceptor site in exon 3 (denoted as 3′), this transcript is deduced to encode a protein containing the NTD and a unique 33 amino acid sequence at the COOH-terminus.

AR8 transcript was detected in a panel of benign and malignant tumor specimens. The level of AR8 transcript was also elevated in the castration-resistant LNCaP sublines, C4-2 and C4-2b, as well as CWR22 xenograft tumors compared to their androgen-dependent counterparts. Using a polyclonal antibody specific for the unique COOH-terminus of the AR8 protein, endogenous AR8 protein expression was detected in a number of castration-resistant PCa cell lines including CWR-R1, VCaP, C4-2 and C4-2b (Yang *et al.*, 2011).

1.5.2 Identification of AR variants in the VCaP model

The VCaP cell line has been an additional cell-based model for studying alternatively spliced AR variants, following the observation of a smaller 75-80 kDa AR species that is immunoreactive with antibodies directed to the NTD, but not the LBD (Dehm *et al.*, 2008). VCaP cells are derived from PCa tissue harvested at autopsy from a metastatic lesion to a lumbar vertebral body of a patient with CRPC, aseptically xenografted into SCID mice, and later harvested and plated onto tissue culture dishes (Korenchuk *et al.*, 2001).

1.5.2.1 AR-V8 to AR-V11

Watson and colleagues employed an adaptation of the 3′ RACE method whereby 3′ RACE products produced using a forward primer spanning the AR exon 2/3 junction in VCaP cells were sequenced using two next-generation sequencing platforms, 454 and SOLiD (Watson *et al.*, 2010). This approach confirmed splicing of AR exons CE1 and CE3 downstream of AR exon 3 and also identified four additional isoforms, named AR-V8 through AR-V11. These novel AR transcripts resulted from AR exon 3 read-through, splicing of novel cryptic exons in AR intron 3, or use of a different splice acceptor site in exon CE3. As in previous studies, these novel AR mRNAs would be predicted to encode variable-length COOH-terminus extensions following the exon 3 sequence**.** Using this approach, additional novel AR transcripts were identified in VCaP cells that appear to result from exon skipping (AR exons 4/6, 4/7, 4/8 and 6/8) and splicing of a novel cryptic exon within AR intron 5 (Watson *et al.*, 2010). However, no full-length cDNAs were isolated from VCaP cells that

correspond to these newly identified AR transcripts, so it is unclear whether these novel AR isoforms represent splicing intermediates, splicing errors or are actual alternatively spliced AR transcripts that encode functional proteins.

1.5.3 Identification of AR variants in the mouse Myc-CaP model

AR variant discovery has also been extended into murine models of PCa using the Myc-CaP model. The Myc-CaP cell line is derived from tumors that arose in mice with a prostate-specific c-myc transgene, which drives carcinogenesis in a stepwise fashion from prostatic intraepithelial neoplasia to invasive cancer (Ellwood-Yen *et al.*, 2003; Watson *et al.*, 2005).

1.5.3.1 mAR-V1 to mAR-V4

Watson and colleagues employed their 3' RACE approach combined with next generation sequencing in the Myc-CaP model and demonstrated the synthesis of four novel alternatively spliced AR isoforms in these cells, designated mAR-V1 to mAR-V4 (Watson *et al.*, 2010). mAR-V1 was composed of contiguously spliced mouse exons 1/2/3 followed by read-through into intron 3, with a unique 40 amino acid COOH-terminus sequence before an in-frame stop codon. Unlike the human AR variants already described, mAR-V2 and mAR-V4 were found to harbor novel exons that map outside the mouse AR gene to distal regions on the X chromosome. The first exon, located ~250 kb downstream of the AR locus was shown to be spliced after AR exon 3 in the mAR-V2 variant. The other exon, located ~1 Mb upstream of the AR locus was shown to be spliced after AR exon 4 to yield mAR-V3 and mAR-V4. The structural basis for the generation of these isoforms at the genomic level remains unclear. However, the fact that Myc-CaP cells have AR gene amplification raises the possibility of intrachromosomal gene rearrangements.

Protein expression of mAR-V2 and mAR-V4 and also isoform-specific knockdown of these species were confirmed in Myc-CaP cells (Watson *et al.*, 2010). Functionally, mAR-V4 was shown to be constitutively transcriptionally active in a ligand-independent manner in an androgen-responsive promoter-reporter assay and also localized to the nucleus. However, mAR-V2 did not display transcriptional

activity and was found to be localized to the cytoplasm. Consistent with these findings, LNCaP xenografts stably expressing mAR-V4, but not mAR-V2 could support tumor growth in castrated mice.

1.5.4 Identification of AR variants in the LuCaP xenograft model

The LuCaP series is an extensive panel of human PCa xenograft models derived from a range of primary PCa, soft tissue, lymph node and osseous bone metastases (obtained during surgery or during a rapid autopsy) developed at the University of Washington, Seattle. These xenograft models have proved to be valuable resource for furthering our understanding of PCa progression from an androgen-dependent phenotype to an androgen-independent CRPC phenotype.

1.5.4.1 ARV567es

Sun and colleagues performed a comprehensive RT-PCR analysis on a panel of 25 LuCaP xenografts, mostly derived from metastases obtained from men with CRPC after prolonged ADT (Sun *et al.*, 2010). Two of the LuCaP xenografts, 86.2 and 136 (derived from a bladder metastasis and cells from ascitic fluid, respectively from two different men who had relapsed following prolonged ADT) express shorter AR transcripts in the region spanning AR exons 2-8. Sequencing the short AR transcripts from both xenografts found identical cDNA sequences, lacking AR exons 5, 6 and 7, hence the designation ARV567es, in which "es" denotes exons skipped (Sun *et al.*, 2010). While the full nucleotide sequence of AR exon 8 is present, due to the splicing of exon 4 to exon 8 a frame-shift occurs in the open reading frame of ARV567es, resulting in a premature stop codon after the first 30 nucleotides, and leading to a shortened AR exon 8 sequence of 10 amino acids. Interestingly, in the LuCaP 86.2 xenograft both full-length AR transcript and protein not appear to be expressed, which is in contrast to 22Rv1 and other models where full-length AR expression remains high. Using primers designed to specifically detect the AR exon 4-8 junction present in ARV567es; RT-PCR detected this AR variant isoform in patient specimens of primary PCa, CRPC, as well as benign prostate epithelium (Sun *et al.*, 2010). Functionally, AR567es was shown to act as a constitutively active, ligand-

independent transcription factor that could support CRPC cell growth *in vitro* and *in vivo* (Sun *et al.*, 2010).

1.5.5 Identification of AR variants in CRPC specimens

1.5.5.1 AR-Q640X

AR-Q640X is a nonsense mutation that results in the substitution of the glutamine residue in AR position 640 by a premature stop codon identified using a yeast-based functional assay on metastatic bone marrow aspirate from a man who had progressed on ADT (Ceraline *et al.*, 2004). This mutation was subsequently identified in other CRPC patients (Lapouge *et al.*, 2007). This AR isoform was immunoreactive at ~75 kDa with an AR NTD-directed antibody confirming that the Q640X mutation led to a CTD truncated protein (Ceraline *et al.*, 2004). The AR-Q640X mutation is located downstream of the NLS within the second zinc finger of the DBD and the hinge region, therefore a nonsense mutation at this codon may be predicted to confer constitutive transcriptional capabilities. Consistent with this hypothesis it was demonstrated that AR-Q640X displays constitutive transcriptional activity in a ligand-independent manner in an androgen-responsive promoter-reporter assay and is also constitutively localized to the nucleus (Ceraline *et al.*, 2004).

1.5.5.2 AR23

This functional yeast assay approach has also detected an AR splice variant named AR23, from a bone metastasis in a man with CRPC (Jagla *et al.*, 2007). This AR isoform results from an exon/intron splicing error resulting in a 69 bp insertion which corresponds to the 3'-part of AR intron 2. The insertion occurred at the junction of exon 2 and exon 3, which encode for the first and the second zinc finger respectively. Analysis of this novel AR variant cDNA sequence did not reveal any mutation in the classical donor or acceptor splice sites. As a result of this aberrant splicing, the last 69 bp of intron 2 persist in AR transcripts. These supplementary nucleotides do not lead to a premature stop codon but give rise to an AR variant, containing an insertion of 23 amino acids between the two zinc fingers of the DBD (Jagla *et al.*, 2007). Functionally, AR23 remains in the cytoplasm and nuclear

translocation is impaired following androgen binding. Despite the predicted lack of DNA-binding function, expression of AR23 in LNCaP cells increased transcriptional activity at androgen-responsive promoters. The mechanism remains unclear; however AR23 was able to stimulate NF-κB transcriptional activity and decrease AP-1 transcriptional activity, suggesting that non-genomic functions of novel AR splice variants may be important in PCa disease progression (Jagla *et al.*, 2007).

1.5.5.3 AR-V12 to AR-V14

Another novel method has recently been developed to identify additional alternatively spliced AR isoforms in CRPC specimens based on selective linear amplification of sense RNA (SLASR) using an AR exon 3-anchored forward primer, followed by detection of amplified transcripts using an AR gene tiling array (Hu *et al.*, 2011). Using this approach, the synthesis of mRNAs containing AR exons 2b/CE4, CE1, CE2 and CE3 were confirmed in 22Rv1 cells along with a novel isoform termed AR-V9, which was produced by contiguous splicing of AR exons 1/2/3 and a novel cryptic exon located in AR intron 3, termed CE5. This AR isoform was identical to the AR-V9 isoform reported in VCaP cells (Watson *et al.*, 2010). This novel workflow was then deployed in RNA derived from two CRPC specimens, leading to the identification of an additional novel AR exon, termed 'exon 9'. This exon was located downstream of AR exon 8 and was shown to be the most 3' exon in three novel AR mRNA species, termed AR-V12 (AR exons 1/2/3/4/8/9, which encodes the ARV567es variant), AR-V13 (AR exons 1/2/3/4/5/6/9) and AR-V14 (AR exons 1/2/3/4/5/6/7/9; Hu *et al.*, 2011). Functional characterization of AR-V9 and AR-V12 found varying degrees of constitutive ligand-independent transcriptional activity depending on the cell line studied. AR-V9 was also shown to function independently of full-length AR in LNCaP cells. AR-V9 and AR-V12 transcript expression is significantly elevated in CRPC specimens compared to hormone-naïve and benign specimens. Follow-up data available for patients with hormone-naïve PC found no association between AR-V9 or AR-V12 expression and biochemical recurrence. Moreover, neither AR-V9 nor AR-V12 was associated with pre-operative PSA values or disease stage. However, significantly higher AR-V12 expression levels, but not AR-V9, was found in patients

with Gleason score 8 and above, when compared to those with Gleason score 6 and below or Gleason score 7 (Hu *et al.*, 2011).

1.6 Characterization of AR variants

1.6.1 Function of AR variants

A particularly intriguing observation made by several groups is that although AR variants are expressed at the highest levels in CRPC specimens, AR variants are expressed at variable levels in benign prostate epithelium, as well as hormone-naïve primary PCa specimens (Hu *et al.*, 2009; Guo *et al.*, 2009; Sun *et al.*, 2010; Hu *et al.*, 2011). In one study, full-length AR was detected in 35 out of 36 samples of laser captured microdissected (LCM) benign prostate biopsies from men (aged 35-55) without PCa and serum PSA less than 2 ng/ml who were administered ADT as part of a contraception study. In this study a total of 10 samples (27.8%) also expressed ARV567es or AR-V7 (Sun *et al.*, 2010). In a separate cohort of non-castrate men (ages 59-70) who had radical prostatectomy for PCa, analysis of paired LCM benign and malignant prostate epithelial samples found that ARV567es and AR-V7 transcript expression could be found in benign, malignant or both specimens (Sun *et al.*, 2010). These data suggest that AR variants are not necessarily etiologic in PCa initiation, but may serve to facilitate progression when ADT is applied. This phenomenon has been demonstrated using LNCaP cells stably expressing ARV567es (Sun *et al.*, 2010), VCaP (Watson *et al.*, 2010), LuCaP 35 (Sun *et al.*, 2010; Watson *et al.*, 2010) and LuCaP 136 (Sun *et al.*, 2010) xenografts grown in castrate versus intact SCID mice.

Another hypothesis is that AR variant expression may be a mechanism of resistance to current ADT therapies and development of CRPC. Overexpression of AR-V7 in LNCaP cells and specific siRNA targeted knock down of endogenous AR variants in 22Rv1 cells result in increased and decreased growth respectively, under CRPC-like condition *in vivo* and *in vitro* (Dehm *et al.*, 2008; Guo *et al.*, 2009; Watson *et al.*, 2010). ARV567es increases proliferation of LNCaP cells in the absence of androgens as well as enhanced proliferation in response to very low levels of androgen (Sun *et al.*, 2010). Similarly, in intact non-castrate mice there were no differences in tumor growth between LNCaP xenografts and LNCaP xenografts

stably expressing ARV567es, however in castrated mice, ARV567es tumors were larger compared to controls (Sun *et al.*, 2010). Furthermore, in support of this hypothesis, AR-V7 (Hu *et al.*, 2009; Guo *et al.*, 2009; Zhang *et al.*, 2011), AR-V1 (Hu *et al.*, 2009), AR-V9, AR-V12 (Hu *et al.*, 2011) and ARV567es (Zhang *et al.*, 2011) are all expressed at higher levels in CRPC specimens. These data suggest that AR variants may be a valid therapeutic target in those men who progress while on ADT.

1.6.2 Subcellular location of AR variants

AR variants are structurally similar, however as it has already been elaborated, may have distinct biological activity and functions. One obvious difference is the subcellular location of specific AR variants may predict their activity. However, the cellular localization of AR variants may be cell-specific, dependent on the presence of full-length AR or androgen status. AR-V7 has been found to be localized to the nucleus in culture under androgen-depleted conditions and constitutively active in driving the expression of genes including canonical androgen-responsive genes using reporter assays and microarray analysis (Hu *et al.*, 2009; Guo *et al.*, 2009). However, analysis of clinical specimens reveals a more complicated expression pattern. In benign prostate tissues, AR-V7 is mainly expressed in the basal and stromal cells with minimal staining in luminal epithelial cells. However, in malignant tissues, AR-V7 stained the majority of luminal cells in the cytoplasm and was nuclear in CRPC tissues (Guo *et al.*, 2009). This suggests additional variables are required for AR-V7 to localize to the nucleus.

AR8 does not contain a DBD and is therefore unlikely to function as a transcription factor. This lack of transcriptional activity was confirmed by testing several reporters driven by androgen-responsive promoters (Yang *et al.*, 2011). A further surprising observation from this study was that AR8 localized to the plasma membrane when overexpressed in AR-null COS-1 cells, whereas AR-V7 localized to the cytosol or the nucleus of COS-1 cells under the same conditions. Similar results were obtained when AR8 was overexpressed in LNCaP and CWR-R1 cells, suggesting that AR8 is preferentially associated with the plasma membrane (Yang *et al.*, 2011). This is possibly mediated through palmitoylation of two cysteine residues

at codons 558 and 560 located within the unique COOH-terminus amino acid sequence of AR8. Mutation of these putative palmitoylation sites led to loss of membrane localization of AR8 (Yang *et al.*, 2011). The primarily plasma membrane localization of AR8, coupled with the lack of a DBD, would suggest AR8 most likely functions through non-genomic mechanisms. This may be accomplished through EGFR, Src, full-length AR and AR8 forming a dynamic signaling complex in response to EGF, and the level of AR8 may modulate kinetics of the assembly and dissociation of this complex, allowing sequential phosphorylation and subsequent nuclear translocation (Yang *et al.*, 2011).

AR-V1 remains in the cytoplasm regardless of androgen status when overexpressed in AR-null COS-7 cells (Watson *et al.*, 2010). AR-V9 is also exclusively cytoplasmic (Hu *et al.*, 2011). AR-V12 (Hu *et al.*, 2011) and ARV567es (Sun *et al.*, 2010) localize to the nucleus regardless of the presence of ligand. Both these AR variants retain the NLS in the AR exon 4-encoded hinge region, necessary for nuclear translocation (Sun *et al.*, 2010; Hu *et al.*, 2011). However, recent data suggests AR variants such as AR-V7 which lack the canonical NLS, contain unique CTD sequences that reconstitute classical NLS activity and this is sufficient for nuclear localization and androgen-independent transcriptional activation of endogenous AR target genes (Chan *et al.*, 2012).

1.6.3 AR variant interaction with full-length AR

AR variants have been shown to be co-expressed with full-length AR in various cell lines, xenografts and clinical samples. Analysis of AR-V7 and AR-V1 transcript expression by semi-quantitative PCR in PCa metastases and several xenografts has shown that these AR variants are expressed at much lower levels (~0.1-2.5%) relative to full-length AR (Watson *et al.*, 2010). Therefore, the question whether AR variants interact with full-length AR is highly relevant. ARV567es has been shown to be constitutively transcriptionally active and nuclear in the absence of androgens (Sun *et al.*, 2010). Importantly, when full-length AR and HA-tagged AV567es are overexpressed in AR-null M12 cells, both full-length AR and ARV567es are immunoprecipitated with an anti-HA antibody (Sun *et al.*, 2010). This suggests

that ARV567es can functionally interact with full-length AR, perhaps resulting in increased stability of full-length AR proteins as well as causing nuclear localization of full-length AR in the absence of androgens (Sun *et al.*, 2010). However, full-length AR/AR-V7 complexes have not been detected in 22Rv1 cells using standard co-immunoprecipitation techniques (Tepper *et al.*, 2002; Guo *et al.*, 2009). Structural differences may account for the differences in interaction with full-length AR between ARV567es and AR-V7 due to AR-V7 lacking the hinge region encoded by AR exon 4 that is present in ARV567es. However, cell specificity cannot be ruled out. Furthermore, the tandem duplication of AR exon 3 in full-length AR in 22Rv1 cells which encodes an additional zinc finger, has not been detected in other clinical specimens and may present a confounding factor by impeding interactions that could potentially be different than that observed with the wild-type full-length AR.

Overexpression of AR-V7 in LNCaP cells which express full-length AR confers anchorage-independent growth in the absence of androgens (Watson *et al.*, 2010). This phenotype was reversed with treatment with the anti-androgen MDV3100 which binds directly to the LBD, suggesting that AR-V7 is dependent on full-length AR (Watson *et al.*, 2010), despite the observation that AR-V7 does not interact with full-length AR. More recently, it was observed that full-length AR is not required for AR-V7 transcriptional activity as demonstrated by silencing full-length AR with siRNA in LNCaP cells that overexpress AR-V7 (Hu *et al.*, 2011). Consistent with these data it has been shown that androgen-independent growth of 22Rv1 cells is not dependent on full-length AR by knocking down its expression using two siRNAs targeting AR exon 4 and exon 6 encoding the LBD (Mashima *et al.*, 2010). Taken together, these observations from two independent model systems suggest that at least some AR variants do not require full-length AR. However, further work to delineate the relationship between full-length AR and different AR variants is needed.

1.6.4 AR variant target genes

Correlating androgen-regulated genes with genes expressed in CRPC has been a challenge. One potential explanation comes from the recent discovery that the

gene expression program regulated by AR in androgen-dependent PCa cells is distinct from the one in androgen-independent cells (Wang *et al.*, 2009). Cell-specific and gene-specific transcription is thought to result from recruitment of different coregulatory proteins to the AR (Wang *et al.*, 2009). Therefore, it would be expected that truncated AR variants lacking regions of AR would be devoid of protein interfaces and/or have new interfaces as a result of the variable COOH-terminus sequences. As a result, it is possible that specific AR variants may mediate distinct transcriptional programs in CRPC.

One hypothesis is that AR variants simply substitute for androgens and activate an identical repertoire of full-length AR target genes. Microarray analysis performed on RNA isolated from LNCaP cells ectopically expressing AR-V7 cDNA found that AR-V7 induced canonical androgen-responsive genes such as PSA, KLK2, NKX3.1, FKBP5 and TMPRSS2 in the absence of androgens (Table 1.2; Hu *et al.*, 2009). Similarly, PSA, TMPRSS2 and FKBP5 induced by DHT treatment in cells expressing full-length AR were also induced by ARV567es expression in the absence of androgens (Sun *et al.*, 2010). However, the authors of this study also identified a subset of genes regulated by ARV567es that are unique and not influenced by androgens in context of full-length AR. Analysis of gene ontology (GO) terms enriched specifically in cells overexpressing ARV567es revealed that the GO molecular function "transcription factor activity" is significantly increased in the absence of androgen, signifying activation of other growth and survival pathways. Among the transcription factors upregulated by ARV567es that are known to induce a proliferative program of gene expression are STAT3 and JUN (Sun *et al.*, 2010).

In order to identify AR-V7-regulated genes, Guo and colleagues employed shRNA constructs to selectively knock down AR-V7 or full-length AR in CWR-R1 and 22Rv1 cells and analyzed differential gene expression using microarray analysis. The differential expression of 188 genes was consistently detected in both cell lines when AR-V7 was knocked down, whereas the expression of 412 genes was altered in both cell lines when full-length AR was specifically inhibited. Among them, 71 genes are commonly regulated by both full-length AR and AR-V7 (e.g. IGFBP3, FKBP5). However, many well characterized AR-regulated genes such as CLU, TMEPAI, PSA

Table 1.2 Expression profiling of AR variant genes

Study	Tissue Source	Experimental Conditions	Profiling Platform	GEO Accession
Hu *et al.*, 2009	LNCaP	Transfection of pcDNA-AR-V7 in 10% CSS for 24 hr ± 10 nM R1881 for 24 hr		
Guo *et al.*, 2009	CWR-R1, 22Rv1	Transfection of shAR3-1 (knocks down AR-V7) or shARa (knocks down full-length AR)	Agilent 4x44K (G4112F)	GSE13919
Sun *et al.*, 2010	LNCaP	Transfection of pcDNA-ARV567es or pcDNA-AR-FL in 10% CSS for 24 hr ± 1nM DHT for 24 hr	Agilent 4x44K	
Hornberg *et al.*, 2011	CRPC bone metastases (n=30)	Comparison of CRPC bone metastases with AR-V567es and/or AR-V7 transcript levels in the upper quartile compared to other CRPC bone metastases	Illumina HumanHT-12 V3.0	GSE29650
Hu *et al.*, 2012	LNCaP	Transduction of lenti-AR-V7 in 10% CSS for 24 hr ± 1 nM R1881 for 24 hr	Agilent 4x44K (G4112F)	GSE36549
	LNCaP	Transfection of pcDNA-AR-V7 or pcDNA-ARV567es in cells expressing stable clones of full-length AR positive and full-length AR negative cells in 10% CSS		
	LNCaP95	Transfection of siRNA targeting AR-LBD or AR-DBD for 24 h ± 1 nM R1881 for 48 hr		
	VCaP	Transfection of siRNA targeting AR-DBD or treatment with MDV3100 in the presence of 1 nM R1881		

and CLDN4 were not affected by knockdown of AR-V7. Among the genes preferentially regulated by AR-V7, a number of genes such as MAP4K4, HOXB7 and ELK1 have been found to be upregulated in CRPC (Guo *et al.*, 2009).

Recent data from gene expression profiling of CRPC bone metastases found 60 genes differentially expressed in bone metastases with high transcript expression of ARV567es and/or AR-V7 compared to other CRPC bone metastases. These genes included several gene transcripts known to be positively regulated by AR such as C-MYC, UBE2C, CDK1, CYCLINA2 and HSP27 but specifically not some classical androgen-regulated genes such as PSA, TMPRSS2 and FKBP5 (Hornberg *et al.*, 2011).

To dissect the transcriptional programs induced by AR variant-mediated signaling, Hu and colleagues (2012) identified two core sets of AR variant and full-length AR-regulated genes, by assessing the transcriptional changes in LNCaP cells overexpressing AR-V7 in the presence and absence of full-length AR signaling by microarray and gene set enrichment analysis. Transient expression of exogenous AR-V7 in LNCaP cells induced expression of cell cycle genes under both androgen-depleted and androgen-stimulated conditions. Whereas, the top gene sets induced by ligand-dependent full-length AR were related to biosynthesis, metabolism and secretion. The AR variant gene set was further characterized by overexpressing AR-V7 and ARV567es in stable clones of full-length AR positive and full-length AR negative LNCaP cells. Under these conditions, the AR variant gene set was induced by both AR-V7 and ARV567es and the absence of full-length AR did not attenuate induction of the AR variant gene signature. This AR variant gene set was also upregulated in LNCaP95 cells (which endogenously express AR-V7 as well as full-length AR) and also VCaP cells, following siRNA targeted knockdown of the AR-LBD. These data suggest that the presence of full-length AR is not required for the induction of cell cycle genes by AR-V7 and ARV567es (Hu *et al.*, 2012).

Therefore, it is likely that distinct gene expression profiles are mediated by AR variants versus full-length AR with some overlap. Clinically, these differences in gene expression profiles may be relevant to biomarker discovery to identify patients with tumors that express AR variants that would not respond to therapies that target the LBD.

1.6.5 Mechanism of AR variant production

The first mechanism to account for the production of naturally occurring truncated AR variants in the CWR22 models was calpain-mediated proteolysis (Libertini *et al.*, 2007; Chen *et al.*, 2010). However, contradictory data also exists (Sivanandam *et al.*, 2011). More recently an alternative mechanism for the generation of AR variants involving intragenic rearrangement within the AR locus within a region that is downstream of AR exon 2 and upstream to AR exon 4 has been reported in 22Rv1 cells (Li *et al.*, 2011). Interestingly, long term culture of androgen-dependent CWR22Pc cells, which do not express AR variants or display intragenic rearrangement, in androgen-depleted media, leads to concurrent emergence of both (Li *et al.*, 2011). These data suggest that AR intragenic rearrangements in CRPC may be associated with pathologic expression of AR variants in disease progression. Further intragenic rearrangements have been identified in the LuCaP 86.2 xenograft which expresses ARV567es. Using deletion-spanning PCR, an 8,579 bp deletion of AR exons 5, 6 and 7 in the LuCaP 86.2 xenograft was discovered, which provides a rational explanation for the synthesis of the truncated ARV567es variant (Li *et al.*, 2012). The association of these genetic events with enhanced expression of AR variants in a large cohort of clinical specimens is required to determine the clinical importance of this mechanism. Nevertheless, a genomic basis for the pathologic expression of AR variants may serve as a stable mechanism-based marker for resistance to ADT.

1.6.6 Implications for treatment of CRPC

Currently, all hormonal therapies used in the clinic target the AR LBD directly with anti-androgens or indirectly by reducing circulating levels of androgens. Although this approach is initially effective, it is unclear why these approaches eventually fail. Compelling evidence from the last five years suggest that truncated AR variants lacking the LBD may be an important mechanism in disease progression to CRPC. AR variants retain the domain of the receptor that harbors the main transcriptional transactivation function (i.e. AF-1); therefore there is a strong rationale for developing compounds that specifically bind the AR NTD to block

activity. The first indication that the NTD is a viable target for *in vivo* intervention in CRPC was demonstrated using a decoy molecule encoding residues 1 to 558 of the AR NTD (Quayle *et al.*, 2007). One particular challenge in developing NTD antagonists is the fact that this region of the receptor is intrinsically disordered and therefore not amenable to crystallographic structural determination, which in turn precludes virtual screening of small molecule libraries (Sadar, 2011). Recent studies have overcome this problem using cell-based screening protocols leading to the development of EPI-001, a potent and specific AR NTD antagonist (Andersen *et al.*, 2010). Impressive data from preclinical studies has shown EPI-001 treatment leads to cytoreduction of CRPC in xenografts dependent on AR for growth and survival without causing toxicity (Andersen *et al.*, 2010). Analogues of EPI-001 are currently in development for phase 1 clinical trials (Sadar, 2012). These significant findings bring forth a new era in targeting AR and the first to definitively demonstrate the feasibility for targeting the AR NTD *in vivo*.

1.7 Objectives

The objectives of this thesis are: (1) to assess the levels of a number of AR variants by absolute quantification real-time RT-PCR in a panel of androgen-dependent and androgen-independent human PCa cell lines; (2) to demonstrate the constitutive transcriptional activity of AR variants using shRNA constructs targeting both full-length AR and AR variants and full-length AR alone in more than one cell-based model of CRPC; (3) to assess the transcriptional activity of individual AR variants and also investigate the transcriptional activity of AR variants independent of full-length AR; (4) to identify a unique AR variant gene signature, independent of full-length AR using a microarray-based approach in a cell-based model of CRPC; and (5) to determine the clinical significance of the AR variant gene signature in a cohort of PCa specimens.

Chapter 2

Materials and Methods

2.1 Materials

Acetic acid	EMD
Actin antibody	Cell Signaling
Activated charcoal powder	Mallinckrodt
Adenosine-5′-triphosphate	Sigma
Agarose	Sigma
Ampicillin	Sigma
Antibiotic-Antimycotic	Invitrogen
Anti-mouse IgG HRP	GE Healthcare
Anti-rabbit IgG HRP	GE Healthcare
Aprotinin	Roche
AR (N-20) antibody	Santa Cruz
BglII	New England BioLabs
Bovine serum albumin	Sigma
Bromophenol blue	Sigma
Calcium chloride	Sigma
CellTiter 96 Aqueous Cell Proliferation Assay	Promega
Chloroform	Fisher
DC Protein Assay kit	Bio-Rad
DEPC-water	Ambion
Dextran T70	Pharmacia
DH5α	Invitrogen
Dithiothreitol	Sigma
DNase I	Invitrogen
Dulbecco's Modified Eagle Medium	Cellgro
Dynabeads	Invitrogen
EDTA	Sigma
Ethanol	Invitrogen
Fetal bovine serum	Invitrogen
Formaldehyde	Fisher
GAPDH PCR primers	Applied Biosystems

Glycerol	Fisher
HEPES	Sigma
HindIII	New England BioLabs
HyBlot CL autoradiography film	Denville
Isopropanol	Sigma
Leupeptin	Roche
Lipofectamine 2000	Invitrogen
Luciferase assay reagent	Promega
Luria broth	Becton Dickinson
Magnesium chloride	Sigma
Mercaptoethanol	Bio-Rad
Methanol	Sigma
Nitrocellulose membrane	Bio-Rad
Non-fat dry milk	Bio-Rad
Novex Sharp Protein Standard	Invitrogen
NP-40	Sigma
NuPAGE LDS sample buffer	Invitrogen
NuPAGE MOPS SDS running buffer	Invitrogen
NuPAGE Novex minigels	Invitrogen
NuPAGE sample reducing agent	Invitrogen
NuPAGE transfer buffer	Invitrogen
Opti-MEM	Invitrogen
pFRT-H1p	Invitrogen
Phenylmethylsulphonyl fluoride	Sigma
Phosphate-buffered saline	Invitrogen
Plasmid Plus Mega kit	Qiagen
Potassium chloride	Sigma
QIAquick Gel Extraction kit	Qiagen
R1881	Sigma
Reporter lysis buffer	Promega
RNeasy kit	Qiagen

RPMI-1640	Invitrogen
Silver nitrate	Sigma
Sodium carbonate	Sigma
Sodium chloride	Sigma
Sodium dodecylsulphate	Bio-Rad
Sodium thiosulphate (pentahydrate)	Sigma
Standard grade No. 3 filter paper	Whatman
SuperScript III First-Strand Synthesis System	Invitrogen
SuperSignal West Dura Chemiluminescent Substrate	Pierce
SYBR Green PCR Master Mix	Applied Biosystems
T4 polynucleotide kinase	New England BioLabs
TAE buffer	Bio-Rad
Terrific broth	Invitrogen
Tris-HCl	Sigma
TRIzol reagent	Invitrogen
Trypsin/EDTA	Invitrogen
Tween-20	Sigma

2.2 Cell culture

2.2.1 Origin of cell lines

22Rv1, LNCaP, PC-3, DU-145, VCaP and COS-7 cells were obtained from the American Type Culture Collection (ATCC). C4-2 and C4-2B cells (Wu *et al.*, 1994) were a kind gift from Dr. Leland W. K. Chung (Cedars-Sinai Medical Center, Los Angeles, California).

2.2.2 Maintenance of cell lines

22Rv1, PC-3, DU-145, COS-7, C4-2 and C4-2b cells were grown in RPMI-1640 containing, 0.01 mM phenol red, 2 mM glutamine, supplemented with 10% (v/v) fetal bovine serum (FBS; Invitrogen), 10 U/ml penicillin and 10 μg/ml streptomycin. LNCaP cells were maintained in identical conditions without phenol red. VCaP cells were grown in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L

glucose, supplemented with 10% FBS, 10 U/ml penicillin and 10 μg/ml streptomycin. Cells were maintained in a humidified 37°C /5% CO_2 incubator. To sub-culture adherent cells, the medium was aspirated; the monolayer washed twice with phosphate-buffered saline (PBS; pH 7.4) then 0.25% trypsin/EDTA solution (1 ml per 25 cm^2) was added and incubated at 37°C until cell detachment. Upon detachment the cells were resuspended in the appropriate culture medium and transferred to tissue culture plates. Cells were monitored daily using a Leitz inverted microscope (x10 magnification) for morphology and growth characteristics. Cells were discarded once they reached passage number 40.

2.3 Plasmids

2.3.1 shRNA AR suppression constructs

2.3.1.1 Oligonucleotide design

The human AR gene sequence was analyzed for potential shRNA targets based on the predicted secondary structure of the mRNA transcript (Brodsky *et al.*, 1995). Energetically favorable stems were predicted from the input sequence by the GeneBee program (http://www.genebee.msu.su/services/rna2_reduced.html). Target oligonucleotide sequences for AR exon 1 (CAAGGGAGGTTACACCAAA) and AR exon 8 (TCAGTTCACTTTTGACCTGCTAATCA) were identified (Appendix I, Figure I.1). Each of the sequences was modified to contain a BglII site (5'-AGATCT-3'), a HindIII site (5'-AAGCTT-3') and 5' phosphate group. The oligonucleotide strands were synthesized by Integrated DNA Technologies and dissolved in nuclease-free water and stored at -80°C.

2.3.1.2 Phosphorylation and annealing of oligonucleotides

30 pmol of each complementary single-stranded oligonucleotide was phosphorylated with 1 μl T4 polynucleotide kinase, 2 μl kinase buffer (70 mM Tris-HCl, pH 7.6, 10 mM MgCl2, 5 mM DTT), 2 μl 10mM ATP and 14 μl water for 60 min at 37°C. The oligonucleotides were annealed with annealing buffer (200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 500 mM NaCl) for 10 min at 70°C and then cooled to

room temperature over 60 min. The annealed inserts were stored at -20°C until needed.

2.3.1.3 Preparation of vector

The pFRT-H1p vector (Invitrogen) was linearized by sequential digestion with HindIII and BglII enzymes (New England BioLabs). 1 µl of vector was incubated with 2 x 1 µl of each restriction enzyme, 4 µl of NEBuffer #2 (50 mM NaCl, 10 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 1 mM DTT) and 33 µl of water. The digest was performed at 37°C for 60 min and then heat inactivated at 65°C for 20 min. The linearized vector was purified by separating the DNA on a 0.7% (w/v) agarose gel in TAE (Tris Acetic EDTA) buffer (Bio-Rad), excised from the gel and extracted using the QIAquick Gel Extraction Kit (Qiagen). The purified linearized vector was stored at -20°C until needed.

2.3.1.4 Ligation of vector and insert

The vector DNA was dephosphorylated followed by ligation of the vector and annealed inserts using the Rapid DNA Dephos & Ligation Kit (Roche). 1 µg of vector DNA was dephosphorylated with 1 µl alkaline phosphatase, 2 µl alkaline phosphatase buffer and 4 µl water for 30 min at 37°C, followed by 2 min at 75°C to inactivate the enzyme. 50 ng of vector DNA and 150 ng of each annealed insert were ligated for 5 min at room temperature with T4 DNA ligase and stored at -20°C until needed.

2.3.1.5 Bacterial transformation

3-5 µl of the ligation reaction was added to 50 µl DH5α competent cells (Invitrogen), mixed gently and incubated on ice for 30 min. The cells were then heat shocked at 42°C for 45 sec, followed by a further 2 min on ice. 1 ml Luria broth (LB; Becton Dickinson) medium was added and incubated in an orbital shaker (225 rpm) for 1 hr at 37°C. 200 µl of the transformation reaction was spread evenly onto LB agar plates containing 100 µg/ml ampicillin and incubated at 37°C overnight. A single colony was used to inoculate 3 ml LB medium containing 100 µg/ml ampicillin, and

grown for 6 hr at 37°C. The resulting bacterial culture was used to inoculate 500 ml Terrific broth (Invitrogen) containing 100 μg/ml ampicillin and incubated for a maximum of 16 hr at 37°C.

2.3.1.6 Plasmid DNA purification, quantification and sequencing

Bacterial cultures were harvested by centrifugation at 5,000 x g for 15 min in a Beckman J-6B refrigerated centrifuge and the plasmid DNA purified using the Qiagen Plasmid Plus Mega Kit. The purified DNA was dissolved in 300-500 μl TE (Tris EDTA) buffer. Plasmid DNA concentration was determined by measuring the absorbance at 260 nm on a Tecan GENios microplate spectrophotometer. The plasmid DNA was sequenced at the Mayo Clinic DNA Sequencing Core Facility on an ABI Prism 3730xl DNA Analyzer (Applied Biosystems).

2.3.2 Expression plasmids

2.3.2.1 p5HBhAR-A

The full-length AR expression vector (p5HBhAR-A) was a kind gift from Dr. Frank S. French (University of North Carolina at Chapel Hill) and constructed by cloning full-length human AR cDNA into the pCMV5 expression vector (Simental *et al.*, 1991).

2.3.2.2 hAR Ex 1/2/2b and hAR Ex 1/2/3/2b

The AR-V3 and AR-V4 expression plasmids (hAR Ex 1/2/2b and hAR Ex 1/2/3/2b respectively) were constructed by Dr. Scott M. Dehm (Department of Urology, Mayo Clinic) by mutating p5HBhAR-A (Simental *et al.*, 1991) to generate a restriction site within exon 2 or exon 3, followed by synthesis of a cassette containing the exon 2b sequence and insertion downstream of the mutated restriction site in exon 2 and exon 3 respectively (Dehm *et al.*, 2008).

2.3.2.3 pcDNA-AR-V1 and pcDNA-AR-V7

The pcDNA-AR-V1 and pcDNA-AR-V7 expression plasmids were a kind gift from Dr. Jun Luo (Johns Hopkins Hospital, Baltimore, Maryland) and constructed by

cloning the PCR products derived from primers targeting the AR-V1 and AR-V7 full-length open reading frame in 22Rv1 cells into the TopoTA vector (Invitrogen; Hu *et al.*, 2009).

2.3.2.4 MMTV-Luc

The mouse mammary tumor virus (MMTV)-luciferase (Luc) expression plasmid was a kind gift from Dr. Frank Claessens (Catholic University of Leuven, Belgium) and contains MMTV long terminal repeats that drives the transcription of the firefly luciferase reporter gene in response to activation of several nuclear receptors including AR.

2.3.2.5 pGL3-4XARE-E4-Luc

The pGL3-4XARE-E4-Luc luciferase reporter plasmid was a kind gift from Dr. Michael F. Carey (University of California at Los Angeles) and contains four tandem copies of the ARE of the PSA gene upstream of the adenovirus E4 core promoter and the luciferase gene.

2.3.2.6 Plasmid DNA transformation, purification, quantification and sequencing

All plasmid DNA sequences were confirmed using restriction enzymes and sequencing at the Mayo Clinic DNA Sequencing Core Facility on an ABI Prism 3730xl DNA Analyzer (Applied Biosystems). Plasmid DNA was transformed, purified and quantified as described previously (see Section 2.3.1.6).

2.4 Transient transfection of plasmid DNA

2.4.1 Electroporation

22Rv1 and LNCaP cells were transfected with plasmid DNA by electroporation. Adherent cells were washed twice with PBS (pH 7.4) then 0.25% trypsin/EDTA solution (1 ml per 25 cm2) was added and incubated at 37°C until cell detachment. Upon detachment the cells were resuspended in fresh medium and the cell count determined using a Bright-Line hemacytometer (Sigma). An appropriate volume of suspended cells was pelleted at 3,000 rpm in a clinical centrifuge,

resuspended in fresh media and aliquoted (400 μl) into 1.5 ml microcentrifuge tubes. 4 μg shRNA plasmid or 2 μg expression plasmid DNA per 10^6 cells was incubated with the cell suspension for 15 min at room temperature. The cell suspension was transferred to an electroporation cuvette (4 mm gap) and electroporated using a BTX ECM 830 electroporator (BTX-Harvard Apparatus) at either 350V (for 22Rv1 cells) or 305V (for LNCaP cells) for 10 msec. Cells were allowed to recover for 20 min before being plated in fresh medium. Cells transfected with shRNA constructs were harvested after 72 hr and cells transfected with expression constructs were harvested after 48 hr.

2.4.2 Liposome-based transfection

VCaP and COS-7 cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen). Lipofectamine 2000 is a proprietary cationic liposome-based transfection reagent. 2 x 10^6 COS-7 cells were plated in 6-well tissue culture plates in 2 ml media containing no antibiotics 24 hr prior to transfection. In the case of VCaP cells, 2 x 10^6 cells were plated in 6-well tissue culture plates in 2 ml media containing no antibiotics 72 hr prior to transfection. 1 μg plasmid DNA was diluted in 250 μl reduced serum medium (Opti-MEM, Invitrogen) and incubated for 5 min at room temperature. Lipofectamine was diluted at a DNA (μg):Lipofectamine (μl) ratio of 1:2 for COS-7 cells and a ratio of 1:5 for VCaP cells, in 250 μl reduced serum medium and incubated for 5 min at room temperature. The diluted DNA and diluted Lipofectamine solutions were combined, mixed and incubated for 20 min at room temperature. 100 μl DNA/Lipofectamine complexes was added directly to the medium in each well. Media was replaced after 6 hr and harvested at 48-72 hr depending on the experiment.

2.5 Androgen withdrawal and stimulation *in vitro*

In certain experiments, cells were grown in androgen-depleted conditions by culture in media supplemented with 10% (v/v) charcoal-stripped serum (CSS). CSS was prepared by mixing 5% (w/v) activated charcoal (Mallinckrodt) and 0.5% (w/v) dextran T70 (Pharmacia) in 1 mM HEPES buffer (pH 7.4) for 3 hr at 4°C. The dextran-

coated charcoal was centrifuged at 500 x g for 10 min to form a pellet. The dextran-treated charcoal was mixed with FBS (1% w/v) for 1 hr at room temperature and then centrifuged at 500 x g for 10 min to pellet the charcoal. This procedure was repeated twice with fresh dextran-treated charcoal followed by sterile-filtration through a 0.22 μm pore membrane (Millipore).

Cells were stimulated *in vitro* with a physiological concentration of R1881 (Sigma). R1881 (17β-17-hydroxy-17-methyl-estra-4,9,11-trien-3-one), also known as metribolone is a stabile, highly potent, synthetic androgen (Brinkmann *et al.*, 1986). Cells were cultured in media without phenol red, supplemented with 10% CSS and treated with 1 nM R1881 dissolved in ethanol for 24 hr. Previous studies have demonstrated that androgen regulated genes are maximally expressed at 24 hr of androgen exposure (Dehm and Tindall, 2006)

2.6 Protein Expression

2.6.1 Preparation of whole-cell lysates

Cells were washed twice with ice-cold PBS and then lysed with ice-cold radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% SDS) containing the protease inhibitors aprotinin (1 μg/ml), leupeptin (5 μg/ml) and PMSF (1 mM). Lysed cells were scraped with a cell scraper, transferred to a 1.5 ml microcentrifuge tube and disrupted by sonication for 15 sec with a Branson Sonifier Cell Disruptor 200. Lysates were then cleared by centrifugation at 16,000 x g for 10 min at 4°C.

2.6.2 Protein quantification

Protein concentration of cell lysates was determined using a modified Lowry assay (Lowry *et al.*, 1951) in a 96-well microplate (Nunc) with the Bio-Rad DC Protein Assay kit. Standards were prepared from a stock solution of 1.5 mg/ml bovine serum albumin (BSA), serially diluted in RIPA buffer from 1.5 mg/ml to 23.4 μg/ml and also a blank. 25 μl working Reagent A was added to duplicate lysate samples (5 μl) and standards (5 μl), followed by 200 μl Reagent B. After 15 min at room temperature, absorbances were read at 750 nm on a Tecan GENios microplate spectrophotometer.

A regression line was plotted and the concentration of protein was calculated and converted to mg/ml.

2.6.3 Co-immunoprecipitation

Co-immunoprecipitation was performed using the superparamagnetic Dynabeads system, covalently coupled with recombinant Protein A (Invitrogen). Whole-cell lysates were prepared in RIPA buffer (see Section 2.6.1). 50 μl resuspended Dynabeads were incubated with 1 μg antibody in 200 μl binding buffer (PBS, pH7.4, 0.02% Tween-20) for 1 hr at 4°C. The antibody-bound beads were separated on a DynaMag-2 magnet and washed with 200 μl wash buffer (PBS, pH7.4, 0.02% Tween-20). 1-1.5 μg whole-cell lysate was incubated with the beads-antibody complex by rotation for 4 hr at 4°C to allow antigen to bind. The beads-antibody-antigen complex was then washed three times with 200 μl wash buffer. Target antigen was eluted from the beads with 20 μl elution buffer (50 mM glycine, pH 2.8) and 10 μl premixed NuPAGE LDS Sample Buffer (Invitrogen) and NuPAGE Sample Reducing Agent (Invitrogen) and heated to 70°C for 10 min. The beads were separated and the supernatant loaded on to pre-cast NuPAGE Novex mini gels (Invitrogen).

2.6.4 SDS-polyacrylamide gel electrophoresis

Electrophoresis was performed under denaturing conditions using pre-cast NuPAGE Novex mini gels (8 x 8 cm) with the XCell SureLock Mini-Cell electrophoresis unit (Invitrogen). Equal amounts of protein (ranging from 10-50 μg depending on the experiment) were denatured in sample buffer (250 mM Tris-HCl, pH 6.8, 30% (v/v) glycerol, 10% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.02% (w/v) bromophenol blue) and heated to 95°C for 5 min. Acrylamide/bisacrylamide separation gels (1.5 mm) with a monomer concentration of 4-12% overlaid with 4% stacking gel were used. Proteins were loaded and separated by application of 175V in MOPS running buffer (50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7) for 90 min.

2.6.5 Semi-wet electrophoretic blotting of proteins

Following electrophoresis, the gel was removed from the cassette and placed on top of a sheet of nitrocellulose membrane (0.2 μm pore size; Bio-Rad), soaked in transfer buffer (25mM Tris-base; 192mM glycine; 10% methanol, 0.05% SDS) and cut to the size of the gel. Filter paper (Standard Grade No. 3, Whatman) was placed on top and beneath the nitrocellulose/gel forming a sandwich. Air bubbles were removed from the sandwich by gently passing a glass pipette over it. The sandwich was soaked in transfer buffer and placed between pre-soaked blotting pads on an XCell II Blot Module (Invitrogen). A constant voltage of 30V was applied for 120 min.

2.6.6 Western immunoblotting

Following transfer, the membrane was blocked for non-specific binding for 1 hr at room temperature in 5% (w/v) non-fat dry milk (Bio-Rad) in Tris-buffered saline (TBS; 20 mM Tris-HCl, pH 7.5, 500 mM NaCl) containing 0.05% (v/v) Tween-20 (TBS-T). Membranes were probed with primary antibodies in 5% milk/TBS-T (1:1,000-5,000), washed three times in TBS-T (5 min each) and incubated with the appropriate horseradish peroxidase conjugated secondary antibodies (1:2,000-10,000) for 1-2 hr. The membranes were again washed three times in TBS-T (15 min each). All washes and incubations were performed with gentle agitation. Proteins were visualized using a chemiluminescent detection reagent (SuperSignal West Dura Chemiluminescent Substrate) followed by autoradiography using HyBlot CL autoradiography film (Denville) and a Kodak M35A X-OMAT processor.

2.6.7 Protein identification

2.6.7.1 Silver staining

Following electrophoresis (see Section 2.6.4), the gel was removed from the cassette and fixed in 50% methanol/10% acetic acid (v/v) over night. The fixed gel was incubated in 5% (v/v) methanol for 15 min and washed three times (5 min each) with ultrapure water (Milli-Q). The gel was sensitized with 0.02% (w/v) sodium thiosulphate for 2 min and washed twice with ultrapure water (30 sec each). The gel was incubated with 0.2% (w/v) silver nitrate solution for 30 min and washed twice

with ultrapure water (1 min each). The gel was developed with 3% (w/v) sodium carbonate containing 0.05% (v/v) formaldehyde and terminated with 1% (w/v) sodium-EDTA once an appropriate staining intensity had been achieved. The gel was washed with ultrapure water and stored in 1% (v/v) acetic acid at 4°C prior to mass spectrometry.

2.6.7.2 Mass spectrometry

Silver-stained gel bands were excised using a clean scalpel, destained and digested with trypsin. The tryptic peptides were identified using nanoHPLC-electrospray tandem mass spectrometry (nanoLC-ESI-MS/MS) performed at the Mayo Clinic Proteomics Core Facility. Identified peptides were compared with the Swiss-Prot protein sequence database.

2.7 Luciferase Assay

Firefly luciferase is a monomeric 61 kDa protein which catalyzes luciferin oxidation using ATP•Mg^{2+} as a co-substrate to produce oxyluciferin and light. Quantification of firefly luciferase reporter activity was performed using the Luciferase Assay System (Promega). Cells were transfected with an expression construct containing luciferase under the control of specific promoters and plated in 6-well tissue culture plates (see Section 2.4). Culture media was aspirated from the cells to be assayed and washed twice with PBS (pH 7.4). Cells were lysed in 200 µl reporter lysis buffer (a proprietary formulation of bicine buffer and Tween; Promega) and incubated with gentle agitation for 10 min at room temperature. Attached cells were scraped with a cell scraper into a 1.5 ml microcentrifuge tube and vortexed for 10-15 sec. Cell lysates were subjected to a freeze-thaw cycle to achieve complete lysis, followed by protein quantification (see Section 2.6.2). The lysate was assayed by dispensing 10 µl of cell lysate (in triplicate) to an opaque 96-well plate (Berthold). A Centro LB 960 microplate luminometer (Berthold) was programmed to inject 50 µl luciferase assay reagent (Promega), followed by a delay of 2 sec and a read time of 10 sec. Luciferase activity was expressed as relative luciferase units per mg protein (RLU/mg).

2.8 Cell viability assay

Cell viability was determined using the CellTiter 96 Aqueous Cell Proliferation Assay (Promega). This is a colorimetric method for determining the number of viable cells in proliferation using MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt], which is converted to formazan by dehydrogenase enzymes found in metabolically active cells. The quantity of formazan product is directly proportional to the number of living cells in culture. Cells were plated in 96-well tissue culture plates (5,000 cells/well in 100 μl medium). At appropriate time points, 20 μl of the combined MTS/PMS solution was added to each well and incubated for 2 hr at 37°C in a humidified, 5% CO_2 atmosphere. To determine the amount of soluble formazan produced, absorbance at 490 nm was measured using a Tecan GENios microplate spectrophotometer.

2.9 Gene expression

2.9.1 Real-time RT-PCR

2.9.1.1 Isolation of total RNA

Cells were washed twice in ice-cold PBS (pH 7.4) and 1 ml TRIzol reagent (Invitrogen) was added directly to the cells. Cells were scrapped into 1.5 ml microcentrifuge tubes and incubated for 5 min at room temperature. Phase separation was performed by adding 200 μl chloroform and centrifuging at 9,500 x g for 15 min at 4°C. RNA was precipitated from the aqueous phase over night at 4°C with 500 μl isopropanol and then recovered by centrifuging at 9,500 x g for 10 min at 4°C. The pellet was washed with 1 ml 75% (v/v) ethanol, centrifuged at 5,000 x g for 5 min at 4°C, air dried for 5-10 min and then resuspended in diethylpyrocarbonate (DEPC) treated water. RNA concentration and purity was determined on a Tecan GENios microplate spectrophotometer. Those samples with an absorbance $A_{260/280}$ ratio <1.7 were subjected to an additional purification procedure using the Qiagen RNeasy kit. All RNA preparations were further purified with deoxyribonuclease (DNase) to degrade any remaining genomic DNA. 3 μg RNA was incubated with 1 μl DNase I (1 U/μl), 1 μl DNase I reaction buffer (200 mM Tris-HCl, pH 8.4, 20 mM

$MgCl_2$, 500 mM KCl) for 15 min at room temperature. The enzyme was heat inactivated with 1 µl 25 mM EDTA at 65°C for 10 min.

2.9.1.2 cDNA synthesis from total RNA

cDNA was synthesized from total RNA using the SuperScript III First-Strand Synthesis System (Invitrogen). 3 µg of total RNA was incubated with 1 µg Oligo$(dT)_{20}$ primer and 1 µl 10 mM dNTP mix for 5 min at 65°C and then on ice for 1 min. The cDNA synthesis reaction was composed of the following: reverse transcriptase buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 25 mM $MgCl_2$, 0.1 M DTT, RNase (40 U/µl) and SuperScript III reverse transcriptase (200 U/µl). The cDNA synthesis reaction was incubated with the RNA/primer mixture for 50 min at 50°C. The reaction was terminated at 85°C for 5 min and then placed on ice for 5 min. 1 µl RNase H (2 U/µl) was incubated with the reaction for 20 min at 37°C. cDNA was stored at -20°C until needed.

2.9.1.3 Primer design

All primers were designed using Primer3 (http://primer3.wi.mit.edu, Rozen and Skaletsky, 2000) and synthesized by Integrated DNA Technologies. A complete list of primer sequences can be found in Appendix I (Table I.1). GAPDH control primers were obtained from Applied Biosystems. All primers were dissolved in nuclease-free water to a final concentration of 5 µM and stored at -20°C.

2.9.1.4 Semi-quantitative real-time RT-PCR

Real-time PCR was performed using the SYBR Green PCR Master Mix system (Applied Biosystems) in MicroAmp optical 96-well reaction plates (Applied Biosystems) on an ABI Prism 7900HT Fast Real-Time PCR system, coupled to SDS v2.4 software (Applied Biosystems) at the Mayo Clinic Gene Expression Core Facility. The real-time PCR reaction mix was as follows: 1 µl (150 ng) cDNA template, 4.5 µl forward and reverse primers (final concentration 900 nM), 2.5 µl nuclease-free water and 12.5 µl SYBR Green PCR Master Mix. Thermal cycling (40 cycles) was performed using the following conditions: 50°C for 2 min, 95°C for 10

min, 95°C for 15 sec (denature) and 60°C for 1 min (anneal/extend). Data was extracted from the SDS software and normalized to GAPDH. Relative expression levels were determined by the comparative C_T method using the formula $2^{-\Delta\Delta CT}$ where C_T is the threshold cycle of amplification. PCR products (120-150 bp) were separated on a 1.2% (w/v) agarose gel in TAE to check for nonspecific amplification. Each experiment was performed in triplicate and each replica incorporated three technical repeats.

2.9.1.5 Absolute-quantitative real-time RT-PCR

Absolute-quantitative real-time RT-PCR with serial dilutions of plasmids harboring full-length AR, AR-V3, AR-V4, AR-V1 and AR-V7 was performed using the SYBR Green PCR Master Mix system (see Section 2.9.1.4). Linearized plasmid was quantified using a spectrophotometer and standards were serially diluted in nuclease-free water (from 10^{-3} to 10^{-15} g/µl). C_T values these standards were used to generate C_T versus cDNA concentration standard curves. C_T values from RNA samples were obtained via quantitative real-time RT-PCR and were extrapolated from these standard curves to generate absolute values for copy number using the following formula:

$$Copy\ number/\mu g = \frac{6.022\ x\ 10^{23}(molecules/mole)\ x\ DNA\ concentration\ (g/\mu l)}{Plasmid\ length\ (bp)\ x\ 660\ Da}$$

2.9.2 Whole-genome microarray

2.9.2.1 Total RNA preparation and quality assessment

Total RNA was isolated using TRIzol and further purified on Qiagen RNeasy columns (see Section 2.9.1.1). RNA concentration and purity was determined on a Tecan GENios microplate spectrophotometer. All RNA samples had an absorbance $A_{260/280}$ ratio greater than 1.9. 2 µl total RNA underwent further quality assessment on an Agilent 2100 Bioanalyzer to ensure all total RNA samples used for microarray had an RNA Integrity Number (RIN) greater than 7.5.

2.9.2.2 RNA amplification and hybridization

Prior to hybridization to the array, 500 ng (1 μg/ul) total RNA was amplified using the Illumina TotalPrep RNA Amplification Kit at the Mayo Clinic Gene Expression Core Facility. Double-stranded cDNA template was synthesized from total RNA, followed by *in vitro* transcription generating biotin-labeled cRNA. cRNA was then assembled onto the Illumina HumanHT-12 v4 Expression BeadChip array platform. This platform contains 47,323 probes and 34,694 targets (transcripts) spanning the entire human genome. Either two or three replicates were used and hybridized on to a single array chip. Visualization of the array was performed by staining with streptavidin-Cy3 and scanned using an Illumina BeadArray Reader.

2.9.2.3 Data analysis

Analysis of the raw array data was performed by Dr. Zhifu Sun, Division of Biomedical Statistics & Informatics, Mayo Clinic. Raw array data was imported into the software R and background corrected using R packages *Bioconductor* and *gcrma* (Wu *et al.*, 2004). The intensity data from the perfect-match probes was normalized using *fastlo* (Ballman *et al.*, 2004). Data from all samples were normalized together. The probe-level data for each probe set was summarized using Tukey's median polish (Irizarry *et al.*, 2003). These summarized probe set values represent a measure of expression for the corresponding gene. The $\log_2$ transform of the probe set expression values were then used in all subsequent analysis. The R library *limma* was used to assess differential expression between the different treatment groups. The function eBayes, which uses an empirical Bayes method by Smyth (2004) to shrink the gene-wise sample variances towards a common value, was used for group comparisons. The false discovery rate (FDR; Benjamini and Kochberg, 1995), which is the expected proportion of false discoveries amongst the rejected hypotheses, was used at a 0.05 level. To account for both statistical and biological significance, genes with a FDR < 0.05 and an absolute value of the $\log_2$ transform of the fold change-ratio > 1 in either direction were considered candidate differentially expressed genes. To be considered different a probe set had to be considered statistically different with a minimum log-fold change of more than 1.0 or less than -1.0. This implies at least a

100% increase or 50% decrease in expression, respectively. These lists of probe sets were then used to build heatmaps using the dChip software package where distance metric used was 1 minus correlation and linkage method was centroid. Each row on the heatmap represents a probe set; each column represents an individual sample. Gene expression on each probe set was standardized to the mean of samples where red is higher than the mean and green is lower than the mean.

2.9.2.4 Network generation and functional analysis

Data from the microarray study containing gene identifiers and corresponding expression values were analyzed using Ingenuity Pathway Analysis (IPA; http://www.ingenuity.com; Ingenuity Systems). Each identifier was mapped to its corresponding object in the Ingenuity Knowledge Base. Network Eligible molecules were overlaid onto a global molecular network developed from information contained in the Ingenuity Knowledge Base. Networks of Network Eligible Molecules were then algorithmically generated based on their connectivity. The Functional Analysis identified the biological functions that were most significant to the data set. Right-tailed Fisher's exact test was used to calculate a p value determining the probability that each biological function assigned to the data set is due to chance alone.

2.9.3 Patient specimens

Surgical specimens were obtained with written patient consent from the Mayo Clinic/NIH Specialized Program of Research Excellence (SPORE) in Prostate Cancer tumor bank with institutional review board approval (#1937-00). Ninety-five prostate tissue specimens representing normal epithelium (n=17), benign prostatic hyperplasia (BPH) (n=10), prostate cancers of Gleason Pattern (GP) 3 (n=31), GP4 (n=20), GP5 (n=10) and lymph node (LN) metastasis (n=7) were laser capture microdissected (LCM) from 81 individual treatment-naïve patients who underwent prostatectomy at Mayo Clinic between 1999 and 2004. Eighteen patients had an additional specimen other than primary prostate cancer tissue. Clinical characteristics for all patients can be found in Appendix III (Tables III.2 and III.3). A

detailed description of the procedure for tissue procurement, RNA isolation and generation of microarray data for this cohort has previously been reported (Kube *et al.*, 2007). Briefly, fresh prostatectomy tissue was flash frozen in liquid nitrogen and stored at -80°C. Tissue sections were cut at 5 μm in a -20°C cryostat and stained using the HistoGene LCM Frozen Section Staining Kit (Applied Biosystems). Immediately after staining a tissue section, LCM was performed by a consultant pathologist using a PixCell II instrument (Arcturus). Thirty minutes on average (± 15 min) were spent collecting cells from a tissue section. Extracts from homogeneous populations of cells dissected on multiple captures from one to six serial sections of each case were combined onto a single RNA purification column to achieve a sufficient yield and concentration of RNA for linear amplification. Since the starting amount of total RNA was only 2 ng, a modified cDNA synthesis and cRNA biotin-labeling protocol was used (Affymetrix). 13 μg cRNA was hybridized to Affymetrix U133 Plus 2.0 arrays, washed, stained and scanned according to the Affymetrix protocol.

2.9.3.1 Data analysis

Analysis of the raw array data was performed by Dr. Zhifu Sun, Division of Biomedical Statistics & Informatics, Mayo Clinic. The raw data files were pre-processed similarly to those described in Section 2.9.2.3. Differential expression analysis for the 297 probe sets between combined malignant samples (primary and metastatic prostate cancer, n=68) and benign (BPH and normal prostate, n=27) samples was conducted using parametric t test. A p value less than 0.05 was used to define a differentially expressed gene between the two conditions. The differentially expressed genes between each of malignant cohorts (GP3, GP4, GP5 and LN metastases) and normal samples were identified using R package *limma* as described in Section 2.9.2.3. Nominal p value less than 0.05 was used as the cut-off for significantly changed genes for this validation. Similarly, the differentially expressed genes between paired malignant and benign samples (9 pairs) and between paired primary and metastatic cancers (4 pairs) were selected using the *limma* package by fitting both tissue type and pairing information in the linear model (equivalent to paired t-test). Genes correlated with tumor aggressiveness were selected using the

Spearman linear regression model between GP, LN status and gene expression in the R package. Genes with correlation p value less than 0.05 were selected and correlation coefficients were used to find positively or negatively correlated genes with the tumor aggressiveness. These genes were plotted in a stacked graph for visualization.

Chapter 3

Results

3.1 Full-length AR and truncated AR variant mRNA expression in prostate cancer cell lines

The absolute quantification of full-length mRNA was determined in RNA isolated from a panel of androgen-dependent and androgen-independent human PCa cell lines that mimic various stages, conditions of cancer progression and tissue specific metastatic sites (Figure 3.1A) using specific PCR primers in real-time RT-PCR reactions. Threshold amplification cycle (C_T) values were converted to copy number by extrapolating against standard curves generated from serial dilutions of plasmids harboring full-length AR cDNA. Full-length AR transcript was expressed in the androgen-dependent LNCaP cell line. Full-length AR mRNA was also expressed in the androgen-independent cell lines 22Rv1, VCaP, C4-2 and C4-2b. As a control, full-length AR transcript was detected at negligible levels in the AR null cell lines DU-145 and PC-3.

This technique of absolute quantification real-time RT-PCR was then used to quantify the abundance of several well characterized truncated AR variant mRNA isoforms in the same panel of PCa cell lines (Figure 3.1B). 22Rv1 cells expressed high levels of all four AR variants. Although VCaP cells expressed both AR-V1 and AR-V7, this was to a much lesser extent compared to 22Rv1 cells. LNCaP, C4-2, C4-2b, PC-3 and DU-145 cells expressed minimal or negligible amounts of the four AR variants.

3.2 Truncated AR isoforms mediate ligand-independent constitutive AR transcriptional activity in 22Rv1 cells

The AR variant isoforms are distinct in that they lack the LBD and retain the transactivation domains encoded by AR exon 1. Therefore, it was investigated if this is sufficient to sustain AR transcriptional activity in a ligand-independent manner in 22Rv1 cells, which express both full-length AR and multiple AR variants. shRNA constructs targeting AR exon 1 (which knock down both full-length AR and AR variant isoforms) and also AR exon 8 (which knocks down full-length AR selectively) were designed (Figure 3.2A). 22Rv1 cells were transiently transfected with both shRNA constructs, harvested and lysates subjected to Western immunobloting with

an antibody targeted to the AR NTD (Figure 3.2B). shRNA targeting AR exon 1 specifically and efficiently knocks down both full-length AR (110 kDa) and also truncated AR variant isoforms (75 kDa). However, shRNA targeting AR exon 8 selectively knocks down full-length AR only. A mouse mammary tumor virus (MMTV)-regulated promoter reporter was then employed to assess AR activity in 22Rv1 cells. 22Rv1 cells transfected with empty vector (shCon) and treated with a physiological dose (1 nM) of the stable synthetic androgen, R1881 for 24 hr displayed a 10-fold induction in MMTV-Luc activity compared to ethanol-treated controls (Figure 3.2C). This activity was abolished with shRNA targeted to AR exon 1. However, shRNA targeted to AR exon 8 which selectively knocks down full-length AR, abolished MMTV-Luc androgen-responsiveness but had no effect on ligand-independent MMTV-Luc activity. This demonstrates that truncated AR variant isoforms in 22Rv1 cells mediate constitutive, albeit selective, ligand-independent AR activity. Furthermore, differential ablation by exon 1-targeted versus exon-8 targeted shRNA confirms that the truncated AR variant isoforms are derived from distinct mRNA species.

3.3 Androgen withdrawal induces AR variant expression in VCaP cells

The VCaP cell line also express AR variant mRNA, however at a much lower level compared to the 22Rv1 cell line (Figure 3.1B). As AR variants are expressed at highest levels in CRPC, it was hypothesized that AR variants may be stimulated *in vitro* by mimicking castration by withdrawing androgens from the culture medium. Therefore, VCaP cells were cultured in medium supplemented with either 10% FBS or in androgen-depleted conditions in medium supplemented with 10% CSS and harvested at a number of time points. Cell lysates were subjected to Western immunoblot with an antibody that recognizes the AR NTD. Truncated AR variant isoforms (75 kDa) were upregulated following 72 and 96 hr culture in androgen-depleted conditions compared to 24 and 48 hr. As a control, there was negligible expression of AR variant isoforms following culture in medium supplemented with 10% FBS and harvested at either 24 hr or 96 hr. Expression of full-length AR (110

kDa) was unaffected by any of the culture conditions at any of the time points examined (Figure 3.3A).

To confirm these observations, RNA was isolated from VCaP cells cultured in growth medium supplemented with either 10% FBS or 10% CSS for 96 hr. Absolute quantification of full-length AR mRNA was determined using specific primers in quantitative real-time RT-PCR reactions. Androgen withdrawal had no effect on the expression of full-length mRNA in VCaP cells, compared to controls cultured in media supplemented in 10% FBS for 96 hr (Figure 3.3B).

Absolute quantification was then used to quantify the abundance of several AR variants under the same conditions. The transcript expression of all four AR variants examined, AR-V1, AR-V3, AR-V4 and AR-V7 were increased following culture in androgen-depleted conditions for 96 hr compared to controls cultured in media supplemented with 10% FBS for 96 hr (Figure 3.3C).

3.4 Androgen withdrawal does not affect VCaP cell viability

The effect of androgen withdrawal on VCaP cell viability was determined using the MTS assay. Cell viability was assessed at time points from 24-96 hr in VCaP cells cultured in androgen-depleted conditions in media supplemented with 10% CSS and compared to cells cultured in control conditions in media supplemented with 10% FBS. The viability of VCaP cells was unaffected by androgen withdrawal over the time course of the experiment (Figure 3.4).

3.5 Truncated AR variants induced by androgen withdrawal are constitutively transcriptionally active

VCaP cells cultured in androgen depleted conditions induce the expression of truncated AR variants at both the protein and transcript level. Based on the observations in 22Rv1 cells, that truncated AR variants can mediate constitutive ligand-independent AR activity, it was hypothesized, that VCaP cells cultured in androgen-depleted conditions may also be able to support these features. VCaP cells were cultured in 10% FBS for 72 hr followed by transfection with MMTV-Luc with either empty vector (shCon) or shRNAs targeting AR exon 1 (shAR Ex 1) or AR exon

8 (shAR Ex 8) for 48 hr. Cells were then stimulated with 1 nM R1881 for 24 hr. Empty vector control transfected cells responded with a durable increase in luciferase activity following R1881 treatment compared to ethanol control. shRNA targeting both AR exon 1 and AR exon 8 abolished both MMTV-Luc androgen-responsiveness and ligand-independent MMTV-Luc activity (Figure 3.5A). These experimental conditions were repeated in VCaP cells which had been cultured in androgen-depleted conditions in media supplemented with 10% CSS for 72 hrs prior to transfection with MMTV-Luc, shCon, shAR Ex 1 or shAR Ex 8, followed by androgen stimulation with 1 nM R1881 as before. Empty vector control transfected cells responded with a robust increase in luciferase activity following R1881 treatment compared to ethanol control. This activity was abolished with shRNA targeted to AR exon 1. However, shRNA targeted to AR exon 8 which selectively knocks down full-length AR, abolished MMTV-Luc androgen-responsiveness but had no effect on ligand-independent MMTV-Luc activity (Figure 3.5B). These observations were replicated using a different androgen-responsive promoter reporter, 4XARE-E4-Luc under identical experimental conditions. The 4XARE-E4-Luc reporter contains four tandem copies of the ARE of the PSA gene upstream of the adenovirus E4 core promoter and the luciferase gene (Figures 3.6A and 3.6B). This clearly establishes that truncated AR variant isoforms that lack the ligand binding domain are expressed in multiple human prostate cancer cell lines and can support ligand-independent, constitutive AR transcriptional activity using multiple AR transcriptional reporters in both 22Rv1 and VCaP cell lines.

3.6 Individual AR variants display different ligand-independent constitutive transcriptional activities

In order to investigate the transcriptional activities of full-length AR and different AR variants, a number of expression constructs were obtained (see Section 2.3.2). To verify the expression these constructs, each plasmid was transiently transfected in COS-7 cells. Cell lysates were subjected to Western immunoblot with an antibody that recognizes the AR NTD. The expression of each construct

corresponded to the expected molecular weight of each AR variant protein (Appendix II, Figure II.2)

COS-7 cells which do not express full-length AR, were transfected with MMTV-Luc, including expression plasmids encoding full-length AR (AR-FL), AR-V1, AR-V3, AR-V4 or AR-V7 in androgen-depleted conditions, followed by treatment with 1 nM R1881 or ethanol for 24 hr (Figure 3.7). COS-7 cells transfected with AR-FL responded with a durable increase in MMTV-Luc activity following treatment with R1881 compared with ethanol control. Expression of all AR variant expression plasmids abolished this MMTV-Luc androgen-responsiveness. However, differences in constitutive ligand-independent AR activity were observed between the different AR variants. Both AR-V4 and AR-V7 displayed a much greater constitutive activity compared to both AR-V1 and AR-V3 (Figure 3.7).

3.7 AR variant transcriptional activity is negatively regulated by full-length AR *in vitro*

LNCaP cells express robust levels of full-length AR but negligible levels of truncated AR variant isoforms as determined by absolute quantification using real-time RT-PCR (Figures 3.1A and 3.1B). In order to assess the effect of full-length AR on the transcriptional activities of AR variants, LNCaP cells were co-transfected with empty vector (EV) control, vectors expressing AR-V1, AR-V3, AR-V4 or AR-V7 with either an empty vector shRNA construct (shCon) or an shRNA construct targeting AR exon 8 in conjunction with MMTV-Luc for 72 hr in androgen-depleted conditions. Greater constitutive AR activity was observed by AR-V4 and AR-V7 compared to AR-V1 and AR-V3. This is consistent with the ligand-independent AR activity observations in AR null COS-7 cells. However, there was an increase in this constitutive MMTV-Luc activity following knockdown of full-length AR in all AR variants (Figure 3.8). This suggests that AR variant transcriptional activity is negatively regulated by full-length AR *in vitro*.

3.8 Identification of full-length AR and AR variant protein by co-immunoprecipitation and mass spectrometry in 22Rv1 and VCaP cells

Protein expression of full-length AR and truncated AR variants was assessed in 22Rv1 cells and VCaP cells using a combined co-immunoprecipitation and mass spectrometry approach. Whole cell extract of 22Rv1 cells were subjected to co-immunoprecipitation with an antibody targeting the AR NTD which recognizes full-length AR (110 kDa) and also truncated AR variant isoforms (75 kDa). The precipitated proteins were separated by SDS-PAGE and visualized using silver staining (Figure 3.9A) and Western immunoblot with the same antibody used to perform the immunoprecipitation (Figure 3.9B). The Western immunoblot confirms that full-length AR and also truncated AR variant isoforms were immunoprecipitated using this technique. Protein bands corresponding to full-length AR (110 kDa) and truncated AR variants (75 kDa) were excised from the silver stained gel for identification by nanoLC-tandem mass spectrometry. Mass spectrometry identified a total of 12 unique peptides corresponding to 176 amino acids out of a total of 919 amino acids (19% coverage) in the full-length AR protein (Figure 3.10A). This approach also identified 13 unique peptides corresponding to 204 amino acids (22% coverage) from the band excised at 75 kDa (Figure 3.10B). A large number of peptide sequences identified were located with in AR exon 1. Significantly, part of the unique COOH terminus peptide sequence of AR-V7 was also identified in this band (K.C#YEAGMTLGEK.F).

VCaP cells were grown in androgen-depleted conditions for 96 hrs and the whole cell extract was subjected to an identical co-immunoprecipitation procedure. The precipitated proteins were separated and visualized using silver staining (Figure 3.11A) and Western immunoblot with the same antibody used to perform the immunoprecipitation (Figure 3.11B). The Western immunoblot confirms that full-length AR and also truncated AR variant isoforms were immunoprecipitated using this technique. Protein bands corresponding to full-length AR (110 kDa) and truncated AR variants (75 kDa) were excised from the silver stained gel for identification by nanoLC-tandem mass spectrometry. Mass spectrometry identified a total of 20 unique peptides corresponding to 307 amino acids out of a total of 919

amino acids (33% coverage) in the full-length AR protein (Figure 3.12A). This approach also identified 5 unique peptides corresponding to 76 amino acids (8% coverage) from the band excised at 75 kDa (Figure 3.12B). A large number of peptide sequences identified were located with in AR exon 1; however none of the unique COOH terminus peptide sequences of the AR variants were identified.

3.9 Identification of an AR variant regulated gene expression profile in 22Rv1 cells

22Rv1 cells endogenously express full-length AR, as well as a number of truncated AR variant isoform mRNAs which are translated into protein. Therefore, 22Rv1 cells are an ideal model system to delineate the AR variant gene expression profile. As 22Rv1 cells express a number of AR isoforms, shRNA constructs targeting exon 1 and exon 8 were employed as before to differentially knockdown full-length AR or all AR variant isoforms. 22Rv1 cells were transfected with empty vector (shCon) or shRNAs targeting AR Exon 1 (which knocks down both full-length AR and truncated AR variant transcripts) or AR exon 8 (knocking down just full-length AR) for 72 hr in androgen-depleted conditions with medium supplemented with 10% CSS. Total RNA was isolated, amplified and hybridized to the Illumina human whole-genome array platform. Biological triplicates were processed for each treatment group on a single array platform. Raw microarray data was first imported into the R statistical environment and the relative expression of each sample was visualized using a box plot. Minor variations were present between samples (Appendix IV, Figure IV.3A). The Universal Human Reference (UHR) sample was removed and quantile normalization was applied to the 9 samples. Detection p values were checked across all samples and probes with $p \geq 0.05$ were excluded from further analysis, leaving 26, 507 probes in the final analyses. Unsupervised clustering of the filtered data found that the samples clustered by treatment groups (Appendix IV, Figure IV.3B), indicating good quality, comparable samples. Pair-wise comparisons were conducted using a moderated t test for differentially expressed genes between shAR Exon 1 and shCon and also shAR Exon 8 and shCon (Figure 3.13A) using the R *limma* package. This identified 5,256 probes differentially expressed between shAR Exon 1 and shCon and 5,086 probes differentially expressed

between shAR Exon 7/8 and shCon (fold change >1.0; $p < 0.05$; Appendix IV, Figure IV.4). In order to establish a gene expression profile unique to the truncated AR variant isoforms, independent of full-length AR, the 3,714 probes common to both groups were subtracted, leaving 1,542 probes uniquely regulated by truncated AR variant isoforms and 1,372 probes uniquely regulated by full-length AR (Figure 3.13B). For the purposes of further analyses, only probes which were at least 1.3 fold different to the control shCon condition were included ($p < 0.05$). This approach detected 297 probes (156 upregulated and 141 downregulated) which were uniquely regulated by AR variants and 254 probes (68 upregulated and 186 downregulated) uniquely regulated by full-length AR. A full list of these probes can be found in Appendix IV (Tables IV.4 and IV.5).

Figure 3.14 shows a heat map representing differential expression of probe sets (fold change >1.5; $p < 0.05$) between shRNA conditions. Each row on the heatmap represents a probe set; each column represents an individual sample. Gene expression on each probe set was standardized to the mean of samples where 'red' is higher than the mean and 'green' is lower than the mean.

The microarray results were confirmed using real-time RT-PCR in two ways. Firstly, in order to confirm the knock down of full-length AR and the truncated AR variant isoforms, RNA from the same experiment was subjected to quantitative real-time PCR using primers specific for full-length AR and also AR-V1, AR-V3, AR-V4 and AR-V7. This confirmed the specificity and durability of the AR exon 1 targeted-shRNA knockdown of full-length AR (Figure 3.15A) and also the truncated AR variant isoforms (Figure 3.15B) endogenously expressed in 22Rv1 cells. Furthermore, to independently validate the microarray results, 22Rv1 cells were transfected with empty vector (shCon) or shRNAs targeting AR exon 1 for 72 hr in androgen-depleted conditions with medium supplemented with 10% CSS. RNA was isolated and subjected to real-time RT-PCR using primers specific for a panel of the most significantly upregulated and downregulated AR variant regulated genes identified in the microarray gene expression study (see Appendix IV, Table IV.5). shRNA targeted to AR exon 1 upregulated ACPP, LEF1, NR4A2, BNIP3L, PLA2G10, NRP1 and SGK compared to empty vector control (Figure 3.16). MKNK2, PALLD,

SIGMAR1, PMEPA1 were all down regulated by shAR exon 1 compared to empty vector control (Figure 3.16). These results corroborated the findings of the microarray gene expression analysis.

3.10 Full-length and truncated AR variants regulate distinct molecular and cellular functions in 22Rv1 cells

In order to gain a further insight into the large numbers of genes identified from the microarray gene expression analysis in 22Rv1 cells, network generation and functional analysis of both the full-length and AR variant gene signatures were performed using the Ingenuity Pathway Analysis (IPA) software. Functional analysis of the three largest networks generated using the full-length AR gene signature were assigned functions in cell death, cellular assembly and cellular proliferation (Figure 3.17A). This is in contrast to the functional analysis of the three largest networks generated using the AR variant gene signature, which were cell signaling, DNA replication and cellular assembly (Figure 3.17B). Detailed results of the networks and functional analysis generated from the full-length AR gene signature and also the AR variant gene signature can be found in Appendix IV (Tables IV.6 and IV.7).

3.11 AR variant regulated gene signature can distinguish benign and malignant prostate specimens

To gain a further understanding of the clinical significance of the AR variant gene signature identified in 22Rv1 cells, the expression of the 297 probes (284 unique genes) were analyzed against mRNA expression data sets derived from human prostate specimens. Ninety-five laser capture microdissected prostate specimens were included in this profiling study and consisted of normal epithelium (n=17), benign prostatic hyperplasia (BPH) (n=10), prostate cancers of Gleason Pattern (GP) 3 (n=31), GP4 (n=20), GP5 (n=10) and lymph node (LN) metastasis (n=7) from 81 individual treatment-naïve patients. The 31 GP3 specimens consisted of 28 Gleason Score (GS) 3+3 cases and 3 GS 3+4 cases; 20 GP4 specimens consisted of 10 GS 4+4 cases, 4 GS 4+3 cases and 6 GS 4+5 cases and 10 GP5 specimens consisted of 5 GS 5+5 and 5 GS 5+4 cases. Detailed clinical characteristics of all patients can be found in

Appendix III (Table III.2). RNA was analyzed using the Affymetrix U133 Plus 2.0 whole genome array platform. The 297 probe AR variant regulated gene signature identified in 22Rv1 cells using the Illumina platform mapped to 512 probe sets (231 genes) in the Affymetrix platform used in the prostate specimen profiling study.

For the initial analysis, normal epithelium and BPH samples were designated as non-cancer cases and GP3, GP4, GP5 and LN were designated as cancer cases. Unsupervised clustering of these 95 specimens by the 297 probe set AR variant gene signature alone generated three major clusters, which strikingly separated "non-cancer" and "cancer" prostate specimens (Figure 3.18). A t test comparison revealed that this separation is due to differential expression of 230 out of 512 probes ($p < 0.05$).

3.12 Differentially expressed AR variant regulated genes between normal epithelium, localized and metastatic prostate cancer

Pair-wise comparison between normal epithelium specimens and GP3, GP4, GP5 and LN specimens, respectively, showed differential expression of 179, 192, 161 and 176 probes ($p < 0.05$, one-way ANOVA; Figure 3.19). Notably, altered expression of some probes was unique to a particular tissue type whereas changes in expression of others were common to two or more tissue types. As shown in Figure 3.19, a core of 58 probes was found to be consistently deregulated with altered expression in GP3, GP4, GP5 and LN specimens compared with normal prostate epithelium. A complete list of all genes can be found in Appendix IV (Table IV.8).

3.13 AR variant regulated genes correlate with prostate cancer disease progression

The significance of the AR variant regulated gene signature in disease progression was assessed by exploring its correlation with histological grade and metastatic status. Correlation analysis using Spearman rank analysis was conducted for the 68 malignant specimens to identify genes responsible for disease progression. In total, 165 probes were significantly associated with disease progression (86 probes positively correlated and 80 probes negatively correlated; $p < 0.05$). A complete list of these genes can be found in Appendix IV (Table IV.9). Figure 3.20A shows the 11

most significantly positively correlated probe sets in a stacked line graph. Figure 3.20B shows the 11 most significantly negatively correlated probe sets in a stacked line graph.

3.14 AR variant regulated genes correlate with biochemical recurrence

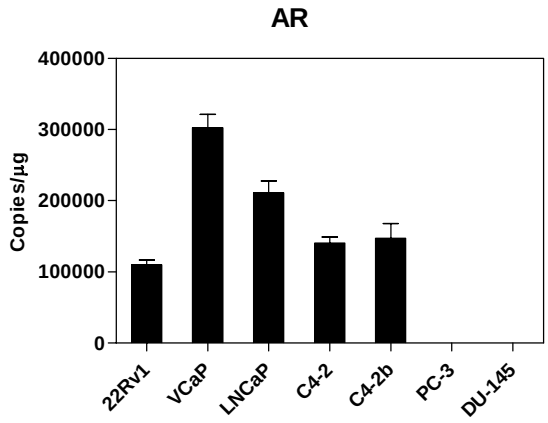
For 56 of 61 patients with localized prostate cancer (GP3, GP4 or GP5) at the time of surgery, follow-up serum prostate-specific antigen (PSA) data was available. All patients were followed up for an average of 42.5 ± 22.42 months and 15 patients had biochemical recurrence within that time (Appendix III, Table III.3). Expression of 56 probes was significantly associated with time to PSA failure ($p < 0.05$; Cox model). A complete list of these genes can be found in Appendix IV (Table IV.10). Figure 3.21 shows six genes significantly associated with time to PSA failure graphed as Kaplan Meier curves. Probe set expression level was extracted and then median expression used to dichotomize patients into low and high so that Kaplan Meier curves could be generated. Log rank test was used to assess for differences between high expression and low expression of the particular probes and time to PSA failure. Low expression of ACOX1 and MAP2K4 was significantly associated with a shorter time to PSA failure ($p < 0.05$). High expression of IGFBP3 and NRP1 was significantly associated with a shorter time to PSA failure ($p < 0.05$). There was no significant difference in the levels of either MBP and TAXPB1 and time to biochemical recurrence.

Figure 3.1 Expression of full-length AR and AR variant mRNA in a panel of prostate cancer cell lines

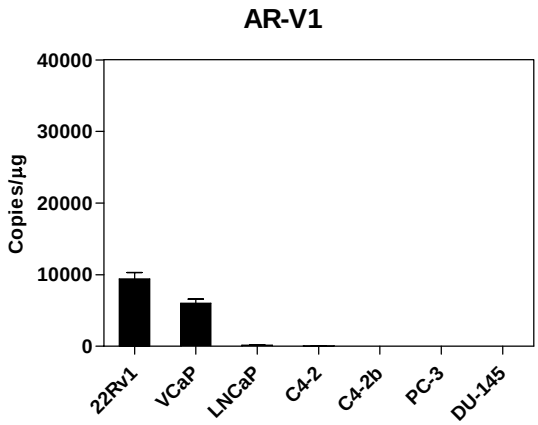
A. Absolute quantification of full-length AR mRNA in a series of androgen-dependent and androgen-independent prostate cancer cell lines was determined using specific PCR primers in quantitative real-time RT-PCR reactions. Threshold amplification cycle (C_T) values were converted to copy number by extrapolating against standard curves generated from serial dilutions of plasmids harboring full-length AR cDNA. Data are expressed as mean (± SEM) copy number/μg total RNA of three independent experiments.

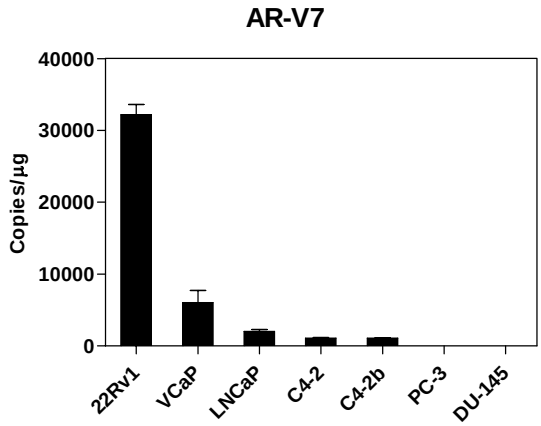
B. Absolute quantification of AR-V1, AR-V7, AR-V3 and AR-V4 mRNA in a series of androgen-dependent and androgen-independent prostate cancer cell lines was determined using specific PCR primers in quantitative real-time RT-PCR reactions. Threshold amplification cycle (C_T) values were converted to copy number by extrapolating against standard curves generated from serial dilutions of plasmids harboring AR-V1, AR-V7, AR-V3 or AR-V4 cDNA. Data are expressed as mean (± SEM) copy number/μg total RNA of three independent experiments.

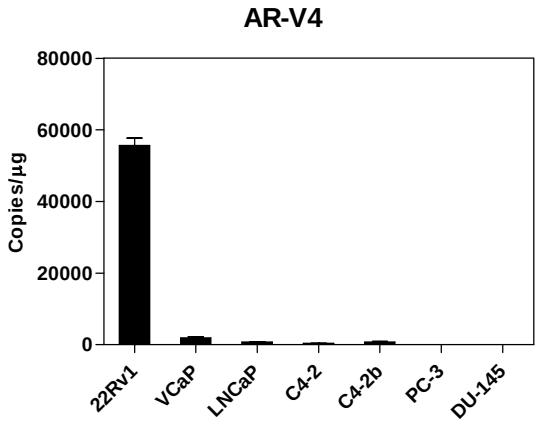
A.


AR
Copies/µg
400000
300000
200000
100000
0
22Rv1
VCaP
LNCaP
C4-2
C4-2b
PC-3
DU-145

B.


AR-V1
Copies/µg
40000
30000
20000
10000
0
22Rv1
VCaP
LNCaP
C4-2
C4-2b
PC-3
DU-145


AR-V7
Copies/µg
40000
30000
20000
10000
0
22Rv1
VCaP
LNCaP
C4-2
C4-2b
PC-3
DU-145


AR-V4
Copies/µg
80000
60000
40000
20000
0
22Rv1
VCaP
LNCaP
C4-2
C4-2b
PC-3
DU-145

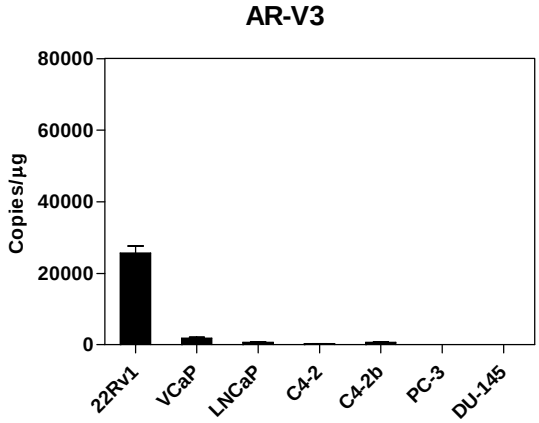

AR-V3
Copies/µg
80000
60000
40000
20000
0
22Rv1
VCaP
LNCaP
C4-2
C4-2b
PC-3
DU-145

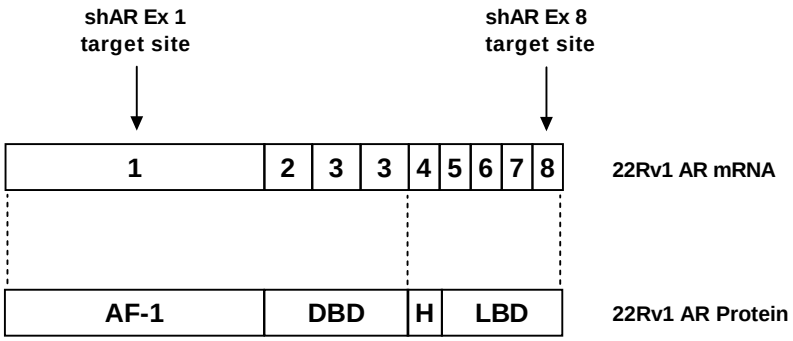
Figure 3.2 22Rv1 cells express alternatively spliced AR variants which mediate ligand-independent constitutive transcriptional activity

A. Schematic representation of full-length AR mRNA and protein in 22Rv1 cells, with target sites for shRNA constructs. 22Rv1 cells have a duplicated exon 3 which codes for a third zinc finger in the DNA-binding domain.

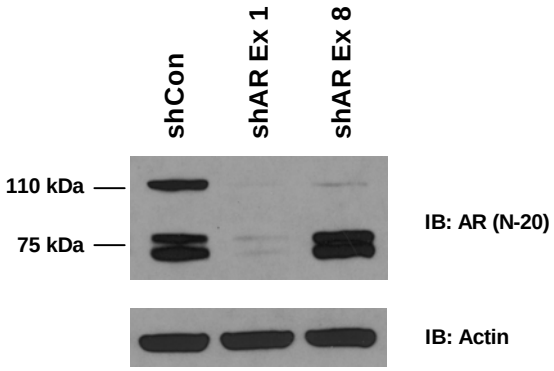
B. 22Rv1 cells were transfected with empty vector (shCon) or shRNAs targeting AR Exon 1 or AR Exon 8 for 72 hr. Lysates were subjected to Western immunobloting using a polyclonal antibody targeted to the AR N-terminal domain. Immunoblots were stripped and reprobed with an actin antibody to confirm equal protein loading.

C. 22Rv1 cells were transfected with MMTV-Luc, empty vector (shCon) or shRNAs targeting AR Exon 1 or AR Exon 8 for 48 hr in medium supplemented with charcoal-stripped serum and treated with 1 nM R1881 or ethanol (vehicle control) for 24 hr. Luciferase activity was determined. Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) from at least three independent experiments performed in triplicate.

A.


shAR Ex 1
target site
shAR Ex 8
target site
1
2
3
3
4
5
6
7
8
22Rv1 AR mRNA
AF-1
DBD
H
LBD
22Rv1 AR Protein

B.


shCon
shAR Ex 1
shAR Ex 8
110 kDa
75 kDa
IB: AR (N-20)
IB: Actin

C.

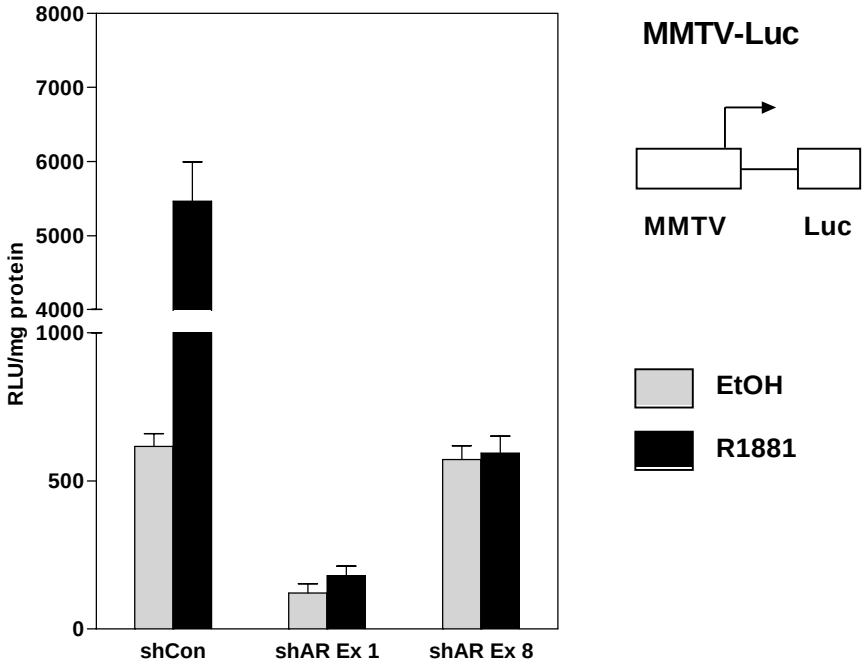

8000
7000
6000
5000
4000
1000
500
0
RLU/mg protein
shCon
shAR Ex 1
shAR Ex 8
MMTV-Luc
MMTV
Luc
EtOH
R1881

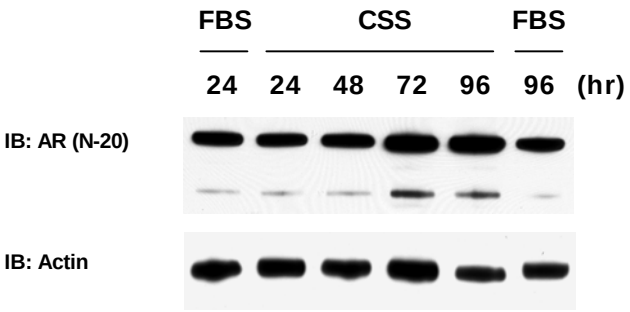
Figure 3.3 Androgen withdrawal induces AR variant expression in VCaP cells

A. VCaP cells were cultured in medium supplemented with either 10% fetal bovine serum (FBS) or 10% charcoal-stripped serum (CSS) and harvested at the time points indicated. Lysates were subjected to Western immunobloting using a polyclonal antibody targeted to the AR N-terminal domain. Immunoblots were stripped and reprobed with an actin antibody to confirm equal protein loading.

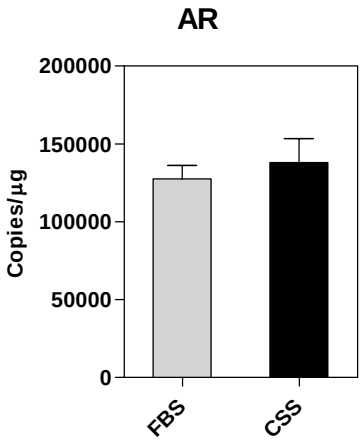
B. RNA was isolated from VCaP cells grown in medium supplemented with either FBS or CSS for 96 hr. Absolute quantification of full-length AR mRNA determined using specific PCR primers in quantitative real-time RT-PCR reactions. Threshold amplification cycle (C_T) values were converted to copy number by extrapolating against standard curves generated from serial dilutions of plasmids harboring full-length AR cDNA. Data are expressed as mean (± SEM) copy number/μg total RNA of three independent experiments.

C. RNA was isolated from VCaP cells grown in medium supplemented with either FBS or CSS for 96 hr. Absolute quantification of AR-V1, AR-V7, AR-V3 and AR-V4 mRNA was determined using specific PCR primers in quantitative real-time RT-PCR reactions. Threshold amplification cycle (C_T) values were converted to copy number by extrapolating against standard curves generated from serial dilutions of plasmids harboring AR-V1, AR-V7, AR-V3 or AR-V4 cDNA. Data are expressed as mean (± SEM) copy number/μg total RNA of three independent experiments.

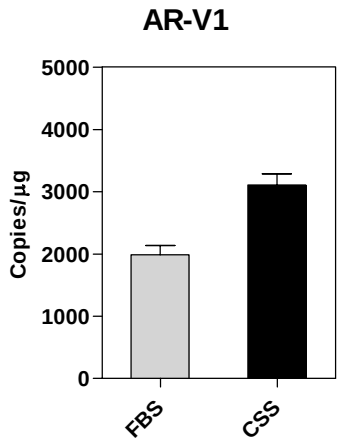
A.

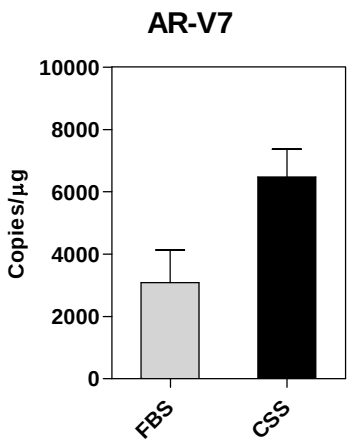

FBS
CSS
FBS
24 24 48 72 96 96 (hr)
IB: AR (N-20)
IB: Actin

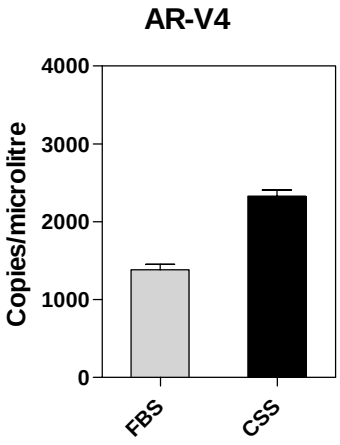
B.


AR
Copies/µg
200000
150000
100000
50000
0
FBS
CSS

C.


AR-V1
Copies/µg
5000
4000
3000
2000
1000
0
FBS
CSS


AR-V7
Copies/µg
10000
8000
6000
4000
2000
0
FBS
CSS


AR-V4
Copies/microlitre
4000
3000
2000
1000
0
FBS
CSS

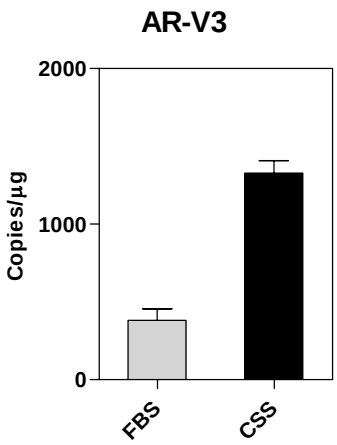

AR-V3
Copies/µg
2000
1000
0
FBS
CSS

Figure 3.4 VCaP cell viability is unaffected by androgen withdrawal

VCaP cells were cultured in medium supplemented with either 10% fetal bovine serum (FBS) or 10% charcoal-stripped serum (CSS). Cell viability was determined by MTS assay at the time points specified. Data are expressed as mean optical density measurements at 490 nm (± SEM) from two independent experiments performed in quadruplicate.

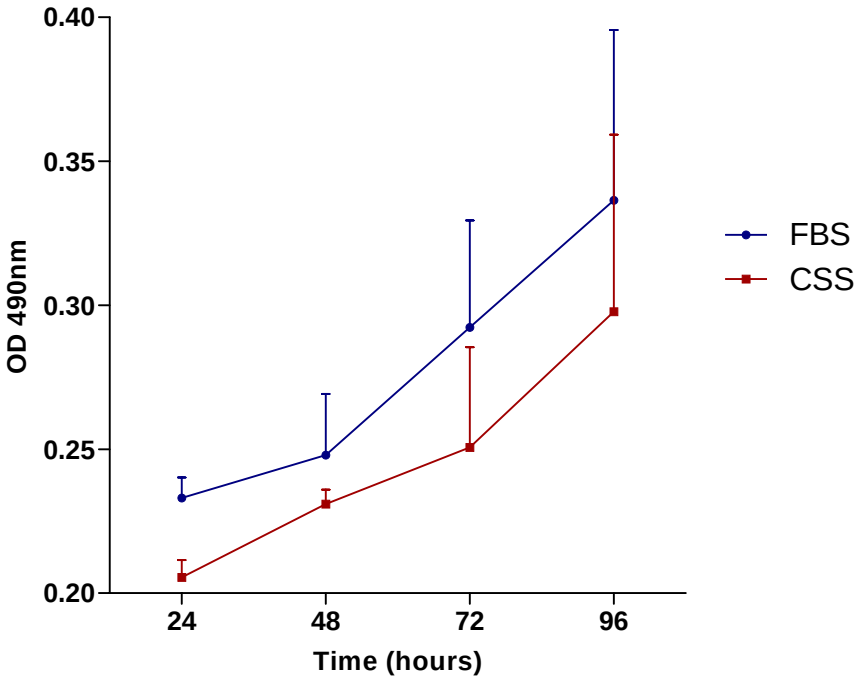

0.40
0.35
0.30
0.25
0.20
OD 490nm
24
48
72
96
Time (hours)
FBS
CSS

Figure 3.5 AR variants induced by androgen withdrawal in VCaP cells are constitutively transcriptionally active on the MMTV-Luciferase promoter

A. VCaP cells were grown in medium supplemented with 10% fetal bovine serum (FBS) for 72 hr and then transfected with MMTV-Luc, empty vector (shCon) or shRNAs targeting AR Exon 1 or AR Exon 8 for 48 hr. Cells were treated with 1 nM R1881 or ethanol (vehicle control) for 24 hr in androgen-depleted medium. Luciferase activity was determined. Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) from at least two independent experiments performed in triplicate.

B. VCaP cells were grown in androgen-depleted conditions with medium supplemented with 10% charcoal-stripped serum (CSS) for 72 hr and then transfected with MMTV-Luc, empty vector (shCon) or shRNAs targeting AR Exon 1 or AR Exon 8 for 48 hr. Cells were treated with 1 nM R1881 or ethanol (vehicle control) for 24 hr in androgen-depleted medium. Luciferase activity was determined (*** $p < 0.001$; ANOVA). Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) from at least two independent experiments performed in triplicate.


MMTV-Luc
MMTV
Luc

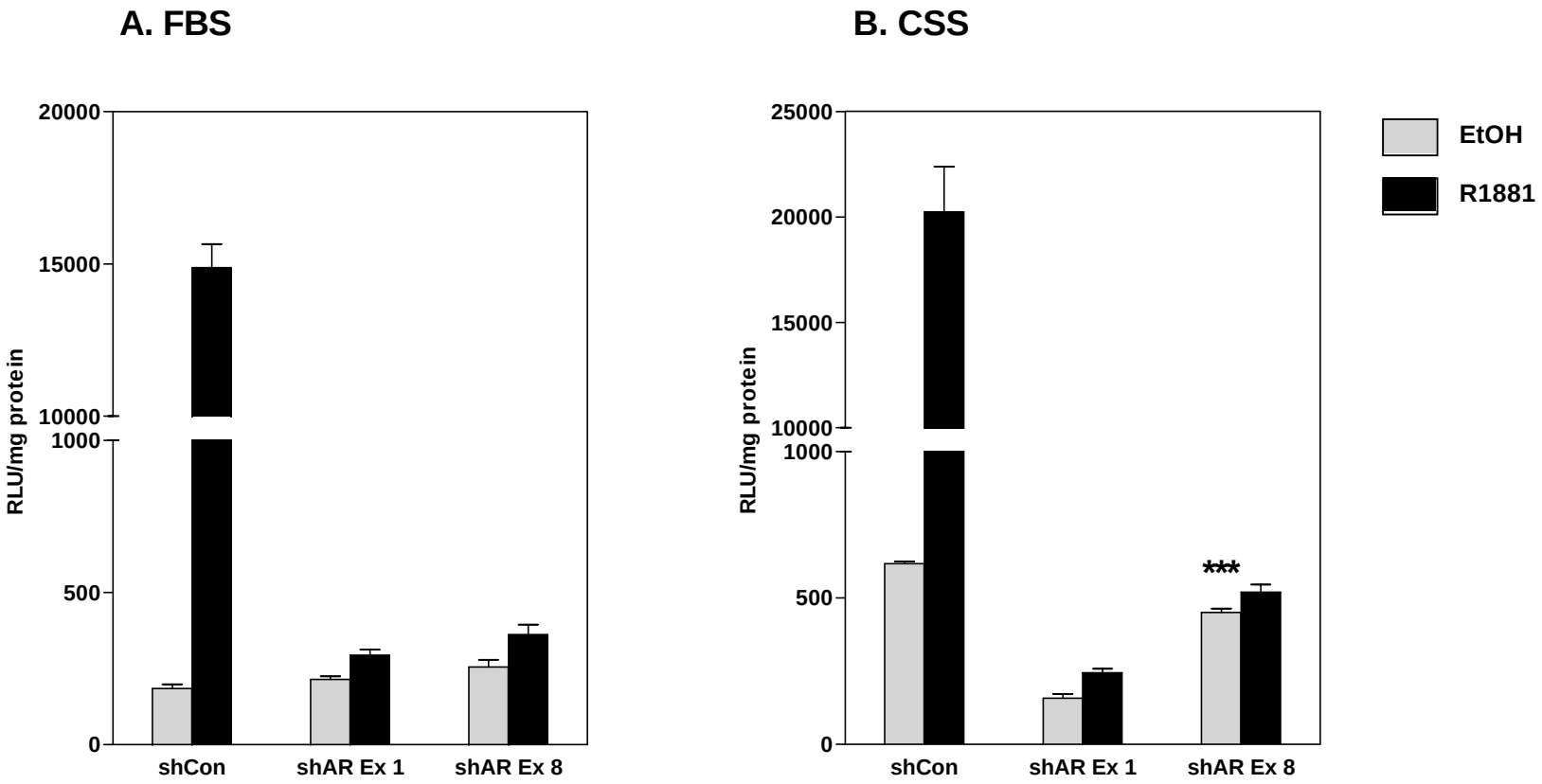
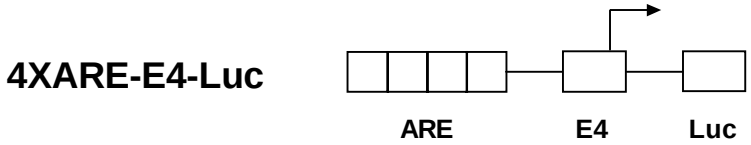
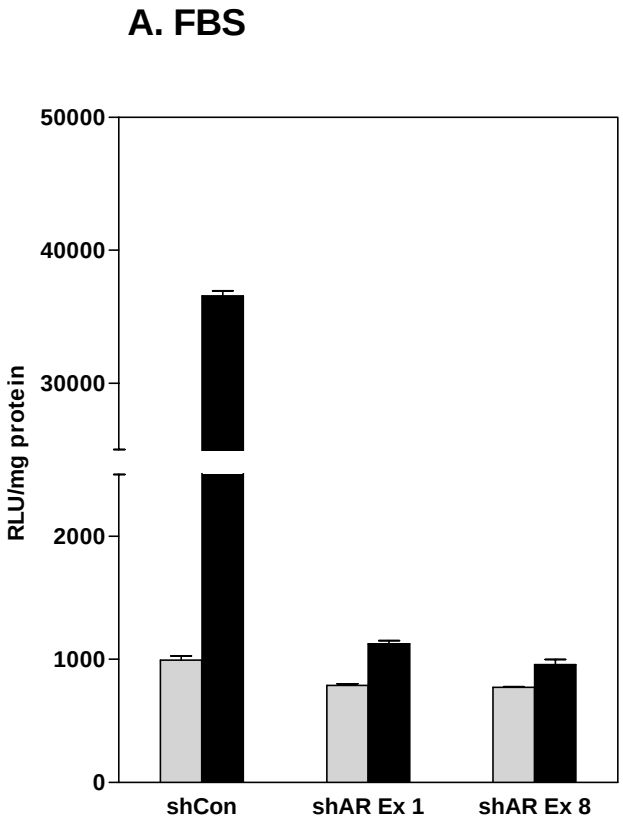

A. FBS
B. CSS
RLU/mg protein
20000
15000
10000
1000
500
0
shCon
shAR Ex 1
shAR Ex 8
25000
***
EtOH
R1881

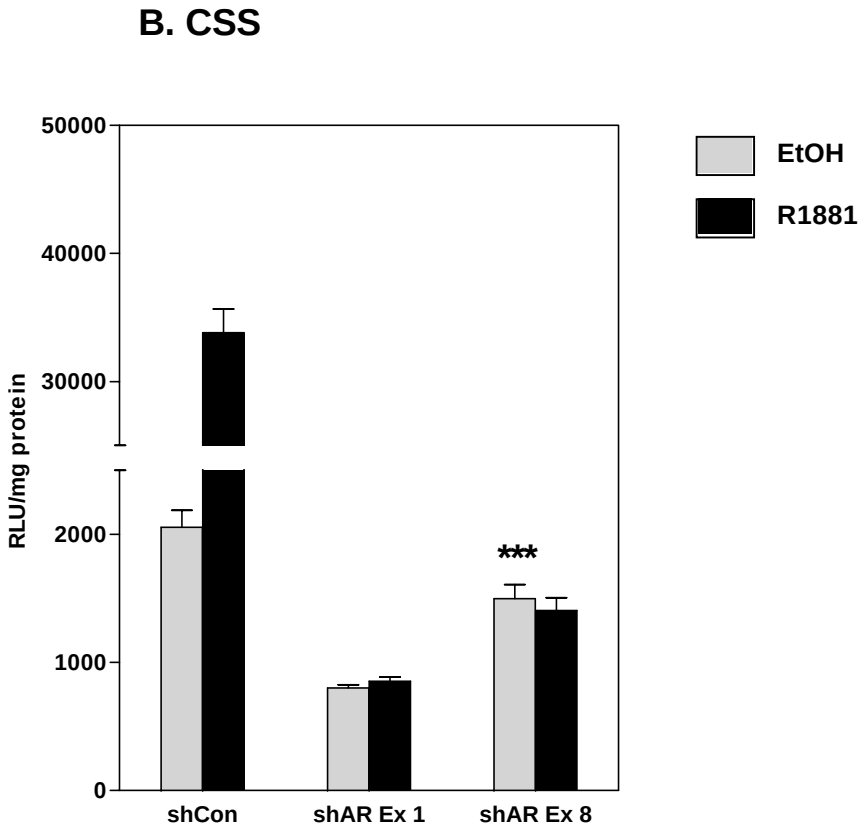
Figure 3.6 AR variants induced by androgen withdrawal in VCaP cells are constitutively transcriptionally active on the 4XARE-E4-Luciferase promoter

A. VCaP cells were grown in medium supplemented with 10% fetal bovine serum (FBS) for 72 hr and then transfected with 4XARE-E4-Luc, empty vector (shCon) or shRNAs targeting AR Exon 1 or AR Exon 8 for 48 hr. Cells were treated with 1 nM R1881 or ethanol (vehicle control) for 24 hr in androgen-depleted medium. Luciferase activity was determined. Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) from at least two independent experiments performed in triplicate.

B. VCaP cells were grown in androgen-depleted conditions with medium supplemented with 10% charcoal-stripped serum (CSS) for 72 hr and then transfected with 4XARE-E4-Luc, empty vector (shCon) or shRNAs targeting AR Exon 1 or AR Exon 8 for 48 hr. Cells were treated with 1 nM R1881 or ethanol (vehicle control) for 24 hr in androgen-depleted medium. Luciferase activity was determined (*** $p < 0.001$; ANOVA). Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) from at least two independent experiments performed in triplicate.


4XARE-E4-Luc
ARE
E4
Luc


A. FBS
50000
40000
30000
2000
1000
0
RLU/mg protein
shCon
shAR Ex 1
shAR Ex 8


B. CSS
50000
40000
30000
2000
1000
0
RLU/mg protein
***
shCon
shAR Ex 1
shAR Ex 8

EtOH
R1881

Figure 3.7 AR variants exhibit different ligand-independent constitutive transcriptional activities

COS-7 cells were transfected with an empty vector (EV) control or expression vectors encoding full-length AR (AR-FL), AR-V1, AR-V3, AR-V4 or AR-V7 in conjunction with MMTV-Luc for 24 hr in androgen-depleted conditions, followed by treatment with with 1 nM R1881 or ethanol (vehicle control) for 24 hr in androgen-depleted medium. Luciferase activity was determined. Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) relative to the empty vector control (with no androgens) from at least two independent experiments performed in triplicate.

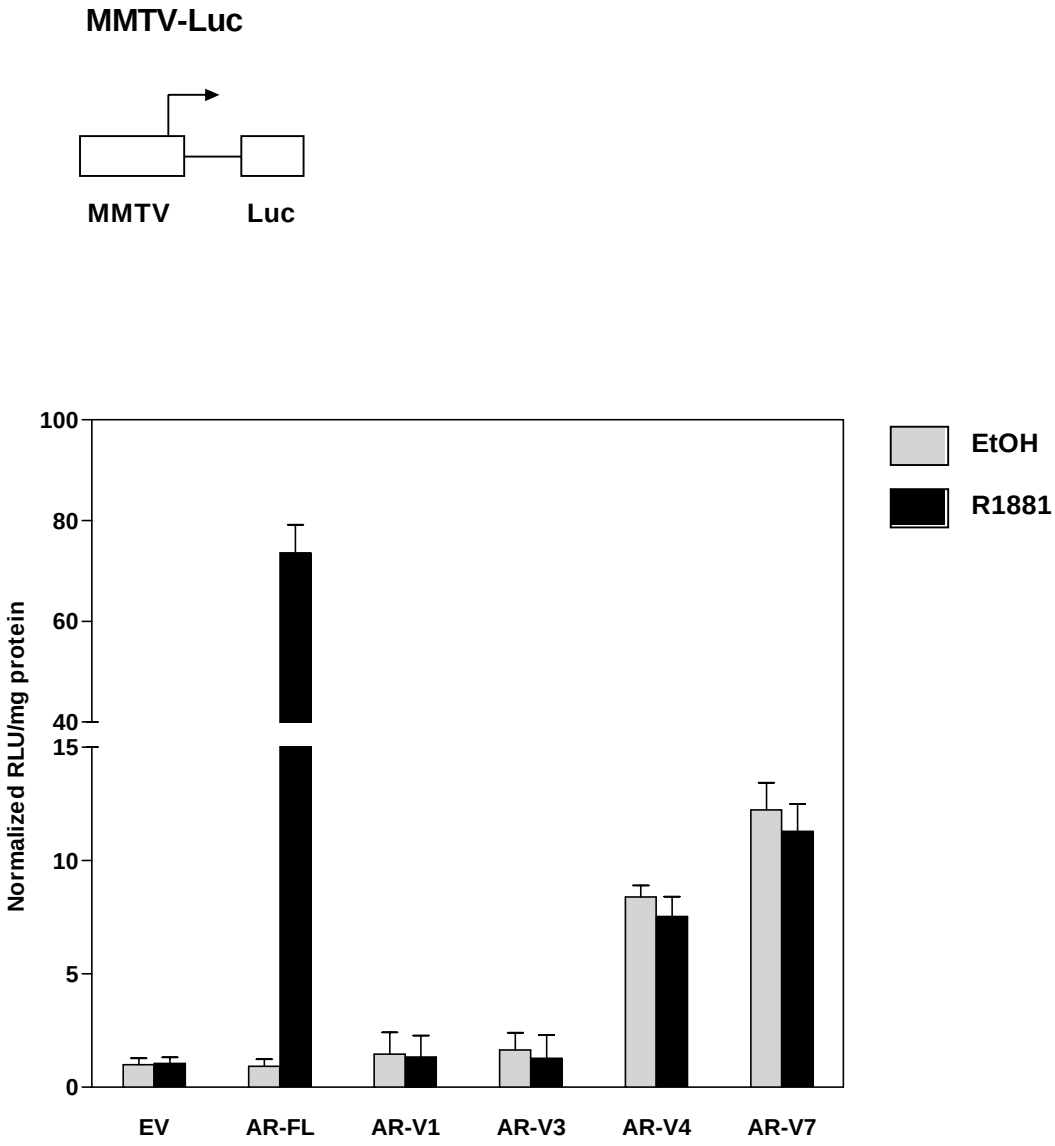

MMTV-Luc
MMTV
Luc
100
80
60
40
15
10
5
0
Normalized RLU/mg protein
EtOH
R1881
EV
AR-FL
AR-V1
AR-V3
AR-V4
AR-V7

Figure 3.8 Transcriptional activity of AR variants is negatively regulated by full-length AR *in vitro*

LNCaP cells were co-transfected with empty vector (EV) control, expression vectors encoding AR-V1, AR-V3, AR-V4 or AR-V7 with either an empty vector shRNA construct (shCon) or an shRNA construct targeting AR Exon 8 in conjunction with MMTV-Luc for 72 hr in androgen-depleted conditions. Data are expressed as mean relative luciferase units (RLU) per mg protein (± SEM) relative to the empty vector control from at least two independent experiments performed in triplicate.

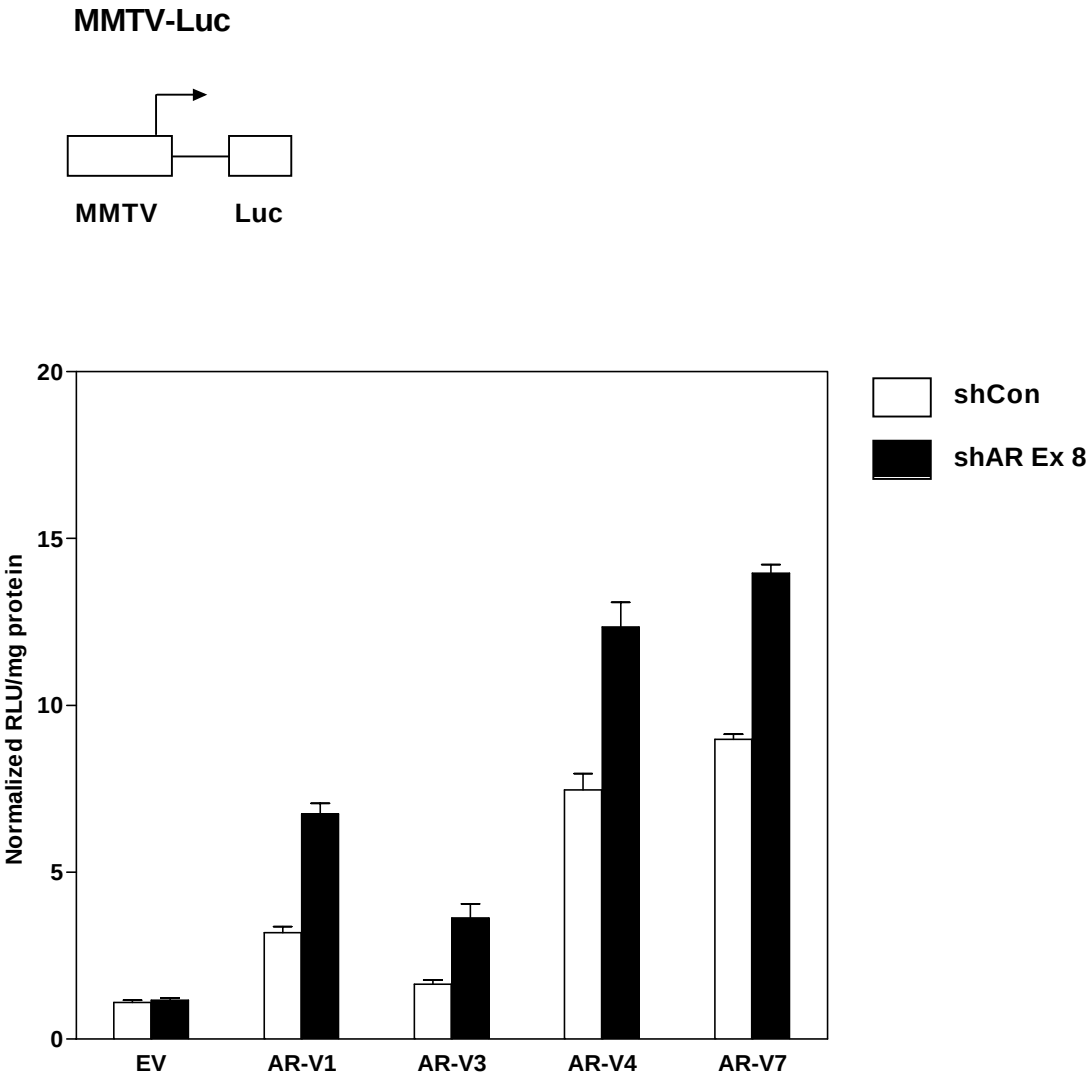

MMTV-Luc
MMTV
Luc
20
15
10
5
0
Normalized RLU/mg protein
EV
AR-V1
AR-V3
AR-V4
AR-V7
shCon
shAR Ex 8

Figure 3.9 Co-immunoprecipitation of full-length AR and truncated AR variants from 22Rv1 cells

A. Co-immunoprecipitation with an antibody which recognizes the AR N-terminal domain was performed with 22Rv1 whole-cell lysate. Eluted proteins were separated on a 4-12% SDS-polyacrylamide gel and visualized by silver staining. Arrows indicate the gel bands corresponding to full-length AR (~110 kDa) and AR variants (~75 kDa) which were excised for nanoLC-tandem mass spectrometry.

B. Eluted co-immunoprecipitated proteins were also subjected to Western immunobloting with the same antibody used to perform the immunoprecipitation. Arrows indicate the expression of full-length AR (~110 kDa) and AR variants (~75 kDa).

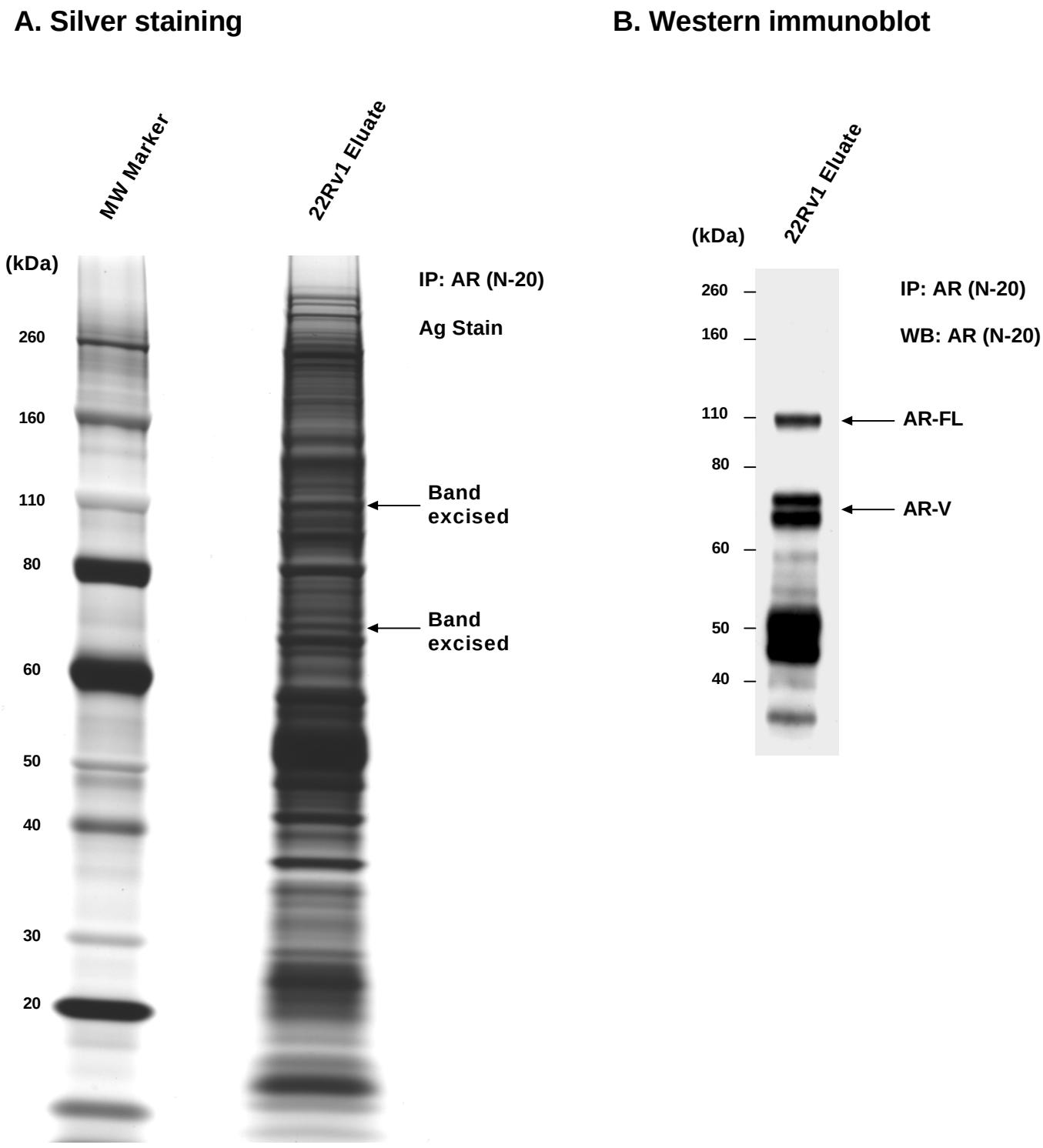

A. Silver staining
B. Western immunoblot
MW Marker
22Rv1 Eluate
(kDa)
260
160
110
80
60
50
40
30
20
IP: AR (N-20)
Ag Stain
Band excised
Band excised
22Rv1 Eluate
(kDa)
260
160
110
80
60
50
40
IP: AR (N-20)
WB: AR (N-20)
AR-FL
AR-V

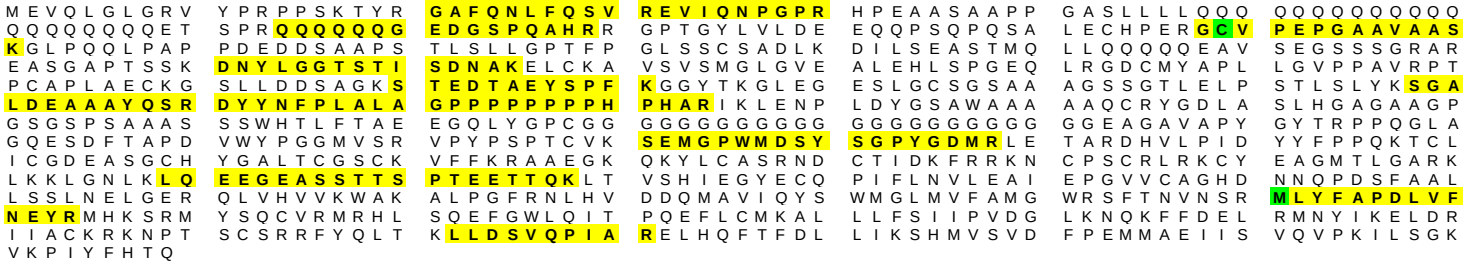
Figure 3.10 Identification of peptide sequences in 22Rv1 cells by mass spectrometry

A. Peptide sequences identified by nanoLC-tandem mass spectrometry (in yellow) and their position within the full-length AR from the gel band excised at ~110 kDa, corresponding to full-length AR.

B. Peptide sequences identified by nanoLC-tandem mass spectrometry (in yellow) and their position within the full-length AR from the gel band excised at ~75 kDa, corresponding to truncated AR variants.

A. 110 kDa band

ANDR_HUMAN (100%), 98,989.6 Da
Androgen receptor OS=Homo sapiens GN=AR PE=1 SV=2 (sp) (P10275)
12 unique peptides, 12 unique spectra, 13 total spectra, 176/919 amino acids (19% coverage)



B. 75 kDa band

ANDR_HUMAN (100%), 98,989.6 Da
Androgen receptor OS=Homo sapiens GN=AR PE=1 SV=2 (sp) (P10275)
13 unique peptides, 13 unique spectra, 13 total spectra, 204/919 amino acids (22% coverage)

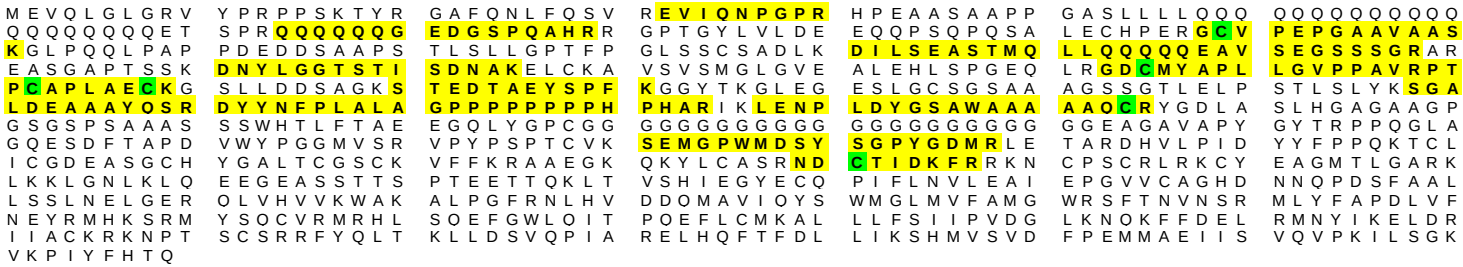


Figure 3.11 Co-immunoprecipitation of full-length AR and truncated AR variants from VCaP cells

A. VCaP cells were grown in androgen-depleted conditions with medium containing 10% CSS for 96 hr, followed by lysis and extraction of whole-cell extract. Co-immunoprecipitation with an antibody which recognizes the AR N-terminal domain was performed with the VCaP cell lysate. Eluted proteins were separated on a 4-12% SDS-polyacrylamide gel and visualized by silver staining. Arrows indicate the gel bands corresponding to full-length AR (~110 kDa) and AR variants (~75 kDa) which were excised for nanoLC-tandem mass spectrometry.

B. Eluted co-immunoprecipitated proteins were also subjected to Western immunobloting with the same antibody used to perform the immunoprecipitation. Arrows indicate the expression of full-length AR (~110 kDa) and AR variants (~75 kDa).

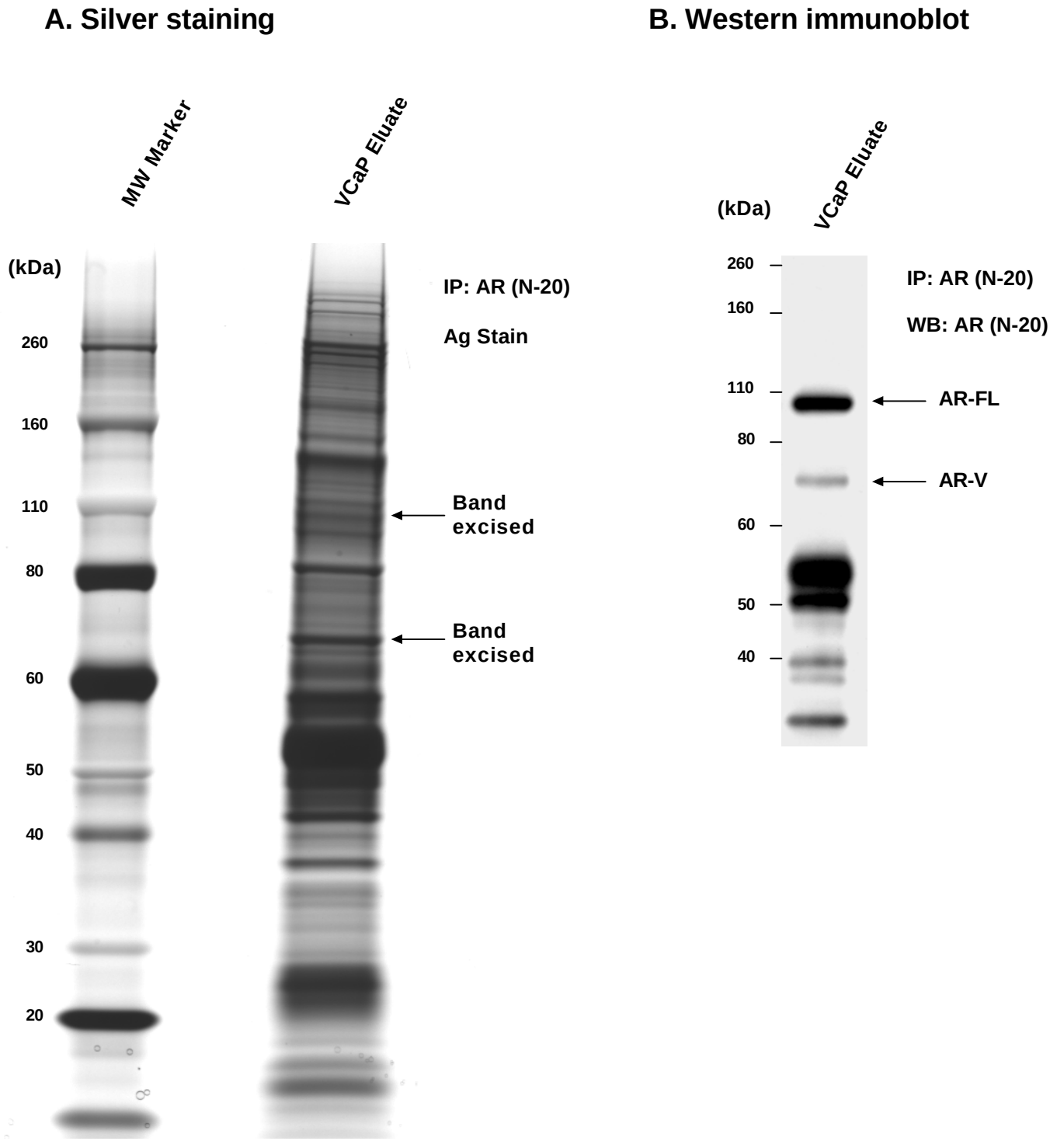

A. Silver staining
B. Western immunoblot
MW Marker
VCaP Eluate
(kDa)
260
160
110
80
60
50
40
30
20
IP: AR (N-20)
Ag Stain
Band excised
Band excised
VCaP Eluate
(kDa)
260
160
110
80
60
50
40
IP: AR (N-20)
WB: AR (N-20)
AR-FL
AR-V

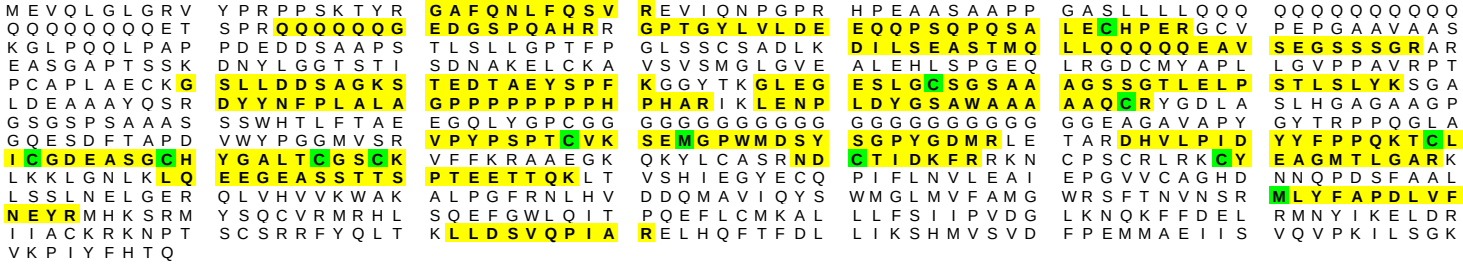
Figure 3.12 Identification of peptide sequences in VCaP cells by mass spectrometry

A. Peptide sequences identified by nanoLC-tandem mass spectrometry (in yellow) and their position within the full-length AR from the gel band excised at ~110 kDa, corresponding to full-length AR.

B. Peptide sequences identified by nanoLC-tandem mass spectrometry (in yellow) and their position within the full-length AR from the gel band excised at ~75 kDa, corresponding to truncated AR variants.

A. 110 kDa band

ANDR_HUMAN (100%), 98,989.6 Da
Androgen receptor OS=Homo sapiens GN=AR PE=1 SV=2 (sp) (P10275)
20 unique peptides, 20 unique spectra, 23 total spectra, 307/919 amino acids (33% coverage)



B. 75 kDa band

ANDR_HUMAN (100%), 98,989.6 Da
Androgen receptor OS=Homo sapiens GN=AR PE=1 SV=2 (sp) (P10275)
5 unique peptides, 5 unique spectra, 5 total spectra, 76/919 amino acids (8% coverage)

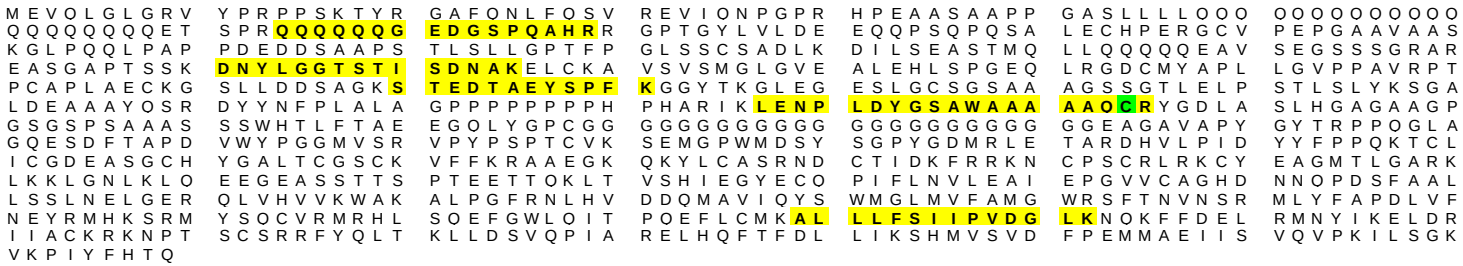
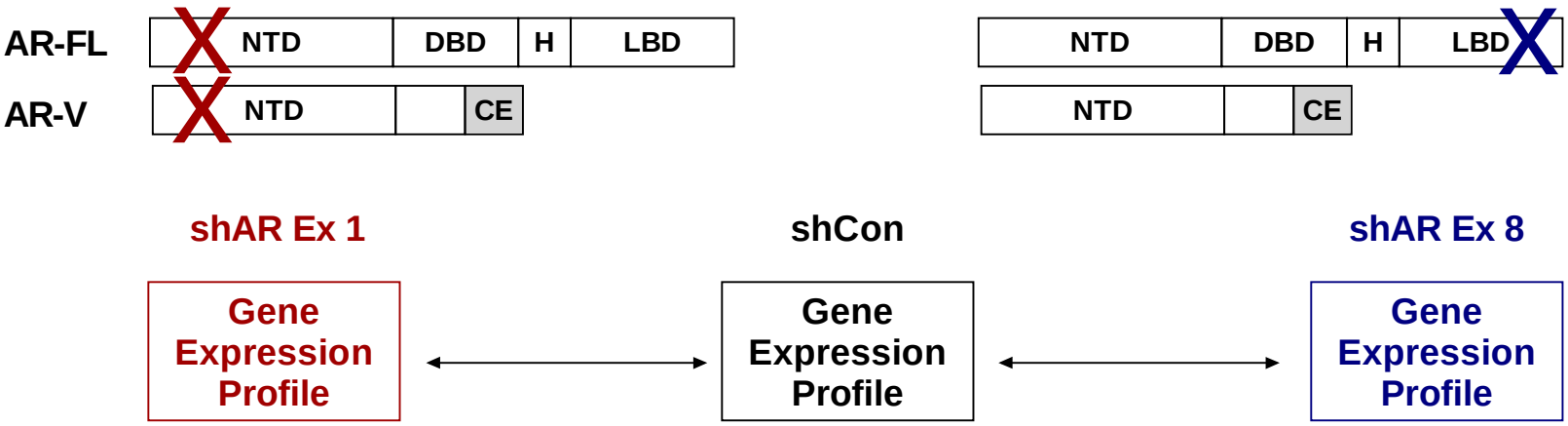


Figure 3.13 Whole genome expression profiling of AR variants and full-length AR by microarray in 22Rv1 cells

A. Experimental design for whole genome microarray study in 22Rv1 cells. 22Rv1 cells were transfected with empty vector (shCon) or shRNAs targeting AR Exon 1 (which knocks down both full-length AR and truncated AR variant transcripts) or AR Exon 8 (knocking down just full-length AR) for 72 hr in androgen-depleted conditions with medium supplemented with 10% CSS. Total RNA was isolated, amplified and hybridized to the Illumina whole-genome array platform.

B. A Venn diagram indicates the number of probe sets differentially expressed (fold change > 1.0; $p < 0.05$) that are in common or unique between shRNA conditions identified by microarray.

A.


AR-FL
AR-V
NTD
DBD
H
LBD
NTD
CE
NTD
DBD
H
LBD
NTD
CE
shAR Ex 1
shCon
shAR Ex 8
Gene Expression Profile
Gene Expression Profile
Gene Expression Profile

B.

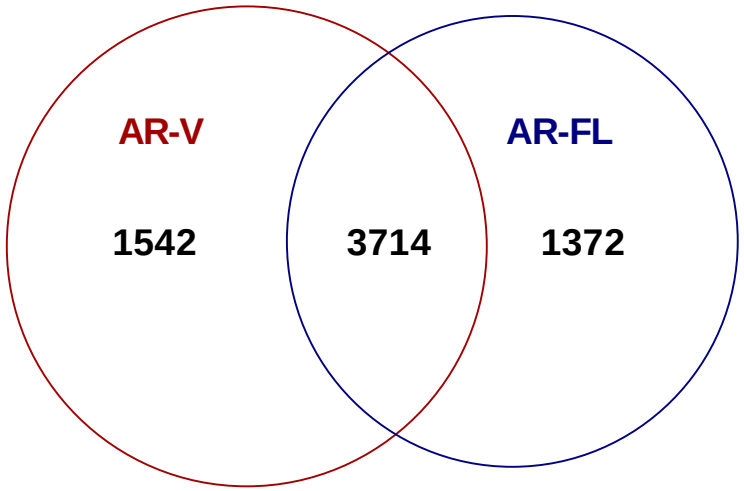

AR-V
AR-FL
1542
3714
1372

Figure 3.14 Identification of differentially expressed probe sets in 22Rv1 cells

Heat map representing differential expression of probe sets (fold change >1.5; $p < 0.05$) between shRNA conditions. Each row on the heatmap represents a probe set; each column represents an individual sample. Gene expression on each probe set was standardized to the mean of samples where red color is higher than the mean and green color is lower than the mean.

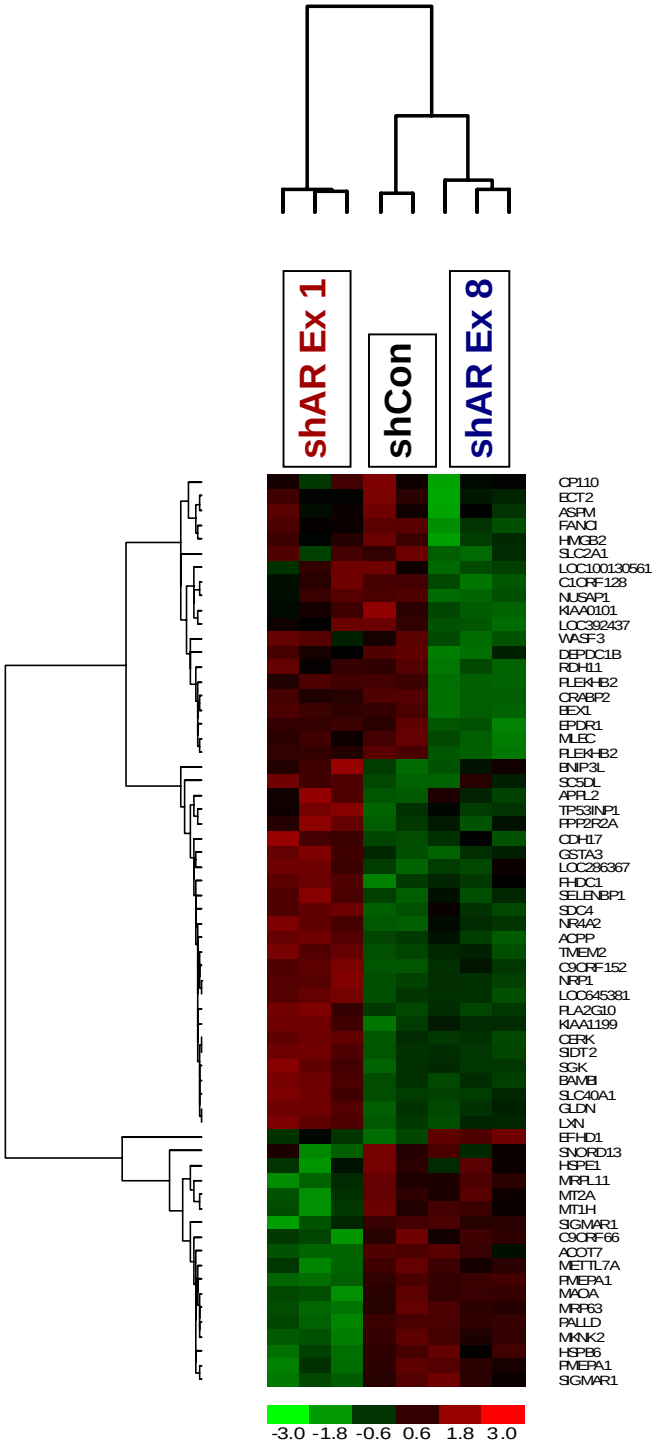
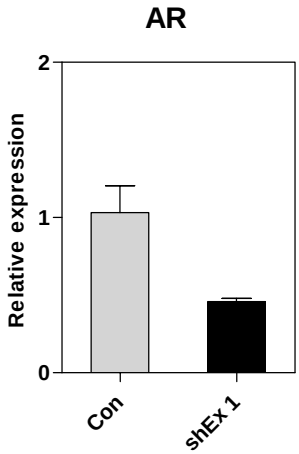

shAR Ex 1
shCon
shAR Ex 8
CP110
ECT2
ASPM
FANCI
HMGB2
SLC2A1
LOC100130561
C1ORF128
NUSAP1
KIAA0101
LOC392437
WASF3
DEPDC1B
RDH11
PLEKHB2
CRABP2
BEX1
EPDR1
MLEC
PLEKHB2
BNIP3L
SC5DL
APRL2
TP53INP1
PPP2R2A
CDH17
GSTA3
LOC286367
FHDC1
SELENBP1
SDC4
NR4A2
ACPP
TMEM2
C9ORF152
NRP1
LOC645381
PLA2G10
KIAA1199
CERK
SIDT2
SGK
BAMBI
SLC40A1
GLDN
LXN
EFHD1
SNORD13
HSPE1
MRPL11
MT2A
MT1H
SIGMAR1
C9ORF66
ACOT7
METTL7A
PMEPA1
MAOA
MRPS3
PALLD
MKNK2
HSPB6
PMEPA1
SIGMAR1
-3.0 -1.8 -0.6 0.6 1.8 3.0

Figure 3.15 Real-time RT-PCR validation of full-length AR and AR variant knockdown in 22Rv1 cells

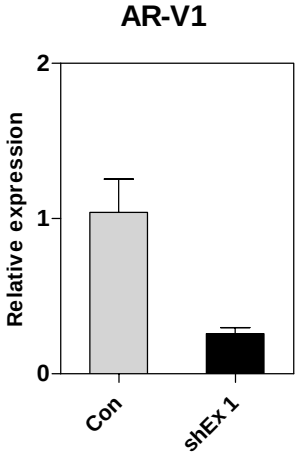
A. 22Rv1 cells were transfected with empty vector control (shCon) or shRNA targeting AR Exon 1 (shAR Ex 1) for 72 hr in androgen-depleted conditions with medium supplemented with 10% CSS. RNA was isolated and subjected to quantitative real-time RT-PCR using primers specific for full-length AR. Expression is relative to empty vector control, which was arbitrarily set to 1. Data are expressed as means (± SEM) from three independent biological replicates.

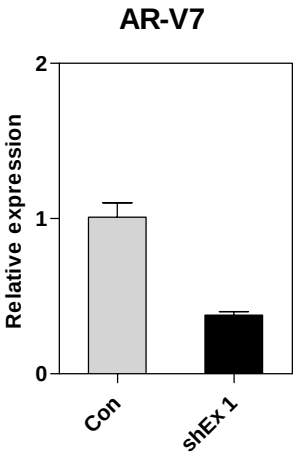
B. 22Rv1 cells were transfected with empty vector control (shCon) or shRNA targeting AR Exon 1 (shAR Ex 1) for 72 hr in androgen-depleted conditions with medium supplemented with 10% CSS. RNA was isolated and subjected to quantitative real-time RT-PCR using primers specific for AR-V1, AR-V3, AR-V4 and AR-V7. Expression is relative to empty vector control, which was arbitrarily set to 1. Data are expressed as means (± SEM) from three independent biological replicates.

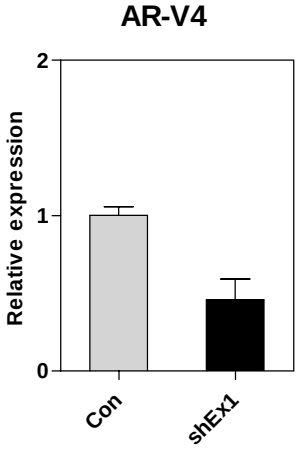
A.


AR
Relative expression
2
1
0
Con
shEx 1

B.


AR-V1
Relative expression
2
1
0
Con
shEx 1


AR-V7
Relative expression
2
1
0
Con
shEx 1


AR-V4
Relative expression
2
1
0
Con
shEx1

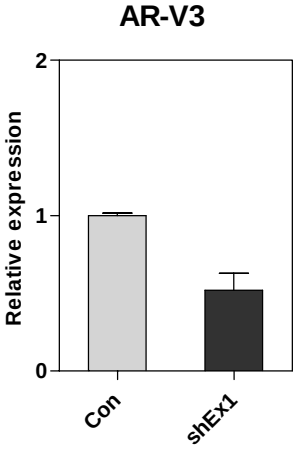

AR-V3
Relative expression
2
1
0
Con
shEx1

Figure 3.16 Real-time RT-PCR validation of microarray results in 22Rv1 cells

22Rv1 cells were transfected with empty vector control (shCon) or shRNA targeting AR Exon 1 (shAR Ex 1) for 72 hr in androgen-depleted conditions with medium supplemented with 10% CSS. RNA was isolated and subjected to quantitative real-time RT-PCR using primers specific for a panel of 11 genes identified from the microarray gene expression experiment in 22Rv1 cells. Expression is relative to empty vector control, which was arbitrarily set to 1. Data are expressed as means (± SEM) from three independent biological replicates.

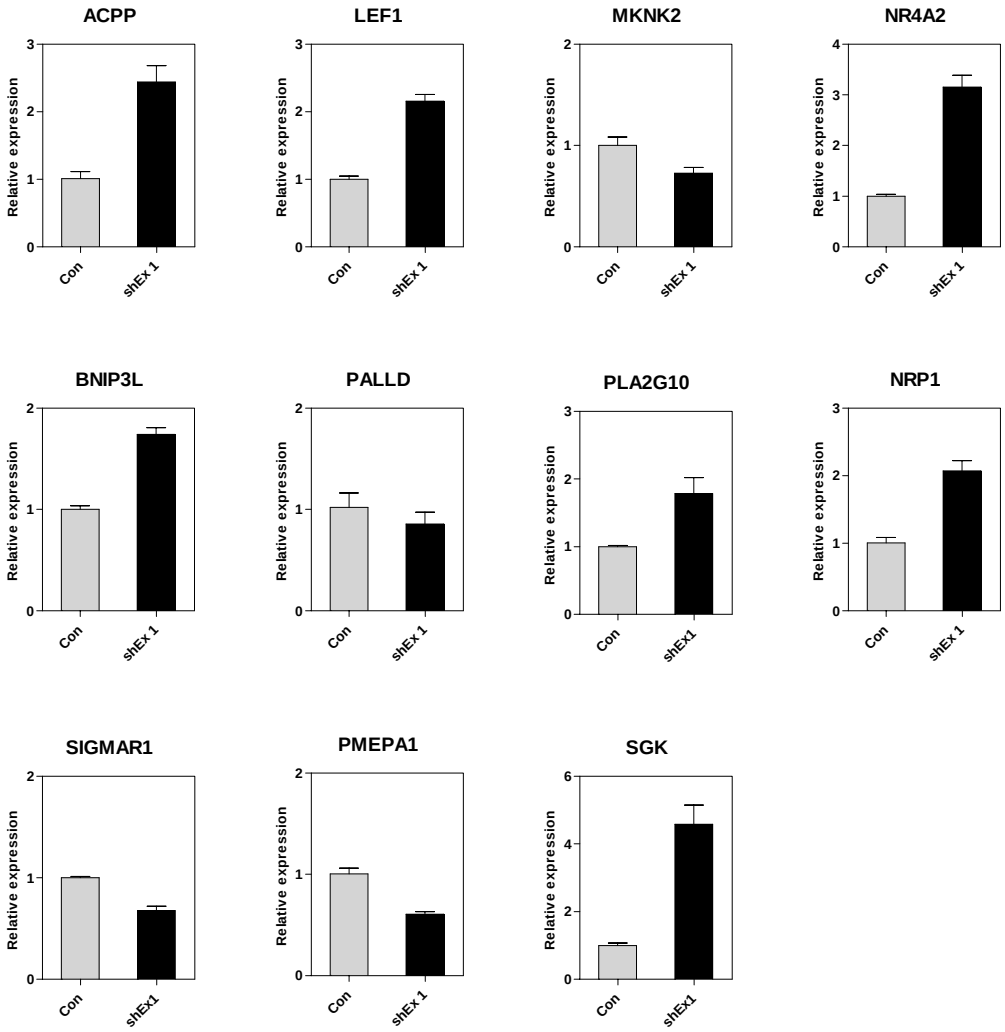
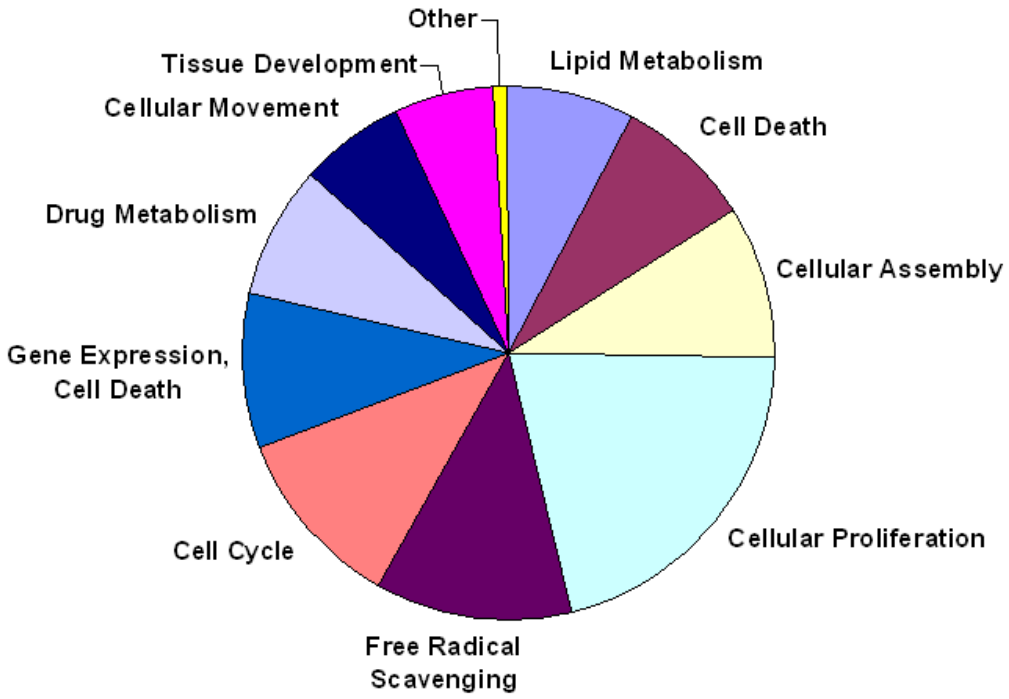

ACPP
LEF1
MKNK2
NR4A2
BNIP3L
PALLD
PLA2G10
NRP1
SIGMAR1
PMEPA1
SGK
Relative expression
Con
shEx 1
shEx1

Figure 3.17 Functional analysis of genes regulated by full-length AR and AR variants in 22Rv1 cells

A. Cellular and molecular functions of the genes uniquely regulated by full-length AR regulated identified by microarray in 22Rv1 cells. Network generation and functional analysis was performed using Ingenuity Pathway Analysis (IPA) using the 254 probes as focus molecules.

B. Cellular and molecular functions of the genes uniquely regulated by truncated AR variants identified by microarray in 22Rv1 cells. Network generation and functional analysis was performed using Ingenuity Pathway Analysis (IPA) using the 297 probes as focus molecules.

A. Full-length AR gene signature



B. AR variant gene signature

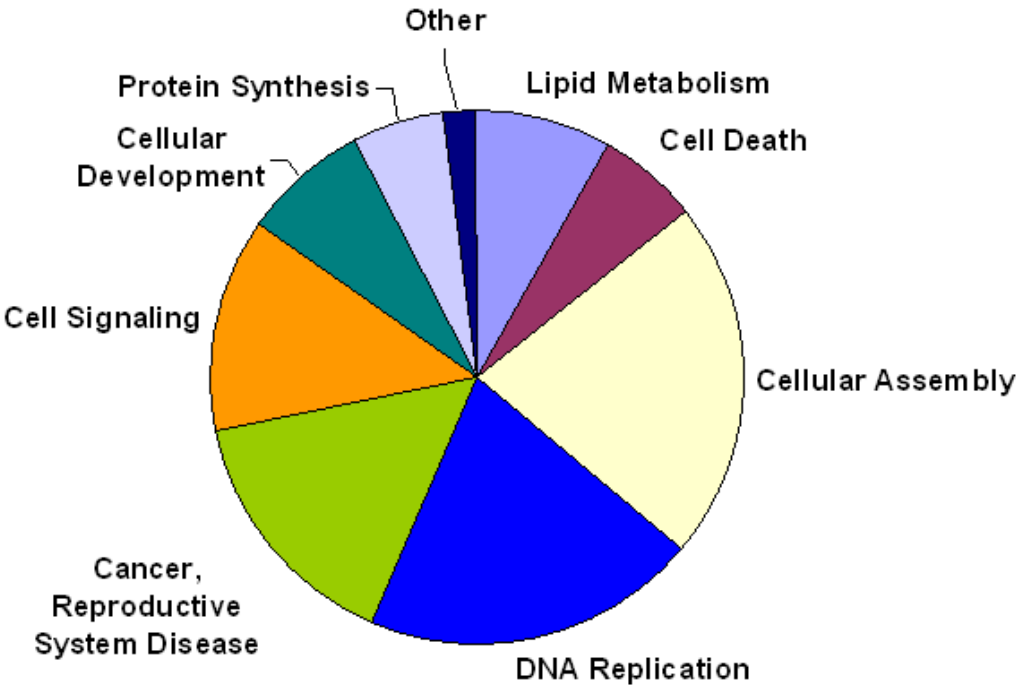


Figure 3.18 Unsupervised clustering of prostate cancer specimens by AR variant regulated gene signature

Unsupervised clustering using the 297 probe set (231 gene) AR variant gene signature identified by microarray in 22Rv1 cells separates benign from malignant prostate cancer specimens. Tissue-type designation N, B, C, L refers to normal epithelium, benign prostatic hyperplasia, localized prostate cancer and lymph node metastasis respectively. Malignant specimens are marked “Y” and benign samples are marked as “N”.

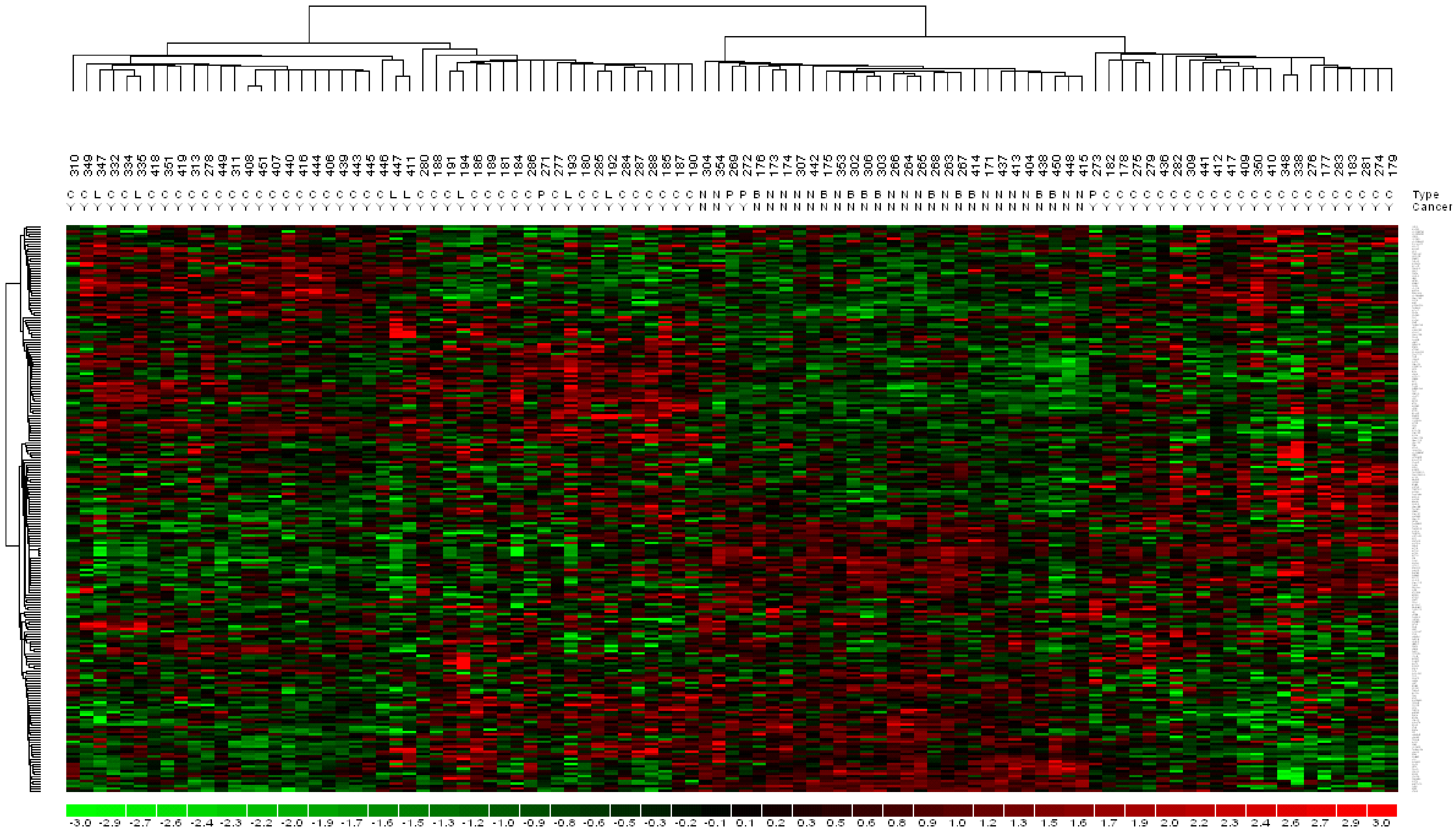

310 349 347 332 334 335 418 351 419 313 278 449 311 408 451 407 440 416 444 406 439 443 445 446 447 411 280 188 191 194 186 189 181 184 286 271 277 193 180 285 192 284 287 288 185 187 190 304 354 269 272 176 173 174 307 442 175 353 302 306 303 266 264 265 268 263 267 414 171 437 413 404 438 450 448 415 273 182 178 275 279 436 282 309 441 412 417 409 350 410 348 338 276 177 283 183 281 274 179
Type
Cancer
-3.0 -2.9 -2.7 -2.6 -2.4 -2.3 -2.2 -2.0 -1.9 -1.7 -1.6 -1.5 -1.3 -1.2 -1.0 -0.9 -0.8 -0.6 -0.5 -0.3 -0.2 -0.1 0.1 0.2 0.3 0.5 0.6 0.8 0.9 1.0 1.2 1.3 1.5 1.6 1.7 1.9 2.0 2.2 2.3 2.4 2.6 2.7 2.9 3.0

Figure 3.19 Differentially expressed AR variant regulated genes in normal epithelium, localized and metastatic prostate cancer

Differentially expressed AR variant regulated probe sets in normal epithelium, localized cancer, and metastatic lymph node (LN) metastases. Pairwise comparisons were made between Gleason Patter (GP) 3, GP4, GP5, and LN metastases and normal epithelium, respectively, using the *limma* package implemented in R to identify differentially expressed probe sets ($p < 0.05$; one-way ANOVA). Expression patterns of probe sets that are common or unique in GP3, GP4, GP5, and LN metastatic specimens are summarized in a 4-way Venn diagram.

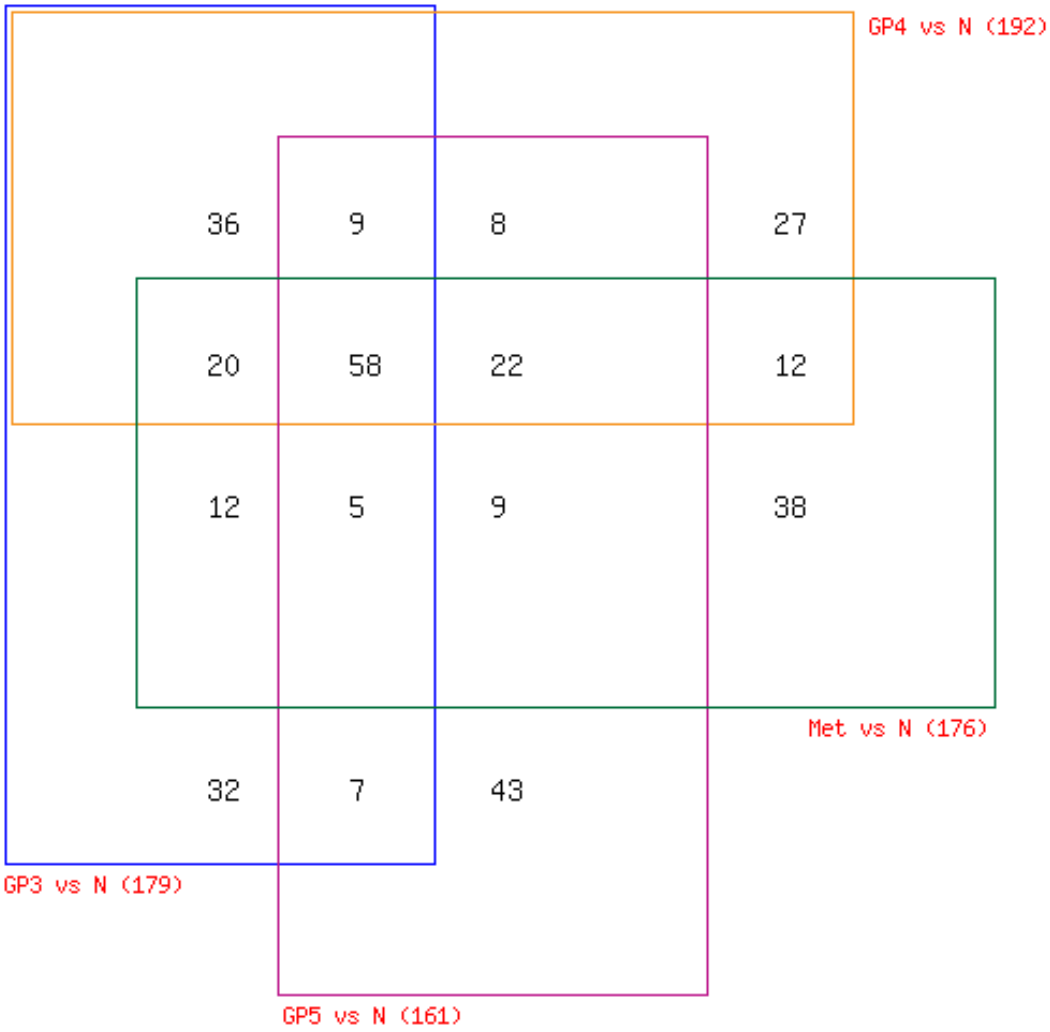
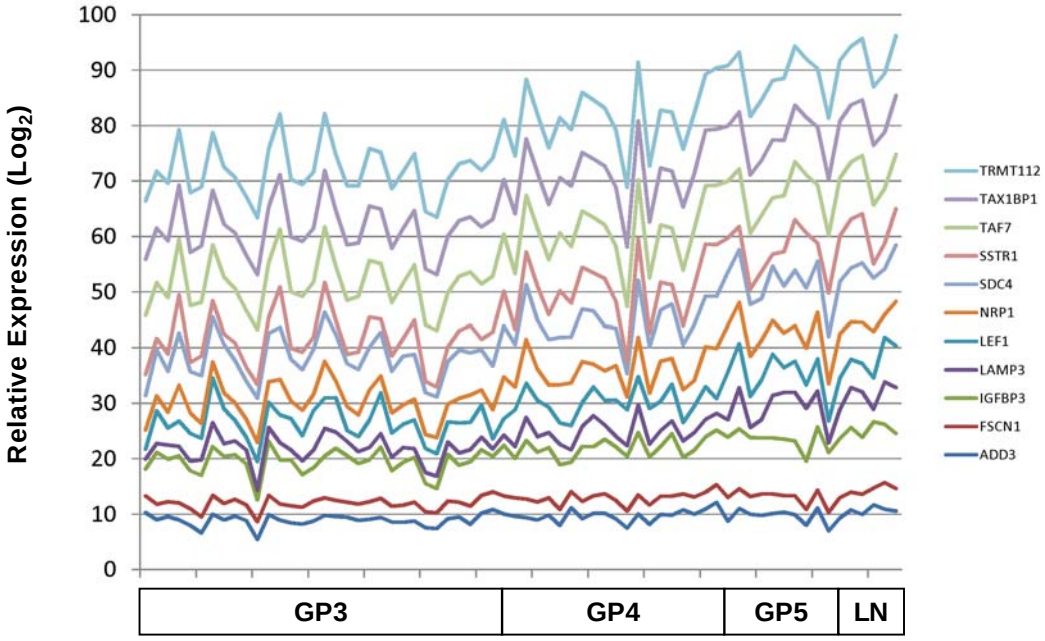

GP4 vs N (192)
36
9
8
27
20
58
22
12
12
5
9
38
Met vs N (176)
32
7
43
GP3 vs N (179)
GP5 vs N (161)

Figure 3.20 Expression of AR variant regulated genes correlates with prostate cancer disease progression

A. A Spearman rank linear regression model assesses the correlation of probe set expression with Gleason Patter and presence of lymph node metastases. The 11 most significantly positively correlated probe sets ($p < 0.05$) are plotted in a stacked line graph. The y-axis represents a Log_2 scale of relative probe set expression.

B. A Spearman rank linear regression model assesses the correlation of probe set expression with Gleason Patter and presence of lymph node metastases. The 11 most significantly negatively correlated probe sets ($p < 0.05$) are plotted in a stacked line graph. The y-axis represents a Log_2 scale of relative probe set expression.

A.


Relative Expression (Log2)
100
90
80
70
60
50
40
30
20
10
0
GP3
GP4
GP5
LN
TRMT112
TAX1BP1
TAF7
SSTR1
SDC4
NRP1
LEF1
LAMP3
IGFBP3
FSCN1
ADD3

B.

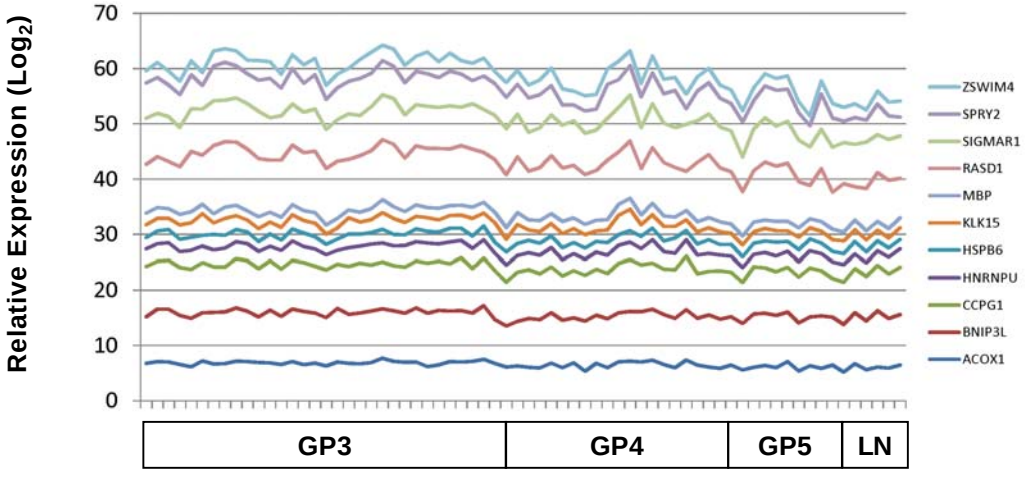
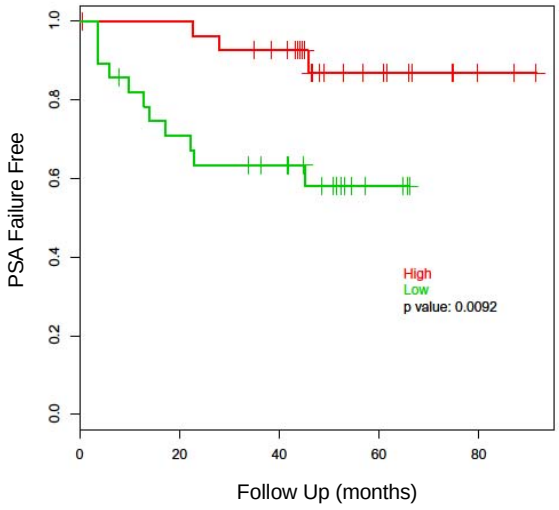

Relative Expression (Log2)
70
60
50
40
30
20
10
0
GP3
GP4
GP5
LN
ZSWIM4
SPRY2
SIGMAR1
RASD1
MBP
KLK15
HSPB6
HNRNPU
CCPG1
BNIP3L
ACOX1

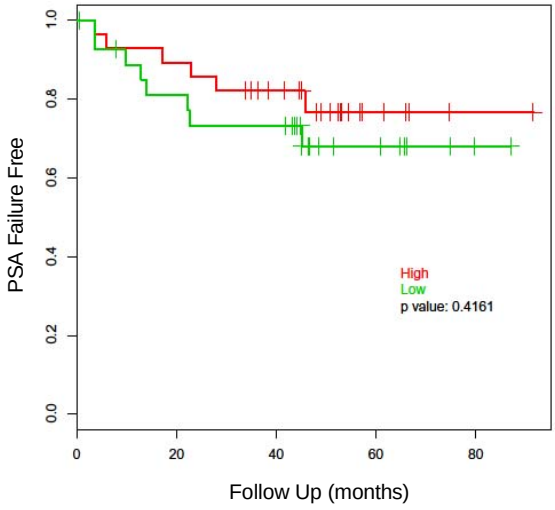
Figure 3.21 Expression of AR variant regulated genes correlates with biochemical recurrence

Follow-up data was available for 56 patients and univariate association was conducted for time to biochemical recurrence ($p < 0.05$). Probe set expression of the six most significantly associated genes were extracted and the median expression was used to dichotomize patients into "high" (red) and "low" (green) so Kaplan Meier curves could be generated. Log rank test was used to determine differences in time to biochemical recurrence.

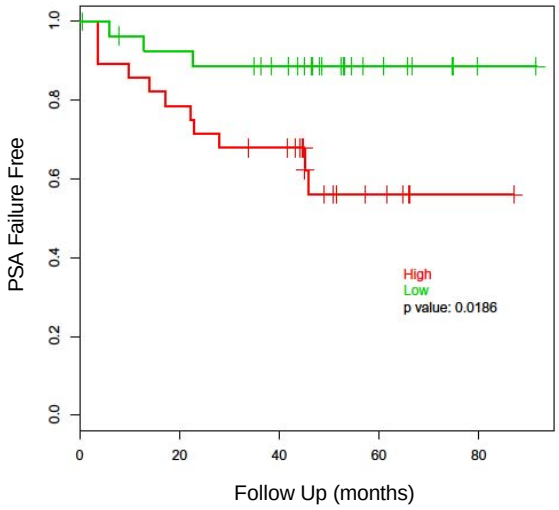
ACOX1



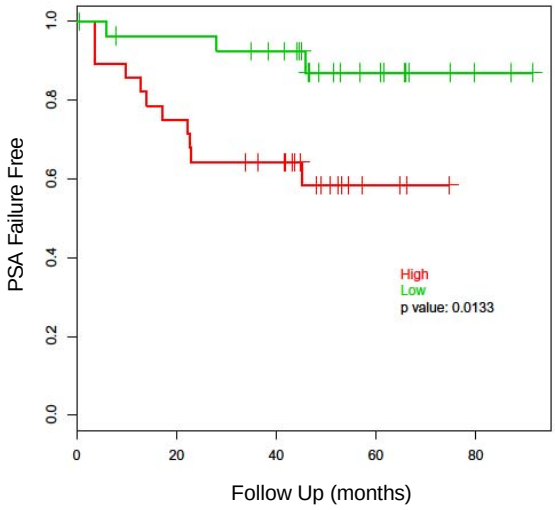
MBP



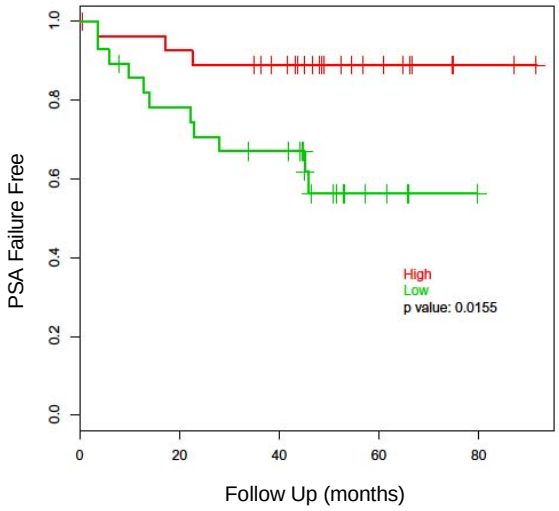
IGFBP3



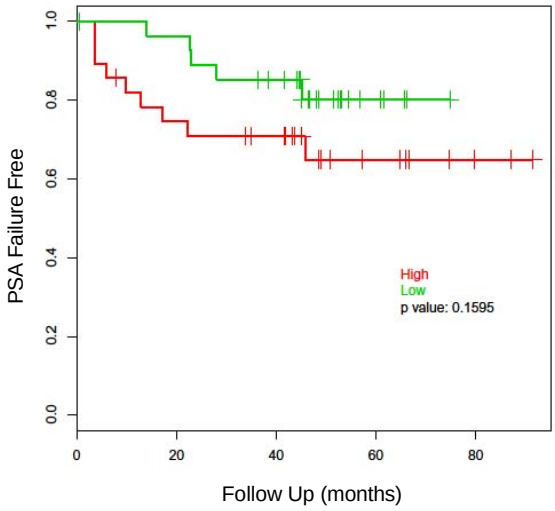
NRP1



MAP2K4



TAX1BP1



Chapter 4

Discussion

The principal objectives of this thesis were (1) to identify an AR variant gene signature in a cell-based model of CRPC and (2) to determine the clinical significance of this AR variant gene signature in a cohort of PCa specimens.

Androgens and AR signaling pathways are regarded as the core oncogenic drivers in prostate carcinogenesis and represent a logical target for PCa treatment. The clinical activity of ADT remains the mainstay of systemic therapy for advanced and metastatic PCa, improving survival in high-risk localized disease and resulting in at least an 80% response rate when initiated in patients with newly diagnosed metastatic disease (Loblaw *et al.* 2007). However, relapse and progression to a castration-resistant phenotype for which no durable treatment currently exists, is inevitable with progression-free survival ranging from 12-20 months (Leuprolide Study Group, 1984; Crawford *et al.*, 1989; Denis *et al.*, 1993; Eisenberger *et al.*, 1998). Unequivocal evidence now exists that restored AR activity is fundamental in the progression to CRPC. Radiological evidence of disease progression is inevitably preceded by a rise in serum PSA, indicating that AR activity has been inappropriately restored despite ADT or combined androgen blockade. Elevated expression of AR protein (Sadi *et al.*, 1991; van der Kwast *et al.*, 1991; Ruizeveld de Winter *et al.*, 1994; Mohler *et al.*, 2004) and AR mRNA (Taplin *et al.*, 1995; Holzbeierlein *et al.*, 2004; Stanbrough *et al.*, 2006) has been consistently found in CRPC tissues. In addition to expressing AR, there is continued expression of multiple AR regulated genes in CRPC, such as PSA, indicating that AR transcriptional activity is reactivated during this stage of the disease (Feldman and Feldman, 2001; Balk, 2002; Mohler *et al.*, 2004; Holzbeierlein *et al.*, 2004; Stanbrough *et al.*, 2006). Multiple mechanisms by which AR is reactivated under androgen-depleted conditions may be involved in the development of this lethal phenotype including AR amplification/overexpression (Buchanan *et al.*, 2001), gain-of-function AR mutations particularly in the LBD (Gottlieb *et al.*, 2012), intratumoral androgen production (Montgomery *et al.*, 2008; Locke *et al.*, 2008) and AR cofactor alteration (Heemers and Tindall, 2007).

In 2008, Dehm and colleagues reported a novel phenomenon whereby alternative splicing of a novel AR exon (from AR intron 2) termed exon 2b, gave rise

to a truncated, constitutively active AR isoform lacking the LBD, that could support androgen-independent gene expression and growth of the CRPC 22Rv1 cell line. Since this initial report, several groups have independently confirmed and extended these findings in various models. To date, over 15 alternatively spliced truncated AR isoforms have been identified and characterized (Dehm *et al.*, 2009; Hu *et al.*, 2009; Guo *et al.*, 2009; Marcias *et al.*, 2010; Sun *et al.*, 2010; Watson *et al.*, 2010; Hu *et al.*, 2011; Yang *et al.*, 2011). These isoforms all have a similar configuration, containing the AR NTD and DBD and unique, short COOH-terminus extensions ranging from 1 to 39 amino acids. AR variant isoforms all lack the LBD, demonstrating that this domain may not be required for AR activity in CRPC cells. Conventional therapy for advanced PCa has focused on androgen-dependent activation of AR through the CTD by either (1) androgen ablation or (2) antiandrogens that compete for the LBD to prevent binding of the endogenous ligand. Therefore, constitutively active, truncated AR splice variants that lack the LBD are a compelling mechanism for CRPC cells to circumvent ADT.

Alternatively spliced AR variants were first identified by 5′ RACE experiments in RNA derived from 22Rv1 cells (Dehm *et al.*, 2008; Hu *et al.*, 2009). Subsequently, AR variants were identified in CWR-R1 cells (Guo *et al.*, 2009; Yang *et al.*, 2011), VCaP cells (Watson *et al.*, 2010), LuCaP xenografts (Sun *et al.*, 2010), CRPC prostatectomy tissues (Hu *et al.*, 2011; Zhang *et al.*, 2011) and CRPC bone metastases (Jagla *et al.*, 2007; Sun *et al.*, 2010; Hornberg *et al.*, 2011). Analyzing the relative AR variant and full-length AR expression by RT-PCR is inherently challenging and may be subject to unintended bias as a result of differences in PCR primer efficiency and product sizes. Approaches such as absolute quantification by real-time PCR using *in vitro* transcribed template as standards may reduce this technical bias (Dehm *et al.* 2008; Li *et al.* 2011). Comparison of AR variant levels with that of full-length AR is further complicated by the possibility of novel variants that have yet to be identified, which may increase the relative proportion of truncated AR variants lacking the LBD. In this study, expression plasmids which contain cDNA for full-length AR, AR-V1 (AR exons 1/2/3/CE1), AR-V7 (AR exons 1/2/3/CE3), AR-V3 (AR exons 1/2/2b) and AR-V4 (AR exons 1/2/3/2b) were used to construct standard curves for absolute

quantification by real-time PCR. Utilizing this method, it was found that both 22Rv1 and VCaP cells express multiple AR variants, as well as full-length AR. 22Rv1 cells express all four AR variant transcripts analyzed with expression of AR-V7 and AR-V4 being the most abundant. However, VCaP cells express lower levels of AR-V1 and AR-V7 compared to 22Rv1 cells and negligible levels of AR-V3 and AR-V4. The androgen-dependent LNCaP cell line, expressed full-length AR transcript but negligible levels of the four AR variants assessed. Interestingly, the androgen independent LNCaP sublines, C4-2 and C4-2b cell lines expressed full-length AR but did not express any of the four AR variant transcripts analyzed. This suggests that factors other than androgen deprivation may be required for AR variant expression. Alternatively, it may suggest that AR variants are cell or tissue site specific. These observations also highlight the heterogeneity of AR variant transcript expression between cell lines. Data from studies examining AR variant expression in clinical specimens also suggest that there is considerable heterogeneity in the AR variant expression patterns between patients (Sun *et al.*, 2010) and also between different metastatic sites within the same patient (Zhang *et al.*, 2011). Moreover, there are also differences between studies in the abundance of specific AR variants. Sun and colleagues found that ARV567es was the most abundant AR variant in 69 metastases samples derived from 13 CRPC patients (Sun *et al.*, 2010), whereas Hornberg and colleagues found that AR-V7 was more abundantly expressed than ARV567es in 30 CRPC bone metastases (Hornberg *et al.*, 2011). This disparity reflects the heterogeneity in AR variant expression, but also may be due to procedural differences and different sensitivities of the real-time PCR techniques used. Furthermore, the intercalating dye reporter technology used in this study may lead to detection of non-specific PCR products, thus potentially leading to lower specificity (Gudnason *et al.*, 2007). Further studies assessing the absolute quantification of AR variants in cell lines and clinical specimens using alternate PCR technologies such as digital PCR are needed to validate the data in this study.

Currently, only AR-V7 has been amenable to antibody production (Hu *et al.*, 2009; Guo *et al.*, 2009), but has yet to be robustly validated in a large cohort of PCa specimens. To overcome this limitation, Zhang and colleagues took advantage of the

fact that novel AR variant COOH termini encoded by alternatively spliced forms of the AR mRNA cannot be recognized by antibodies directed against the wild-type COOH terminus of full-length AR. Using antibodies against the NTD or CTD of the AR protein the authors developed a novel and rapid immunohistochemical approach comparing NTD and CTD AR immunoreactivity to interrogate a large number of benign prostate tissue, primary hormone-naïve PCa and a series of metastatic PCa specimens to ascertain the prevalence of CTD truncated AR variants (Zhang *et al.*, 2011). The most significant observation from the study was the considerable heterogeneity in AR staining between patients and also between different metastatic sites within the same CRPC patient. This suggests that the truncation and/or loss of AR are not necessarily early clonal events in the development of PCa, but rather late stage events occurring independently of one another in CRPC. Therefore, it is tempting to speculate that the generation of truncated AR isoforms, may be associated with mutations, chromosomal rearrangements or splicing events that occur after androgen withdrawal.

In 22Rv1 cells, in addition to the well-characterized 110 kDa full-length AR protein, some lower molecular weight protein bands at 75-80 kDa are immunoreactive with an antibody directed to the NTD, but not the LBD (Tepper *et al.*, 2002). Using this cell line as a model system, custom shRNA constructs were developed targeting sequences in AR exon 1, which would silence both the full-length AR isoform and also shorter AR isoforms which contain the NTD but lack the CTD. Similarly, custom shRNA constructs targeting sequences in AR exon 8 were developed to differentially silence full-length AR which retains the CTD; however truncated AR variant isoforms which lack exon 8 would be unaffected. This approach is unique, as previous studies have knocked down individual AR variants, such as AR-V7 (Guo *et al.*, 2009), whereas this approach has the ability to silence all truncated AR variant isoforms. This is of particular importance in 22Rv1 cells which endogenously express a number of different AR variants as well as full-length AR. The diverse AR variant population of 22Rv1 cells is also more likely to reflect the heterogeneity of AR variant expression patterns found in clinical CRPC specimens (Sun *et al.*, 2010; Hornberg *et al.*, 2011; Zhang *et al.*, 2011). Using these two differential

shRNA constructs in an AR promoter reporter assay also demonstrates that the truncated AR variants expressed in 22Rv1 cells are constitutively transcriptionally active. Furthermore, it confirms previous reports (Dehm *et al.*, 2008) that different mRNA species encode the different AR protein species in 22Rv1 cells.

In addition to 22Rv1 cells, VCaP cells also express truncated AR variants, however at lower levels and with a different AR variant expression pattern. VCaP cells endogenously express AR-V1 and AR-V7 and this is consistent with the original report that used a combination of 3' RACE and next-generation sequencing to identify truncated AR variants in VCaP cells which identified CE1 (found in AR-V1) and CE3 (found in AR-V7; Watson *et al.*, 2010). VCaP cells are a well established cell-based model of CRPC, derived from PCa tissue harvested at autopsy from a metastatic lesion to a lumbar vertebral body of a patient with CRPC (Korenchuk *et al.*, 2001). The VCaP cell line is androgen-sensitive; however, an androgen-independent subline derived from VCaP xenografts which had relapsed following castration of the host has also been reported (Loberg *et al.*, 2006). As a result, the VCaP cell line represents a unique opportunity to investigate the transition to an androgen-independent phenotype. Therefore, it was hypothesized that culturing VCaP cells in steroid-depleted conditions may affect the expression of one or more truncated AR variant isoforms. In this study, protein expression of truncated AR variant isoforms was increased at 72 and 96 hr of culture in charcoal-stripped serum. Protein expression of full-length AR was unchanged by androgen withdrawal at any of the time points examined. Consistent with these data, transcript expression of AR-V1, AR-V7, AR-V3 and AR-V4 were increased in VCaP cells following androgen withdrawal. However, the expression of full-length AR transcript was unchanged. Significantly, this increased expression of truncated AR variants was found to be constitutively transcriptionally active in a ligand-independent manner in two different AR promoter reporter assays. These data expand on the observations in 22Rv1 cells, that constitutively transcriptionally active truncated AR variant isoforms play a role in the progression of PCa to an androgen-independent phenotype that supports CRPC.

Currently, over 15 different truncated AR variants have been identified in a number of cell lines and CRPC specimens using a variety of discovery techniques. As elaborated already, there is considerable heterogeneity present in the expression patterns of AR variants reported in clinical studies, therefore, it is hypothesized that individual variants may have distinct biological activity. To further explore this, the four individual AR variant expression plasmids utilized in this study were transfected into an AR-null cell line and transcriptional activity assessed in an AR promoter reporter assay. This revealed differences in the AR transcriptional activity between individual AR variants with AR-V7 and AR-V4 exhibiting a greater constitutive transcriptional activity compared to AR-V1 and AR-V3. The different transcriptional activities between AR variants may potentially be explained by their cellular localization, with constitutively active AR variants such as AR-V7 localized to the nucleus, while inactive AR variants such as AR-V1 are localized to the cytoplasm (Watson *et al.*, 2010; Hu *et al.*, 2011). The discordance in cellular localization between AR-V1 and AR-V7 is striking given the almost identical exon configuration. Interestingly, both these AR variants lack the NLS in the AR exon 4-encoded hinge region, necessary for nuclear translocation. However, recent a report suggests that AR variants, such as AR-V7 which lack the canonical NLS, contain unique CTD sequences that reconstitute classical NLS activity and this is sufficient for nuclear localization and androgen-independent transcriptional activation of endogenous AR target genes (Chan *et al.*, 2012).

AR variants have consistently been found to be co-expressed with full-length in cell lines, xenografts and clinical specimens (Dehm *et al.*, 2008; Hu *et al.*, 2009; Guo *et al.*, 2009; Sun *et al.*, 2010; Watson *et al.*, 2010; Hu *et al.*, 2011; Zhang *et al.*, 2011; Yang *et al.*, 2011) and also in this study. The relationship between full-length AR and truncated AR variants remains controversial. Currently, only ARV567es has been shown to interact with full-length AR in co-immunoprecipitation experiments (Sun *et al.*, 2010). One report has demonstrated that the activity of AR-V7 is dependent on the presence and function of full-length AR, whereby siRNA-mediated knockdown or inhibition (by the anti-androgen MDV3100) of full-length AR abrogates AR-V7 mediated functions (Watson *et al.*, 2010). However, full-length AR/AR-V7 complexes

have not been detected in 22Rv1 cells using standard co-immunoprecipitation techniques (Tepper *et al.*, 2002; Guo *et al.*, 2009). In order to assess the relationship between full-length AR and the four AR variants used in the study, AR-V1, AR-V7, AR-V3 and AR-V4 were ectopically expressed in LNCaP cells which endogenously express wild-type full-length AR. AR transcriptional activity was assessed following differential knockdown of full-length AR using the shRNA construct targeted to AR exon 8. The transcriptional activity of all AR variants assessed increased following knockdown of full-length AR, therefore this suggests that the transcriptional activity of AR variants is independent of full-length AR. Similar observations have been reported in AR-V9 (Hu *et al.*, 2011), but is in contrast with the observations reported by Watson and colleagues (Watson *et al.*, 2010). However, the data in this study are not directly comparable as different cell lines were used; therefore future studies are required to resolve this important issue.

Although transcript expression of many of the AR variant isoforms have been demonstrated in cell lines, xenografts and clinical specimens, only two studies have demonstrated the translation of truncated AR variant mRNAs into an endogenous protein using an antibody targeted to the unique COOH terminus amino acid sequence of AR-V7 (Hu *et al.*, 2009; Guo *et al.*, 2009). In this present study, a combination of co-immunoprecipitation and mass spectrometry was employed to identify the unique COOH terminus amino acid peptide sequences of known AR variants in 22Rv1 and VCaP cells. Firstly, co-immunoprecipitation with an antibody that recognizes a peptide sequence in the AR NTD was used followed by identification of specific peptide sequences in the protein complex using nanoLC-tandem mass spectrometry. Using this approach, the unique AR-V7 COOH terminus amino acid sequence was identified in 22Rv1 cells, however none of the unique AR variant amino acid sequences were identified in VCaP cells which had been cultured in androgen-depleted conditions. However, a number of peptide sequences corresponding to sequences in AR exons 1-3 were identified in the immunoprecipitated protein complexes in both 22Rv1 and VCaP cells. This immunoprecipitation/mass spectrometry approach requires further refinement but is an exciting possibility for demonstrating translation of AR variant mRNAs into

endogenous proteins. This approach also has the potential to identify cofactor recruitment to the AR variant transcriptional complex, which may provide a further insight the mechanism of AR variant action in CRPC cells (Malovannaya *et al.*, 2010).

The elaboration of the gene expression programs regulated by truncated AR variants has been addressed in several recent reports (Hu *et al.*, 2009, Guo *et al.*, 2009; Sun *et al.*, 2010; Hornberg *et al.*, 2011; Hu *et al.*, 2012). However, conflicting data exists. Microarray analysis performed on RNA isolated from LNCaP cells ectopically expressing AR-V7 cDNA found that AR-V7 induced canonical androgen-responsive genes such as PSA, KLK2, NKX3.1, FKBP5 and TMPRSS2 in the absence of androgens (Hu *et al.*, 2009). A more comprehensive gene expression analysis by Sun and colleagues reported that PSA, TMPRSS2 and FKBP5 were induced by ARV567es expression in LNCaP cells in the absence of androgens. Significantly, a subset of genes regulated by ARV567es that are unique and not influenced by androgens in context of full-length AR were also identified in this study (Sun *et al.*, 2010). In contrast to these studies, a gene expression study performed in CRPC bone metastases expressing ARV567es and/or AR-V7 identified several gene transcripts known to be positively regulated by AR such as C-MYC, UBE2C, CDK1, CYCLINA2 and HSP27 but specifically not some classical androgen-regulated genes such as PSA, TMPRSS2 and FKBP5 (Hornberg *et al.*, 2011). In this study, the 22Rv1 cell model was used for the gene profiling experiments as this cell line endogenously expresses wild-type full-length AR in addition to several truncated AR variants. The endogenous co-expression of both full-length AR and truncated AR variants in this cell line is more reflective of clinical PCa and represents an ideal cell-based model for gene profiling of AR variant target genes. This approach is in contrast to previously reported AR variant gene expression studies were specific AR variants are either overexpressed in PCa model systems (Hu *et al.*, 2009; Sun *et al.*, 2010; Hu *et al.*, 2012) or a specific AR variant is knocked down in a cell line endogenously expressing multiple truncated AR variants (Guo *et al.*, 2009). Therefore, the approach employed in this present study is a more comprehensive analysis of AR variant target genes.

As 22Rv1 cells express a number of AR isoforms, shRNA constructs targeting exon 1 and exon 8 were employed as before to differentially knockdown full-length

AR and all AR isoforms. Whole genome microarray expression analysis revealed 5,256 probes differentially expressed when all AR isoforms are knocked down using shRNA targeted to exon 1 and 5,086 probes differentially expressed when full-length AR is differentially knocked down using shRNA targeted to exon 8. In order to establish a gene expression profile unique to the truncated AR variant isoforms, independent of full-length AR, probes common to both groups were subtracted, leaving 1,542 probes uniquely regulated by truncated AR variant isoforms and 1,372 probes uniquely regulated by full-length AR. The canonical androgen-regulated genes TMPRSS2, hK2, FKBP5 and AR were found to be significantly preferentially regulated in 22Rv1 cells by full-length AR with the expression of these genes downregulated by shAR Ex 1. These data are consistent with a previously reported gene expression profiling study in 22Rv1 cells (Guo *et al.*, 2009). However, in that particular study KLK3 (PSA) was also found to be significantly regulated by full-length AR, whereas PSA was not found to be regulated by either full-length AR or truncated AR variants in this present study. A possible explanation for this discrepancy is that cells were maintained in normal growth conditions whereas cells were cultured in androgen-depleted conditions in this study.

Using microarray, Guo and colleagues identified the differential expression of 188 genes in 22Rv1 cells when AR-V7 was knocked down, where as the expression of 412 genes was altered when full-length AR was specifically inhibited (Guo *et al.*, 2009). Of the genes regulated by AR-V7, SLC40A1, SELENBP1 and METTL7A were also found to be regulated by truncated AR variants in this study. However, DYNLT3 and IGFBP3 were identified as genes commonly regulated by AR-V7 and full-length AR, but identified as preferentially regulated by AR variants in this study. Similarly, KLK15 was found to be regulated by AR variants in this present study, but previously reported as being independently regulated by full-length AR in 22Rv1 cells (Guo *et al.*, 2009). At the time Guo and colleagues reported their findings in 2009, only five alternatively spliced truncated AR variants had been discovered, therefore, their approach to knockdown an individual AR variant in isolation, in a cell line which expresses multiple truncated AR variants, may explain the differences noted in this study, in which all truncated AR isoforms were specifically inhibited.

Furthermore, it suggests that specific individual AR variants and full-length AR may share overlapping target genes.

The first reports examining AR variant target gene expression were performed on RNA isolated from LNCaP cells ectopically expressing AR-V7 (Hu *et al.*, 2009) or ARV567es (Sun *et al.*, 2010) demonstrating that truncated AR variants could induce canonical androgen-regulated genes such as PSA, KLK2, NKX3.1, FKBP5 and TMPRSS2 in the absence of androgens. These data are in contrast to studies in which gene expression profiling was performed in cell-based models of CRPC (Guo *et al.*, 2009) or clinical CRPC specimens (Hornberg *et al.*, 2011) which endogenously express truncated AR variants, revealing some but not all classic AR-regulated genes (e.g. PSA, TMPRSS2, FKPB5) are regulated by AR variants. In combination with the data presented here, it indicates that the repertoire of AR variant target genes is likely to be highly complex, with overlap with known AR-gene targets and implicating novel AR gene targets.

To determine the clinical significance of the truncated AR variant gene expression profile identified in 22Rv1 cells, a subset of the AR variant expression profile with a fold change greater than 1.3 (297 probes in total, corresponding to 284 unique genes) was analyzed against mRNA expression data sets derived from 95 human PCa tissues specimens procured from 81 patients who underwent prostatectomy. Unsupervised clustering of the 95 specimens by the 297 probe set AR variant gene signature alone, generated three major clusters, separating the benign specimens from two clusters of malignant specimens. Interestingly, between the two malignant clusters, the lower Gleason score (6-7) specimens clustered together, whereas the higher Gleason score (8-10) specimens clustered separately. Furthermore, of the 15 patients who experienced biochemical failure, specimens from 12 of these patients clustered in the higher Gleason score group. Therefore, the AR variant gene signature can not only distinguish between benign and malignant PCa specimens, but may potentially be relevant to biomarker discovery.

Several truncated AR variants have consistently been shown to be expressed at the highest levels in CRPC specimens including AR-V7 (Hu *et al.*, 2009; Guo *et al.*, 2009; Zhang *et al.*, 2011), AR-V1 (Hu *et al.*, 2009), AR-V9, AR-V12 (Hu *et al.*, 2011) and

ARV567es (Zhang *et al.*, 2011). However, AR variants have also been reported to be expressed at variable levels in benign prostate epithelium, as well as hormone-naïve primary PCa specimens (Hu *et al.*, 2009; Guo *et al.*, 2009; Sun *et al.*, 2010; Hu *et al.*, 2011). In one study, a subset of hormone-naïve primary PCa specimens expressed AR-V1 and AR-V7 transcript at levels comparable to those in CRPC specimens. Elevated AR-V7 mRNA expression in this subset of patients was associated with biochemical recurrence following prostatectomy (Hu *et al.*, 2009). These data suggest that AR variants are not pathognomonic for CRPC, but may play a role in disease progression to CRPC and/or a more aggressive phenotype of PCa. In order to address this further, the expression of the 297 probe set AR variant gene signature was correlated with two characteristics in the clinical cohort: (1) Gleason pattern (GP) and lymph mode (LN) metastases status and (2) biochemical recurrence. Both of these characteristics represent more aggressive phenotypes of PCa and may identify AR variant regulated genes responsible for more aggressive disease. Spearman rank analysis correlation revealed 86 probes positively correlated and 80 probes negatively correlated with GP and LN status in the 68 malignant tumor specimens (see Table IV.9). Univariate analysis of time to PSA failure revealed 34 probe sets significantly correlated with PSA failure (see Table IV.10).

Many of the upregulated AR variant target genes that are significantly correlated with disease progression in the clinical cohort have not been implicated in PCa disease progression previously, including TAX1BP1, SDC4, NRP1, LEF1, LAMP3 and IGFBP3. However, these genes have been implicated in pro-oncogenic processes in other cancers.

TAX is a regulatory protein of human T-cell leukemia virus type 1 involved in transcriptional regulation of different sets of cellular genes that include some tumor suppressor genes (Yoshida, 1993). TAX has been shown to cooperate with Ras to induce neoplastic transformation (Pozzatti *et al.*, 1990). TAX1BP1 (TAX1 binding protein 1) can modulate some biological effects of TAX1 and its anti-apoptotic activity has been previously reported (De Valck *et al.*, 1999). More recently, it has been reported that TAX1 overexpression in cancer cell lines of liver cancer SMMC-

7721, colon cancer HCT-116, and breast cancer MCF-7 lead to cell proliferation via the Ras-Raf-MEK-ERK signaling pathway (Song *et al.*, 2009).

SDC4 (syndecan-4/ryudocan core protein) is a cell-surface heparan sulfate proteoglycan which play diverse roles in tumor biology by mediating adhesion and migration and cellular responses to mitogenic and angiogenic growth factors (Blackhall *et al.*, 2001). Specifically, SDC4 has been found to be expressed in normal ovary as well as benign and malignant ovarian tumors (Davies *et al.*, 2004). SDC4 has also been implicated in TRAIL-induced apoptosis in PCa acting through a protein kinase Cα-dependent mechanism (Franzen *et al.*, 2009).

NRP1 (neuropillin-1) is a non-tyrosine kinase receptor for class-3 semaphorins in neuronal guidance, and for the angiogenic cytokine vascular endothelial growth factor (VEGF) in vascular development (Pellet-Many *et al.*, 2008). Consistent with the data presented in this study, NRP1 has previously been reported to be positively associated with PCa progression and bone metastasis in both human PCa specimens (Latil *et al.*, 2000; Kosari *et al.*, 2008; Zhang *et al.*, 2010) and also xenograft tissues (Zhang *et al.*, 2010). Functionally, NRP1 has been demonstrated to promote tumor angiogenesis and progression using a tetracycline-inducible promoter model in rats (Miao *et al.*, 2000). NRP1 antagonism also increases the cytotoxic effects of the chemotherapeutic agents 5-FU, paclitaxel and cisplatin on DU145 cells, through inhibition of integrin-dependent cell interaction with the extracellular matrix (Jia *et al.*, 2010). Additionally in this study, increased expression of NRP1 was significantly associated with biochemical recurrence. Taken together, these data indicate that NRP1 may play an important role in truncated AR variant regulated disease progression.

Increased expression of LEF1 (lymphoid enhancer-binding factor 1) was also found to correlate with disease progression. LEF1 which participates in the Wnt signaling pathway has been indicated to regulate AR expression with eight putative LEF1 binding sites described at the AR promoter region (Yang *et al.*, 2006). LEF1 expression has been shown to be increased in an androgen-independent LNCaP subline, LNCaP-AI compared to parental LNCaP cells (Li *et al.*, 2009). Overexpression of LEF1 in LNCaP cells resulted in increased AR expression and

consequently enhanced growth and invasion ability, while LEF1 knockdown in LNCaP-AI cells decreased AR expression and subsequently the growth and invasion capacity (Li *et al.*, 2009). This indicates that LEF1 is associated with androgen-independent PCa and may potentially serve as a marker for androgen-independent disease.

Increased IGFBP3 (insulin-like growth factor binding protein 3) expression was found to correlate significantly with disease progression and also biochemical recurrence. These data are consistent with published findings that IGFBP3 has tumor-suppressive properties (Yamada and Lee, 2009). Several clinical studies have shown that serum IGF-I is elevated and IGFBP3 levels reduced in patients before the diagnosis of PCa and that increased serum levels of IGFBP3 are associated with a decreased risk and better prognosis (Chan *et al.*, 1998; Chan *et al.*, 2002; Shariat *et al.*, 2002).

Although LAMP3 (lysosome-associated membrane glycoprotein 3) has not been implicated in PCa, high expression of LAMP3 has been detected in primary cancers of the esophagus, colon, ovary and breast (Ozaki *et al.*, 1998). Functionally, it has been shown that LAMP3 overexpression is associated with increased metastatic potential both *in vitro* and *in vivo* in uterine cervical cancer (Kanao *et al.*, 2005).

In addition to identifying AR regulated genes upregulated with disease progression, a large number of downregulated genes that correlate with disease progression were also identified, including SPRY2, MBP, KLK15 and CCPG1. Interestingly, these genes have been implicated as tumor suppressors in a variety of cancer models. SPRY2 (sprouty homolog 2) is a tumor suppressor found to be underexpressed in an androgen-independent LNCaP subline compared to an androgen-dependent equivalent using genome-wide expression profiling (Singh *et al.*, 2008). It has also been shown to be epigenetically inactivated in PCa, further implicating a role for SPRY2 as a tumor suppressor locus in PCa (McKie *et al.*, 2005). RASD1 is a member of the Ras superfamily of small G-proteins that often promotes cell growth and tumor expansion (Takai *et al.*, 2001). However, RASD1 has been shown to suppress cell growth *in vitro* and also *in vivo* in breast cancer (Vaidyanathan *et al.*, 2004). BNIP3L (BCL2/adenovirus E1B 19-kDa interacting

protein) is also believed to function as a tumor suppressor in PCa. Using a single nucleotide polymorphic array approach, BNIP3L was found to be homozygously deleted in a panel of PCa cell lines and PCa tumor samples (Liu *et al.*, 2008) and also loss of heterogeneity in PCa tumors and significantly associated with increasing disease stage (Cheng *et al.*, 2012).

Overall, this study has identified a distinct truncated AR variant transcriptome derived from an *in vitro* model system and validated these results using expression profiles from a cohort of clinical PCa specimens. This has uncovered a number of putative AR variant regulated pro-oncogenes and tumor suppressor genes which are upregulated and downregulated respectively with disease progression. Further analysis and characterization of the signaling networks regulated by these genes in PCa may identify novel druggable targets in patients found to express truncated AR variants.

Remarkable advances in our understanding of the underlying mechanisms of AR in disease progression to CRPC have been made over the last decade. There is now incontrovertible evidence that constitutively active truncated AR variants represent another novel mechanism in disease progression, however much work is still required to determine the function and mechanisms of the individual variants. Differences in transcriptional activity of different AR variants have been reported and further studies are required to understand the molecular basis for these differences. It is possible that the different COOH-terminus extensions of AR variants may provide important regulatory information, in addition to harboring the premature stop codon, which may influence DNA-binding activity or explain the cellular localizations of different AR variants. Furthermore, the mechanisms by which full-length AR can influence AR variant function (and *vice versa*) will also require further investigation. Analysis of the 3,714 probes identified in this study thought to be regulated by both truncated AR variants and full-length AR in 22Rv1 cells may provide a critical insight into the mechanism of AR variant regulation in the context of full-length AR. This is likely to be a highly complex relationship and amenable to next generation sequencing technologies.

The gene expression program regulated by AR variants identified in this study remains to be fully characterized. Of particular importance is to validate the AR variant gene signature identified in 22Rv1 cells in a large cohort of well-annotated CRPC specimens. Moreover, a thorough analysis of the AR variant gene expression program in an additional cell line such as the VCaP cell model is warranted. The observations presented here that androgen depletion *in vitro* in this cell line stimulates transcript and protein expression of a number of truncated AR variants may provide an opportunity to identify the AR variant regulated gene expression program in response to androgen withdrawal. This may elucidate the mechanisms critical to disease progression in CRPC patients who have failed ADT. Ultimately, these data may identify patients at high risk of failing ADT and who may benefit from intervention with emerging AR NTD target therapies such as EPI-001.

In the future, one of the most significant challenges will be the development and validation of a reliable, reproducible means of quantifying the levels and frequency of alternatively spliced AR variants relative to full-length AR in clinical specimens. This is of fundamental importance to establish the functional relevance of AR variants in disease progression and for the successful development and implementation of novel targeted strategies for patients with advanced stages of PCa, particularly those with progressive CRPC.

VIII. Bibliography

Cancer in Ireland 2011: Annual Report of the National Cancer Registry. National Cancer Registry Ireland, Cork, Ireland.

Agoulnik I.U., Weigel N.L. (2008) Androgen receptor coactivators and prostate cancer. *Advances in Experimental Medicine and Biology*, **617**, 245-255.

Agus D.B., Cordon-Cardo C., Fox W., Drobnjak M., Koff A., Golde D.W., Scher H.I. (1999) Prostate Cancer Cell Cycle Regulators: Response to Androgen Withdrawal and Development of Androgen Independence. *Journal of the National Cancer Institute*, **91**, 1869-1876.

Ahrens-Fath I., Politz O., Geserick C., Haendler B. (2005) Androgen receptor function is modulated by the tissue-specific AR45 variant. *FEBS Journal*, **272**, 74-84.

Andersen R.J., Mawji N.R., Wang J., Wang G., Haile S., Myung J.-K., Watt K., Tam T., Yang Y.C., Banuelos C.A., Williams D.E., McEwan I.J., Wang Y., Sadar M.D. (2010) Regression of Castrate-Recurrent Prostate Cancer by a Small-Molecule Inhibitor of the Amino-Terminus Domain of the Androgen Receptor. *Cancer Cell*, **17**, 535-546.

Bain D.L., Heneghan A.F., Connaghan-Jones K.D., Miura M.T. (2007) Nuclear Receptor Structure: Implications for Function. *Annual Review of Physiology*, **69**, 201-220.

Balk S.P. (2002) Androgen receptor as a target in androgen-independent prostate cancer. *Urology*, **60**, 132-138.

Ballman K.V., Grill D.E., Oberg A.L., Therneau T.M. (2004) Faster cyclic loess: normalizing RNA arrays via linear models. *Bioinformatics*, **20**, 2778-2786.

Benjamini Y., Hochberg Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, 289-300.

Blackhall F.H., Merry C.L.R., Davies E.J., Jayson G.C. (2001) Heparan sulfate proteoglycans and cancer. *British Journal of Cancer*, **85**, 1094-1098.

Blum H., Beier H., Gross H.J. (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis*, **8**, 93-99.

Brinkmann A.O. (2001) Molecular basis of androgen insensitivity. *Molecular and Cellular Endocrinology*, **179**, 105-109.

Brinkmann A.O., Kuiper G.G.J.M., de Boer W., Mulder E., Bolt J., Van Steenbrugge G.J., Van Der Molen H.J. (1986) Characterization of androgen receptors after photoaffinity labelling with [3H]methyltrienolone (R1881). *Journal of Steroid Biochemistry*, **24**, 245-249.

Brodsky L.I., Ivanov V.V., Kalaydzidis Y.L., Leontovich A.M., Nikolaev V.K., Feranchuk S.I., Drachev V.A. (1995) GeneBee-NET: Internet-based server for analyzing biopolymers structure. *Biochemistry (Moscow)*, **60**, 923-928.

Brooke G.N., Bevan C.L. (2009) The role of androgen receptor mutations in prostate cancer progression. *Current Genomics*, **10**, 18-25.

Bruggenwirth H.T., Boehmer A.L.M., Ramnarain S., Verleun-Mooijman M.C.T., Satijn D.P.E., Trapman J., Grootegoed J.A., Brinkmann A.O. (1997) Molecular Analysis of the Androgen-Receptor Gene in a Family with Receptor-Positive Partial Androgen Insensitivity: An Unusual Type of Intronic Mutation. *American Journal of Human Genetics*, **61**, 1067-1077.

Buchanan G., Irvine R.A., Coetzee G.A., Tilley W.D. (2001) Contribution of the androgen receptor to prostate cancer predisposition and progression. *Cancer Metastasis Reviews*, **20**, 207-223.

Burd C.J., Morey L.M., Knudsen K.E. (2006) Androgen receptor corepressors and prostate cancer. *Endocrine-Related Cancer*, **13**, 979-994.

Callewaert L., Van Tilborgh N., Claessens F. (2006) Interplay between Two Hormone-Independent Activation Domains in the Androgen Receptor. *Cancer Research*, **66**, 543-553.

Ceraline J., Cruchant M.D., Erdmann E., Erbs P., Kurtz J.-E., Duclos B., Jacqmin D., Chopin D., Bergerat J.-P. (2004) Constitutive activation of the androgen receptor by a point mutation in the hinge region: A new mechanism for androgen-independent growth in prostate cancer. *International Journal of Cancer*, **108**, 152-157.

Chamberlain N.L., Whitacre D.C., Miesfeld R.L. (1996) Delineation of Two Distinct Type 1 Activation Functions in the Androgen Receptor Amino-terminal Domain. *Journal of Biological Chemistry*, **271**, 26772-26778.

Chan J.M., Stampfer M.J., Giovannucci E., Gann P.H., Ma J., Wilkinson P., Hennekens C.H., Pollak M. (1998) Plasma Insulin-Like Growth Factor-I and Prostate Cancer Risk: A Prospective Study. *Science*, **279**, 563-566.

Chan J.M., Stampfer M.J., Ma J., Gann P., Gaziano J.M., Pollak M., Giovannucci E. (2002) Insulin-Like Growth Factor-I (IGF-I) and IGF Binding Protein-3 as Predictors of Advanced-Stage Prostate Cancer. *Journal of the National Cancer Institute*, **94**, 1099-1106.

Chan S.C., Li Y., Dehm S.M. (2012) Androgen Receptor Splice Variants Activate Androgen Receptor Target Genes and Support Aberrant Prostate Cancer Cell Growth Independent of Canonical Androgen Receptor Nuclear Localization Signal. *Journal of Biological Chemistry*, **287**, 19736-19749.

Chang C.S., Kokontis J., Liao S.T. (1988) Structural analysis of complementary DNA and amino acid sequences of human and rat androgen receptors. *Proceedings of the National Academy of Sciences*, **85**, 7211-7215.

Chen C.D., Welsbie D.S., Tran C., Baek S.H., Chen R., Vessella R., Rosenfeld M.G., Sawyers C.L. (2004) Molecular determinants of resistance to antiandrogen therapy. *Nature Medicine*, **10**, 33-39.

Chen H., Libertini S.J., Wang Y., Kung H.-J., Ghosh P., Mudryj M. (2010) ERK Regulates Calpain 2-induced Androgen Receptor Proteolysis in CWR22 Relapsed Prostate Tumor Cell Lines. *Journal of Biological Chemistry*, **285**, 2368-2374.

Cheng I., Levin A.M., Tai Y.C., Plummer S., Chen G.K., Neslund-Dudas C., Casey G., Rybicki B.A., Witte J.S. (2012) Copy number alterations in prostate tumors and disease aggressiveness. *Genes, Chromosomes and Cancer*, **51**, 66-76.

Clegg N.J., Wongvipat J., Joseph J.D., Tran C., Ouk S., Dilhas A., Chen Y., Grillot K., Bischoff E.D., Cai L., Aparicio A., Dorow S., Arora V., Shao G., Qian J., Zhao H., Yang G., Cao C., Sensintaffar J., Wasielewska T., Herbert M.R., Bonnefous C., Darimont B., Scher H.I., Smith-Jones P., Klang M., Smith N.D., De Stanchina E., Wu N., Ouerfelli O., Rix P.J., Heyman R.A., Jung M.E., Sawyers C.L., Hager J.H. (2012) ARN-509: A Novel Antiandrogen for Prostate Cancer Treatment. *Cancer Research*, **72**, 1494-1503.

Crawford E.D., Eisenberger M.A., McLeod D.G., Spaulding J.T., Benson R., Dorr F.A., Blumenstein B.A., Davis M.A., Goodman P.J. (1989) A Controlled Trial of Leuprolide with and without Flutamide in Prostatic Carcinoma. *New England Journal of Medicine*, **321**, 419-424.

Culig Z., Hobisch A., Cronauer M.V., Cato A.C., Hittmair A., Radmayr C., Eberle J., Bartsch G., Klocker H. (1993) Mutant androgen receptor detected in an advanced-stage prostatic carcinoma is activated by adrenal androgens and progesterone. *Molecular Endocrinology*, **7**, 1541-1550.

Cutress M.L., Whitaker H.C., Mills I.G., Stewart M., Neal D.E. (2008) Structural basis for the nuclear import of the human androgen receptor. *Journal of Cell Science*, **121**, 957-968.

Davies E.J., Blackhall F.H., Shanks J.H., David G., McGown A.T., Swindell R., Slade R.J., Martin-Hirsch P., Gallagher J.T., Jayson G.C. (2004) Distribution and Clinical Significance of Heparan Sulfate Proteoglycans in Ovarian Cancer. *Clinical Cancer Research*, **10**, 5178-5186.

de Bono J.S., Logothetis C.J., Molina A., Fizazi K., North S., Chu L., Chi K.N., Jones R.J., Goodman O.B., Saad F., Staffurth J.N., Mainwaring P., Harland S., Flaig T.W., Hutson T.E., Cheng T., Patterson H., Hainsworth J.D., Ryan C.J., Sternberg C.N., Ellard S.L., Flechon A., Saleh M., Scholz M., Efstathiou E., Zivi A., Bianchini D., Loriot Y., Chieffo N., Kheoh T., Haqq C.M., Scher H.I. (2011) Abiraterone and Increased Survival in Metastatic Prostate Cancer. *New England Journal of Medicine*, **364**, 1995-2005.

De Valck D., Jin D.Y., Heyninck K., Van De Craen M., Contreras R., Fiers W., Jeang K.T., Beyaert R. (1999) The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. *Oncogene*, **18**, 4182-4190.

Dehm S.M., Regan K.M., Schmidt L.J., Tindall D.J. (2007) Selective Role of an NH2-Terminal WxxLF Motif for Aberrant Androgen Receptor Activation in Androgen Depletion–Independent Prostate Cancer Cells. *Cancer Research*, **67**, 10067-10077.

Dehm S.M., Schmidt L.J., Heemers H.V., Vessella R.L., Tindall D.J. (2008) Splicing of a Novel Androgen Receptor Exon Generates a Constitutively Active Androgen Receptor that Mediates Prostate Cancer Therapy Resistance. *Cancer Research*, **68**, 5469-5477.

Dehm S.M., Tindall D.J. (2006) Molecular regulation of androgen action in prostate cancer. *Journal of Cellular Biochemistry*, **99**, 333-344.

Denis L.J., Carneiro De Moura J.L., Bono A., Sylvester R., Whelan P., Newling D., Depauw M. (1993) Goserelin acetate and flutamide versus bilateral orchiectomy: a phase III eortc trial (30853). *Urology*, **42**, 119-130.

Ding D., Xu L., Menon M., Reddy G.P.V., Barrack E.R. (2004) Effect of a short CAG (glutamine) repeat on human androgen receptor function. *Prostate*, **58**, 23-32.

Ding D., Xu L., Menon M., Reddy G.P.V., Barrack E.R. (2005) Effect of GGC (glycine) repeat length polymorphism in the human androgen receptor on androgen action. *Prostate*, **62**, 133-139.

Dirac A.M.G., Bernards R. (2010) The Deubiquitinating Enzyme USP26 Is a Regulator of Androgen Receptor Signaling. *Molecular Cancer Research*, **8**, 844-854.

Donovan M.J., Osman I., Khan F.M., Vengrenyuk Y., Capodieci P., Koscuiszka M., Anand A., Cordon-Cardo C., Costa J., Scher H.I. (2010) Androgen receptor expression is associated with prostate cancer-specific survival in castrate patients with metastatic disease. *BJU International*, **105**, 462-467.

Dufort I., Rheault P., Huang X.F., Soucy P., Luu-The V. (1999) Characteristics of a highly labile human type 5 17β-hydroxysteroid dehydrogenase. *Endocrinology*, **140**, 568-574.

Eisenberger M.A., Blumenstein B.A., Crawford E.D., Miller G., McLeod D.G., Loehrer P.J., Wilding G., Sears K., Culkin D.J., Thompson I.M., Bueschen A.J., Lowe B.A. (1998) Bilateral Orchiectomy with or without Flutamide for Metastatic Prostate Cancer. *New England Journal of Medicine*, **339**, 1036-1042.

El-Alfy M., Luu-The V., Huang X.F., Berger L., Labrie F., Pelletier G. (1999) Localization of type 5 17β-hydroxysteroid dehydrogenase, 3β- hydroxysteroid dehydrogenase, and androgen receptor in the human prostate by in situ hybridization and immunocytochemistry. *Endocrinology*, **140**, 1481-1491.

Ellwood-Yen K., Graeber T.G., Wongvipat J., Iruela-Arispe M.L., Zhang J., Matusik R., Thomas G.V., Sawyers C.L. (2003) Myc-driven murine prostate cancer shares molecular features with human prostate tumors. *Cancer Cell*, **4**, 223-238.

Estebanez-Perpina E., Moore J.M.R., Mar E., Delgado-Rodrigues E., Nguyen P., Baxter J.D., Buehrer B.M., Webb P., Fletterick R.J., Guy R.K. (2005) The Molecular Mechanisms of Coactivator Utilization in Ligand-dependent Transactivation by the Androgen Receptor. *Journal of Biological Chemistry*, **280**, 8060-8068.

Feldman B.J., Feldman D. (2001) The development of androgen-independent prostate cancer. *Nature Reviews Cancer*, **1**, 34-45.

Ferro P., Catalano M.G., Dell Eva R., Fortunati N., Pfeffer U. (2002) The androgen receptor CAG repeat: a modifier of carcinogenesis? *Molecular and Cellular Endocrinology*, **193**, 109-120.

Franzen C.A., Chen C.-C., Todorović V., Juric V., Monzon R.I., Lau L.F. (2009) Matrix Protein CCN1 Is Critical for Prostate Carcinoma Cell Proliferation and TRAIL-Induced Apoptosis. *Molecular Cancer Research*, **7**, 1045-1055.

Gelmann E.P. (2002) Molecular Biology of the Androgen Receptor. *Journal of Clinical Oncology*, **20**, 3001-3015.

Gottlieb B., Beitel L.K., Nadarajah A., Paliouras M., Trifiro M. (2012) The androgen receptor gene mutations database: 2012 update. *Human Mutation*, **33**, 887-894.

Gottlieb B., Beitel L.K., Wu J.H., Trifiro M. (2004) The androgen receptor gene mutations database (ARDB): 2004 update. *Human Mutation*, **23**, 527-533.

Gregory C.W., Johnson R.T., Mohler J.L., French F.S., Wilson E.M. (2001) Androgen Receptor Stabilization in Recurrent Prostate Cancer Is Associated with Hypersensitivity to Low Androgen. *Cancer Research*, **61**, 2892-2898.

Gudnason H., Dufva M., Bang D.D., Wolff A. (1997) Comparison of multiple DNA dyes for real-time PCR: effects of dye concentration and sequence composition on DNA amplification and melting temperature. *Nucleic Acids Research*, **35**, e127.

Guo Z., Dai B., Jiang T., Xu K., Xie Y., Kim O., Nesheiwat I., Kong X., Melamed J., Handratta V.D., Njar V.C.O., Brodie A.M.H., Yu L.-R., Veenstra T.D., Chen H., Qiu Y. (2006) Regulation of androgen receptor activity by tyrosine phosphorylation. *Cancer Cell*, **10**, 309-319.

Guo Z., Yang X., Sun F., Jiang R., Linn D.E., Chen H., Chen H., Kong X., Melamed J., Tepper C.G., Kung H.-J., Brodie A.M.H., Edwards J., Qiu Y. (2009) A Novel Androgen Receptor Splice Variant Is Up-regulated during Prostate Cancer Progression and Promotes Androgen Depletion–Resistant Growth. *Cancer Research*, **69**, 2305-2313.

Hara T., Miyazaki J.I., Araki H., Yamaoka M., Kanzaki N., Kusaka M., Miyamoto M. (2003) Novel mutations of androgen receptor: A possible mechanism of bicalutamide withdrawal syndrome. *Cancer Research*, **63**, 149-153.

He B., Gampe R.T., Kole A.J., Hnat A.T., Stanley T.B., An G., Stewart E.L., Kalman R.I., Minges J.T., Wilson E.M. (2004) Structural Basis for Androgen Receptor Interdomain and Coactivator Interactions Suggests a Transition in Nuclear Receptor Activation Function Dominance. *Molecular Cell*, **16**, 425-438.

Heemers H.V., Tindall D.J. (2007) Androgen Receptor (AR) Coregulators: A Diversity of Functions Converging on and Regulating the AR Transcriptional Complex. *Endocrine Reviews*, **28**, 778-808.

Heinlein C.A., Chang C. (2004) Androgen Receptor in Prostate Cancer. *Endocrine Reviews*, **25**, 276-308.

Hellwinkel O.J.-C., Bull K., Holterhus P.-M., Homburg N., Struve D., Hiort O. (1999) Complete androgen insensitivity caused by a splice donor site mutation in intron 2 of the human androgen receptor gene resulting in an exon 2-lacking transcript with premature stop-codon and reduced expression. *Journal of Steroid Biochemistry and Molecular Biology*, **68**, 1-9.

Hellwinkel O.J.-C., Holterhus P.-M., Struve D., Marschke C., Homburg N., Hiort O. (2001) A Unique Exonic Splicing Mutation in the Human Androgen Receptor Gene Indicates a Physiologic Relevance of Regular Androgen Receptor Transcript Variants. *Journal of Clinical Endocrinology & Metabolism*, **86**, 2569-2575.

Holzbeierlein J., Lal P., LaTulippe E., Smith A., Satagopan J., Zhang L., Ryan C., Smith S., Scher H., Scardino P., Reuter V., Gerald W.L. (2004) Gene Expression Analysis of Human Prostate Carcinoma during Hormonal Therapy Identifies Androgen-Responsive Genes and Mechanisms of Therapy Resistance. *American Journal of Pathology*, **164**, 217-227.

Hornberg E., Ylitalo E.B., Crnalic S., Antti H., Stattin P., Widmark A., Bergh A., Wikström P. (2011) Expression of Androgen Receptor Splice Variants in Prostate Cancer Bone Metastases is Associated with Castration-Resistance and Short Survival. *PLoS ONE*, **6**, e19059.

Hu R., Dunn T.A., Wei S., Isharwal S., Veltri R.W., Humphreys E., Han M., Partin A.W., Vessella R.L., Isaacs W.B., Bova G.S., Luo J. (2009) Ligand-Independent Androgen Receptor Variants Derived from Splicing of Cryptic Exons Signify Hormone-Refractory Prostate Cancer. *Cancer Research*, **69**, 16-22.

Hu R., Isaacs W.B., Luo J. (2011) A snapshot of the expression signature of androgen receptor splicing variants and their distinctive transcriptional activities. *Prostate*, **71**, 1656-1667.

Hu R., Lu C., Mostaghel E.A., Yegnasubramanian S., Gurel M., Tannahill C., Edwards J., Isaacs W., Nelson P.S., Bluemn E., Plymate S.R., Luo J. (2012) Distinct transcriptional programs mediated by the ligand-dependent full-length androgen receptor and its splice variants in castration-resistant prostate cancer. *Cancer Research*, **72**, 3457-3462.

Huggins C. (1942) Effect of Orchiectomy and Irradiation on Cancer of the Prostate. *Annals of Surgery*, **115**, 1192-1200.

Huggins C., Hodges C.V. (1941) Studies on Prostatic Cancer. I. The Effect of Castration, of Estrogen and of Androgen Injection on Serum Phosphatases in Metastatic Carcinoma of the Prostate. *Cancer Research*, **1**, 293-297.

Huggins C., Stevens R.E., Hodges C.V. (1941) Studies On Prostatic Cancer: II. The Effects Of Castration On Advanced Carcinoma Of The Prostate Gland. *Archives of Surgery*, **43**, 209-223.

Hunter J. (1840) Observations on certain parts of the animal economy. Haswell, Barrington and Haswell, Philadelphia, Pennsylvania.

Hur E., Pfaff S.J., Payne E.S., Gron H., Buehrer B.M., Fletterick R.J. (2004) Recognition and Accommodation at the Androgen Receptor Coactivator Binding Interface. *PLoS Biology*, **2**, e274.

Irizarry R.A., Hobbs B., Collin F., Beazer-Barclay Y.D., Antonellis K.J., Scherf U., Speed T.P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249-264.

Jagla M., Feve M., Kessler P., Lapouge G., Erdmann E., Serra S., Bergerat J.-P., Ceraline J. (2007) A Splicing Variant of the Androgen Receptor Detected in a Metastatic Prostate Cancer Exhibits Exclusively Cytoplasmic Actions. *Endocrinology*, **148**, 4334-4343.

Jenster G., van der Korput H.A.G.M., van Vroonhoven C., van der Kwast T.H., Trapman J., Brinkmann A.O. (1991) Domains of the Human Androgen Receptor Involved in Steroid Binding, Transcriptional Activation, and Subcellular Localization. *Molecular Endocrinology*, **5**, 1396-1404.

Ji Q., Chang L., VanDenBerg D., Stanczyk F.Z., Stolz A. (2003) Selective reduction of AKR1C2 in prostate cancer and its role in DHT metabolism. *Prostate*, **54**, 275-289.

Jia H., Cheng L., Tickner M., Bagherzadeh A., Selwood D., Zachary I. (2010) Neuropilin-1 antagonism in human carcinoma cells inhibits migration and enhances chemosensitivity. *British Journal of Cancer*, **102**, 541-552.

Jiang Y., Palma J.F., Agus D.B., Wang Y., Gross M.E. (2010) Detection of Androgen Receptor Mutations in Circulating Tumor Cells in Castration-Resistant Prostate Cancer. *Clinical Chemistry*, **56**, 1492-1495.

Kaikkonen S., Jaaskelainen T., Karvonen U., Rytinki M.M., Makkonen H., Gioeli D., Paschal B.M., Palvimo J.J. (2009) SUMO-Specific Protease 1 (SENP1) Reverses the Hormone-Augmented SUMOylation of Androgen Receptor and Modulates Gene Responses in Prostate Cancer Cells. *Molecular Endocrinology*, **23**, 292-307.

Kanao H., Enomoto T., Kimura T., Fujita M., Nakashima R., Ueda Y., Ueno Y., Miyatake T., Yoshizaki T., Buzard G.S., Tanigami A., Yoshino K., Murata Y. (2005) Overexpression of LAMP3/TSC403/DC-LAMP Promotes Metastasis in Uterine Cervical Cancer. *Cancer Research*, **65**, 8640-8645.

Klein E.A., Ciezki J., Kupelian P.A., Mahadevan A. (2009) Outcomes for intermediate risk prostate cancer: Are there advantages for surgery, external radiation, or brachytherapy? *Urologic Oncology: Seminars and Original Investigations*, **27**, 67-71.

Koivisto P., Kononen J., Palmberg C., Tammela T., Hyytinen E., Isola J., Trapman J., Cleutjens K., Noordzij A., Visakorpi T., Kallioniemi O.P. (1997) Androgen receptor gene amplification: A possible molecular mechanism for androgen deprivation therapy failure in prostate cancer. *Cancer Research*, **57**, 314-319.

Korenchuk S., Lehr J.E., L M.C., Lee Y.G., Whitney S., Vessella R., Lin D.L., Pienta K.J. (2001) VCaP, a cell-based model system of human prostate cancer. *In vivo*, **15**, 163-168.

Kosari F., Munz J.M.A., Savci-Heijink C.D., Spiro C., Klee E.W., Kube D.M., Tillmans L., Slezak J., Karnes R.J., Cheville J.C., Vasmatzis G. (2008) Identification of Prognostic Biomarkers for Prostate Cancer. *Clinical Cancer Research*, **14**, 1734-1743.

Kube D., Savci-Heijink C., Lamblin A.-F., Kosari F., Vasmatzis G., Cheville J., Connelly D., Klee G. (2007) Optimization of laser capture microdissection and RNA amplification for gene expression profiling of prostate cancer. *BMC Molecular Biology*, **8**, 25.

Lapouge G., Erdmann E., Marcias G., Jagla M., Monge A., Kessler P., Serra S., Lang H., Jacqmin D., Bergerat J.-P., Ceraline J. (2007) Unexpected paracrine action of prostate cancer cells harboring a new class of androgen receptor mutation—A new paradigm for cooperation among prostate tumor cells. *International Journal of Cancer*, **121**, 1238-1244.

Latil A., Bieche I., Pesche S., Valeri A., Fournier G., Cussenot O., Lidereau R. (2000) VEGF overexpression in clinically localized prostate tumors and neuropilin-1 overexpression in metastatic forms. *International Journal of Cancer*, **89**, 167-171.

Leuprolide (1984) Leuprolide versus Diethylstilbestrol for Metastatic Prostate Cancer. *New England Journal of Medicine*, **311**, 1281-1286.

Li Y., Alsagabi M., Fan D., Bova G.S., Tewfik A.H., Dehm S.M. (2011) Intragenic Rearrangement and Altered RNA Splicing of the Androgen Receptor in a Cell-Based Model of Prostate Cancer Progression. *Cancer Research*, **71**, 2108-2117.

Li Y., Hwang T.H., Oseth L.A., Hauge A., Vessella R.L., Schmechel S.C., Hirsch B., Beckman K.B., Silverstein K.A., Dehm S.M. (2012) AR intragenic deletions linked to androgen receptor splice variant expression and activity in models of prostate cancer progression. *Oncogene*, In press (DOI: 10.1038/onc.2011.1637).

Li Y., Wang L., Zhang M., Melamed J., Liu X., Reiter R., Wei J., Peng Y., Zou X., Pellicer A., Garabedian M.J., Ferrari A., Lee P. (2009) LEF1 in Androgen-Independent Prostate Cancer: Regulation of Androgen Receptor Expression, Prostate Cancer Growth, and Invasion. *Cancer Research*, **69**, 3332-3338.

Libertini S.J., Tepper C.G., Rodriguez V., Asmuth D.M., Kung H.-J., Mudryj M. (2007) Evidence for Calpain-Mediated Androgen Receptor Cleavage as a Mechanism for Androgen Independence. *Cancer Research*, **67**, 9001-9005.

Lim J., Ghadessy F.J., Yong E.L. (1997) A novel splice site mutation in the androgen receptor gene results in exon skipping and a non-functional truncated protein. *Molecular and Cellular Endocrinology*, **131**, 205-210.

Lin H.K., Jez J.M., Schlegel B.P., Peehl D.M., Pachter J.A., Penning T.M. (1997) Expression and characterization of recombinant type 2 3α- hydroxysteroid dehydrogenase (HSD) from human prostate: Demonstration of bifunctional 3α/17β-HSD activity and cellular distribution. *Molecular Endocrinology*, **11**, 1971-1984.

Liu W., Xie C.C., Zhu Y., Li T., Sun J., Cheng Y., Ewing C.M., Dalrymple S., Turner A.R., Isaacs J.T., Chang B.L., Zheng S.L., Isaacs W.B., Xu J. (2008) Homozygous deletions and recurrent amplifications implicate new genes involved in prostate cancer. *Neoplasia*, **10**, 897-907.

Loberg R.D., St. John L.N., Day L.L., Neeley C.K., Pienta K.J. (2006) Development of the VCaP androgen-independent model of prostate cancer. *Urologic Oncology: Seminars and Original Investigations*, **24**, 161-168.

Loblaw D.A., Virgo K.S., Nam R., Somerfield M.R., Ben-Josef E., Mendelson D.S., Middleton R., Sharp S.A., Smith T.J., Talcott J., Taplin M., Vogelzang N.J., Wade J.L., Bennett C.L., Scher H.I. (2007) Initial Hormonal Management of Androgen-Sensitive Metastatic, Recurrent, or Progressive Prostate Cancer: 2007 Update of an American Society of Clinical Oncology Practice Guideline. *Journal of Clinical Oncology*, **25**, 1596-1605.

Locke J.A., Guns E.S., Lubik A.A., Adomat H.H., Hendy S.C., Wood C.A., Ettinger S.L., Gleave M.E., Nelson C.C. (2008) Androgen Levels Increase by Intratumoral De novo Steroidogenesis during Progression of Castration-Resistant Prostate Cancer. *Cancer Research*, **68**, 6407-6415.

Lowry O.H., Rosebrough N.J., Farr A.L., Randall R.J. (1951) Protein Measurement with the Folin Phenol Reagent. *Journal of Biological Chemistry*, **193**, 265-275.

Lubahn D.B., Brown T.R., Simental J.A., Higgs H.N., Migeon C.J., Wilson E.M., French F.S. (1989) Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. *Proceedings of the National Academy of Sciences*, **86**, 9534-9538.

Mahajan N.P., Liu Y., Majumder S., Warren M.R., Parker C.E., Mohler J.L., Earp H.S., Whang Y.E. (2007) Activated Cdc42-associated kinase Ack1 promotes prostate cancer progression via androgen receptor tyrosine phosphorylation. *Proceedings of the National Academy of Sciences*, **104**, 8438-8443.

Malovannaya A., Li Y., Bulynko Y., Jung S.Y., Wang Y., Lanz R.B., O'Malley B.W., Qin J. (2010) Streamlined analysis schema for high-throughput identification of endogenous protein complexes. *Proceedings of the National Academy of Sciences*, **107**, 2431-2436.

Marcias G., Erdmann E., Lapouge G., Siebert C., Barthelemy P., Duclos B., Bergerat J.-P., Ceraline J., Kurtz J.-E. (2010) Identification of novel truncated androgen receptor (AR) mutants including unreported pre-mRNA splicing variants in the 22Rv1 hormone-refractory prostate cancer (PCa) cell line. *Human Mutation*, **31**, 74-80.

Mashima T., Okabe S., Seimiya H. (2010) Pharmacological Targeting of Constitutively Active Truncated Androgen Receptor by Nigericin and Suppression of Hormone-Refractory Prostate Cancer Cell Growth. *Molecular Pharmacology*, **78**, 846-854.

Matias P.M., Donner P., Coelho R., Thomaz M., Peixoto C., Macedo S., Otto N., Joschko S., Scholz P., Wegg A., Basler S., Schäfer M., Egner U., Carrondo M.A. (2000) Structural Evidence for Ligand Specificity in the Binding Domain of the Human Androgen Receptor. *Journal of Biological Chemistry*, **275**, 26164-26171.

McKie A.B., Douglas D.A., Olijslagers S., Graham J., Omar M.M., Heer R., Gnanapragasam V.J., Robson C.N., Leung H.Y. (2005) Epigenetic inactivation of the human sprouty2 (hSPRY2) homologue in prostate cancer. *Oncogene*, **24**, 2166-2174.

Miao H.Q., Lee P., Lin H., Soker S., Klagsbrun M. (2000) Neuropilin-1 expression by tumor cells promotes tumor angiogenesis and progression. *FASEB Journal*, **14**, 2532-2539.

Mohler J.L., Gregory C.W., Ford O.H., Kim D., Weaver C.M., Petrusz P., Wilson E.M., French F.S. (2004) The Androgen Axis in Recurrent Prostate Cancer. *Clinical Cancer Research*, **10**, 440-448.

Montgomery R.B., Mostaghel E.A., Vessella R., Hess D.L., Kalhorn T.F., Higano C.S., True L.D., Nelson P.S. (2008) Maintenance of Intratumoral Androgens in Metastatic Prostate Cancer: A Mechanism for Castration-Resistant Tumor Growth. *Cancer Research*, **68**, 4447-4454.

Mostaghel E.A., Marck B.T., Plymate S.R., Vessella R.L., Balk S., Matsumoto A.M., Nelson P.S., Montgomery R.B. (2011) Resistance to CYP17A1 Inhibition with Abiraterone in Castration-Resistant Prostate Cancer: Induction of Steroidogenesis and Androgen Receptor Splice Variants. *Clinical Cancer Research*, **17**, 5913-5925.

Mostaghel E.A., Page S.T., Lin D.W., Fazli L., Coleman I.M., True L.D., Knudsen B., Hess D.L., Nelson C.C., Matsumoto A.M., Bremner W.J., Gleave M.E., Nelson P.S. (2007) Intraprostatic Androgens and Androgen-Regulated Gene Expression Persist after Testosterone Suppression: Therapeutic Implications for Castration-Resistant Prostate Cancer. *Cancer Research*, **67**, 5033-5041.

Moul J.W. (2009) Twenty years of controversy surrounding combined androgen blockade for advanced prostate cancer. *Cancer*, **115**, 3376-3378.

Nagabhushan M., Miller C.M., Pretlow T.P., Giaconia J.M., Edgehouse N.L., Schwartz S., Kung H.-J., de Vere White R.W., Gumerlock P.H., Resnick M.I., Amini S.B., Pretlow T.G. (1996) CWR22: The First Human Prostate Cancer Xenograft with Strongly Androgen-dependent and Relapsed Strains Both in Vivo and in Soft Agar. *Cancer Research*, **56**, 3042-3046.

Ozaki K., Nagata M., Suzuki M., Fujiwara T., Ueda K., Miyoshi Y., Takahashi E.-i., Nakamura Y. (1998) Isolation and Characterization of a Novel Human Lung-specific Gene Homologous to Lysosomal Membrane Glycoproteins 1 and 2: Significantly Increased Expression in Cancers of Various Tissues. *Cancer Research*, **58**, 3499-3503.

Pellet-Many C., Frankel P., Jia H., Zachary I. (2008) Neuropilins: structure, function and role in disease. *Biochemical Journal*, **411**, 211-226.

Penning T.M., Burczynski M.E., Jez J.M., Hung C.F., Lin H.K., Ma H., Moore M., Palackal N., Ratnam K. (2000) Human 3α-hydroxysteroid dehydrogenase isoforms (AKR1C1-AKR1C4) of the aldo-keto reductase superfamily: Functional plasticity and tissue distribution reveals roles in the inactivation and formation of male and female sex hormones. *Biochemical Journal*, **351**, 67-77.

Popov V.M., Wang C., Shirley L.A., Rosenberg A., Li S., Nevalainen M., Fu M., Pestell R.G. (2007) The functional significance of nuclear receptor acetylation. *Steroids*, **72**, 221-230.

Pozzatti R., Vogel J., Jay G. (1990) The human T-lymphotropic virus type I tax gene can cooperate with the ras oncogene to induce neoplastic transformation of cells. *Molecular and Cellular Biology*, **10**, 413-417.

Pratt W.B., Toft D.O. (1997) Steroid Receptor Interactions with Heat Shock Protein and Immunophilin Chaperones. *Endocrine Reviews*, **18**, 306-360.

Quayle S.N., Mawji N.R., Wang J., Sadar M.D. (2007) Androgen receptor decoy molecules block the growth of prostate cancer. *Proceedings of the National Academy of Sciences*, **104**, 1331-1336.

Quigley C.A., De Bellis A., Marschke K.B., El-Awady M.K., Wilson E.M., French F.S. (1995) Androgen receptor defects: Historical, clinical, and molecular perspectives. *Endocrine Reviews*, **16**, 271-321.

Ris-Stalpers C., Kuiper G.G., Faber P.W., Schweikert H.U., van Rooij H.C., Zegers N.D., Hodgins M.B., Degenhart H.J., Trapman J., Brinkmann A.O. (1990) Aberrant splicing of androgen receptor mRNA results in synthesis of a nonfunctional receptor protein in a patient with androgen insensitivity. *Proceedings of the National Academy of Sciences*, **87**, 7866-7870.

Ris-Stalpers C., Verleun-Mooijman M.C., de Blaeij T.J., Degenhart H.J., Trapman J., Brinkmann A.O. (1994) Differential splicing of human androgen receptor pre-mRNA in X-linked Reifenstein syndrome, because of a deletion involving a putative branch site. *American Journal of Human Genetics*, **54**, 609-617.

Rizner T.L., Lin H.K., Peehl D.M., Steckelbroeck S., Bauman D.R., Penning T.M. (2003) Human type 3 3α-hydroxysteroid dehydrogenase (aldo-keto reductase 1C2) and androgen metabolism in prostate cells. *Endocrinology*, **144**, 2922-2932.

Rozen S., Skaletsky H.J. (2000) Primer3 on the WWW for general users and for biologist programmers, vol 132. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, New Jersey.

Ruizeveld de Winter J.A., Janssen P.J., Sleddens H.M., Verleun-Mooijman M.C., Trapman J., Brinkmann A.O., Santerse A.B., Schroder F.H., van der Kwast T.H. (1994) Androgen receptor status in localized and locally progressive hormone refractory human prostate cancer. *American Journal of Pathology*, **144**, 735-746.

Sack J.S., Kish K.F., Wang C., Attar R.M., Kiefer S.E., An Y., Wu G.Y., Scheffler J.E., Salvati M.E., Krystek S.R., Weinmann R., Einspahr H.M. (2001) Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proceedings of the National Academy of Sciences*, **98**, 4904-4909.

Sadar M.D. (2011) Small Molecule Inhibitors Targeting the "Achilles' Heel" of Androgen Receptor Activity. *Cancer Research*, **71**, 1208-1213.

Sadar M.D. (2012) Advances in small molecule inhibitors of androgen receptor for the treatment of advanced prostate cancer. *World Journal of Urology*, **30**, 311-318.

Sadi M.V., Walsh P.C., Barrack E.R. (1991) Immunohistochemical study of androgen receptors in metastatic prostate cancer. Comparison of receptor content and response to hormonal therapy. *Cancer*, **67**, 3057-3064.

Sammarco I., Grimaldi P., Rossi P., Cappa M., Moretti C., Frajese G., Geremia R. (2000) Novel Point Mutation in the Splice Donor Site of Exon-Intron Junction 6 of the Androgen Receptor Gene in a Patient with Partial Androgen Insensitivity Syndrome. *Journal of Clinical Endocrinology & Metabolism*, **85**, 3256-3261.

Scher H.I., Beer T.M., Higano C.S., Anand A., Taplin M.-E., Efstathiou E., Rathkopf D., Shelkey J., Yu E.Y., Alumkal J., Hung D., Hirmand M., Seely L., Morris M.J., Danila D.C., Humm J., Larson S., Fleisher M., Sawyers C.L. (2010) Antitumour activity of MDV3100 in castration-resistant prostate cancer: a phase 1-2 study. *Lancet*, **375**, 1437-1446.

Scher H.I., Halabi S., Tannock I., Morris M., Sternberg C.N., Carducci M.A., Eisenberger M.A., Higano C., Bubley G.J., Dreicer R., Petrylak D., Kantoff P., Basch E., Kelly W.K., Figg W.D., Small E.J., Beer T.M., Wilding G., Martin A., Hussain M. (2008) Design and End Points of Clinical Trials for Patients With Progressive Prostate Cancer and Castrate Levels of Testosterone: Recommendations of the Prostate Cancer Clinical Trials Working Group. *Journal of Clinical Oncology*, **26**, 1148-1159.

Shaffer P.L., Jivan A., Dollins D.E., Claessens F., Gewirth D.T. (2004) Structural basis of androgen receptor binding to selective androgen response elements. *Proceedings of the National Academy of Sciences*, **101**, 4758-4763.

Shariat S.F., Lamb D.J., Kattan M.W., Nguyen C., Kim J., Beck J., Wheeler T.M., Slawin K.M. (2002) Association of Preoperative Plasma Levels of Insulin-Like Growth Factor I and Insulin-Like Growth Factor Binding Proteins-2 and -3 With Prostate Cancer Invasion, Progression, and Metastasis. *Journal of Clinical Oncology*, **20**, 833-841.

Siegel R., Naishadham D., Jemal A. (2012) Cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, **62**, 10-29.

Simental J.A., Sar M., Lane M.V., French F.S., Wilson E.M. (1991) Transcriptional activation and nuclear targeting signals of the human androgen receptor. *Journal of Biological Chemistry*, **266**, 510-518.

Singh A.P., Bafna S., Chaudhary K., Venkatraman G., Smith L., Eudy J.D., Johansson S.L., Lin M.-F., Batra S.K. (2008) Genome-wide expression profiling reveals transcriptomic variation and perturbed gene networks in androgen-dependent and androgen-independent prostate cancer cells. *Cancer Letters*, **259**, 28-38.

Sivanandam A., Murthy S., Chinnakannu K., Bai V.U., Kim S.-H., Barrack E.R., Menon M., Reddy G.P.-V. (2011) Calmodulin protects androgen receptor from calpain-mediated breakdown in prostate cancer cells. *Journal of Cellular Physiology*, **226**, 1889-1896.

Smith D.F., Toft D.O. (2008) Minireview: The Intersection of Steroid Receptors with Molecular Chaperones: Observations and Questions. *Molecular Endocrinology*, **22**, 2229-2240.

Smyth G.K. (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology*, **3**, Article 3.

Song C., Wang W., Li M., Liu Y., Zheng D. (2009) Tax1 enhances cancer cell proliferation via Ras–Raf–MEK–ERK signaling pathway. *IUBMB Life*, **61**, 685-692.

Sramkoski R.M., Pretlow T.G., Giaconia J.M., Pretlow T.P., Schwartz S., Sy M.S., Marengo S.R., Rhim J.S., Zhang D., Jacobberger J.W. (1999) A new human prostate carcinoma cell line, 22Rv1. *In Vitro Cellular & Developmental Biology - Animal*, **35**, 403-409.

Stanbrough M., Bubley G.J., Ross K., Golub T.R., Rubin M.A., Penning T.M., Febbo P.G., Balk S.P. (2006) Increased Expression of Genes Converting Adrenal Androgens to Testosterone in Androgen-Independent Prostate Cancer. *Cancer Research*, **66**, 2815-2825.

Sun S., Sprenger C.C.T., Vessella R.L., Haugk K., Soriano K., Mostaghel E.A., Page S.T., Coleman I.M., Nguyen H.M., Sun H., Nelson P.S., Plymate S.R. (2010) Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. *Journal of Clinical Investigation*, **120**, 2715-2730.

Takai Y., Sasaki T., Matozaki T. (2001) Small GTP-Binding Proteins. *Physiological Reviews*, **81**, 153-208.

Taplin M.E., Bubley G.J., Ko Y.J., Small E.J., Upton M., Rajeshkumar B., Balk S.P. (1999) Selection for androgen receptor mutations in prostate cancers treated with androgen antagonist. *Cancer Research*, **59**, 2511-2515.

Taplin M.-E., Bubley G.J., Shuster T.D., Frantz M.E., Spooner A.E., Ogata G.K., Keer H.N., Balk S.P. (1995) Mutation of the Androgen-Receptor Gene in Metastatic Androgen-Independent Prostate Cancer. *New England Journal of Medicine*, **332**, 1393-1398.

Taplin M.-E., Rajeshkumar B., Halabi S., Werner C.P., Woda B.A., Picus J., Stadler W., Hayes D.F., Kantoff P.W., Vogelzang N.J., Small E.J. (2003) Androgen Receptor Mutations in Androgen-Independent Prostate Cancer: Cancer and Leukemia Group B Study 9663. *Journal of Clinical Oncology*, **21**, 2673-2678.

Tepper C.G., Boucher D.L., Ryan P.E., Ma A.-H., Xia L., Lee L.-F., Pretlow T.G., Kung H.-J. (2002) Characterization of a Novel Androgen Receptor Mutation in a Relapsed CWR22 Prostate Cancer Xenograft and Cell Line. *Cancer Research*, **62**, 6606-6614.

Thomas L.N., Douglas R.C., Lazier C.B., Too C.K.L., Rittmaster R.S., Tindall D.J. (2008) Type 1 and Type 2 5α-Reductase Expression in the Development and Progression of Prostate Cancer. *European Urology*, **53**, 244-252.

Tran C., Ouk S., Clegg N.J., Chen Y., Watson P.A., Arora V., Wongvipat J., Smith-Jones P.M., Yoo D., Kwon A., Wasielewska T., Welsbie D., Chen C.D., Higano C.S., Beer T.M., Hung D.T., Scher H.I., Jung M.E., Sawyers C.L. (2009) Development of a Second-Generation Antiandrogen for Treatment of Advanced Prostate Cancer. *Science*, **324**, 787-790.

Vaidyanathan G., Cismowski M.J., Wang G., Vincent T.S., Brown K.D., Lanier S.M. (2004) The Ras-related protein AGS1//RASD1 suppresses cell growth. *Oncogene*, **23**, 5858-5863.

van der Kwast T.H., Schalken J., de Winter J.A.R., van Vroonhoven J.C.C., Mulder E., Boersma W., Trapman J. (1991) Androgen receptors in endocrine-therapy-resistant human prostate cancer. *International Journal of Cancer*, **48**, 189-193.

Visakorpi T., Hyytinen E., Koivisto P., Tanner M., Keinanen R., Palmberg C., Palotie A., Tammela T., Isola J., Kallioniemi O.-P. (1995) In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nature Genetics*, **9**, 401-406.

Wainstein M.A., He F., Robinson D., Kung H.-J., Schwartz S., Giaconia J.M., Edgehouse N.L., Pretlow T.P., Bodner D.R., Kursh E.D., Resnick M.I., Seftel A., Pretlow T.G. (1994) CWR22: Androgen-dependent Xenograft Model Derived from a Primary Human Prostatic Carcinoma. *Cancer Research*, **54**, 6049-6052.

Wang Q., Li W., Zhang Y., Yuan X., Xu K., Yu J., Chen Z., Beroukhim R., Wang H., Lupien M., Wu T., Regan M.M., Meyer C.A., Carroll J.S., Manrai A.K., Janne O.A., Balk S.P., Mehra R., Han B., Chinnaiyan A.M., Rubin M.A., True L., Fiorentino M., Fiore C., Loda M., Kantoff P.W., Liu X.S., Brown M. (2009) Androgen Receptor Regulates a Distinct Transcription Program in Androgen-Independent Prostate Cancer. *Cell*, **138**, 245-256.

Ward R.D., Weigel N.L. (2009) Steroid receptor phosphorylation: Assigning function to site-specific phosphorylation. *BioFactors*, **35**, 528-536.

Watson P.A., Chen Y.F., Balbas M.D., Wongvipat J., Socci N.D., Viale A., Kim K., Sawyers C.L. (2010) Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. *Proceedings of the National Academy of Sciences*, **107**, 16759-16765.

Watson P.A., Ellwood-Yen K., King J.C., Wongvipat J., LeBeau M.M., Sawyers C.L. (2005) Context-Dependent Hormone-Refractory Progression Revealed through Characterization of a Novel Murine Prostate Cancer Cell Line. *Cancer Research*, **65**, 11565-11571.

Wilson C.M., McPhaul M.J. (1994) A and B forms of the androgen receptor are present in human genital skin fibroblasts. *Proceedings of the National Academy of Sciences*, **91**, 1234-1238.

Wilson J.D., Harrod M.J., Goldstein J.L., Hemsell D.L., MacDonald P.C. (1974) Familial Incomplete Male Pseudohermaphroditism, Type 1. *New England Journal of Medicine*, **290**, 1097-1103.

Wu H.-C., Hsieh J.-T., Gleave M.E., Brown N.M., Pathak S., Chung L.W.K. (1994) Derivation of androgen-independent human LNCaP prostatic cancer cell sublines: Role of bone stromal cells. *International Journal of Cancer*, **57**, 406-412.

Wu Z., Irizarry R.A., Gentleman R., Martinez-Murillo F., Spencer F. (2004) A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association*, **99**, 909-917.

Wu Z.-Y., Chen K., Haendler B., McDonald T.V., Bian J.-S. (2008) Stimulation of N-Terminal Truncated Isoform of Androgen Receptor Stabilizes Human Ether-a-go-go-Related Gene-Encoded Potassium Channel Protein via Activation of Extracellular Signal Regulated Kinase 1/2. *Endocrinology*, **149**, 5061-5069.

Xu K., Shimelis H., Linn D.E., Jiang R., Yang X., Sun F., Guo Z., Chen H., Li W., Chen H., Kong X., Melamed J., Fang S., Xiao Z., Veenstra T.D., Qiu Y. (2009) Regulation of Androgen Receptor Transcriptional Activity and Specificity by RNF6-Induced Ubiquitination. *Cancer Cell*, **15**, 270-282.

Yamada P.M., Lee K.-W. (2009) Perspectives in mammalian IGFBP-3 biology: local vs. systemic action. *American Journal of Physiology - Cell Physiology*, **296**, C954-C976.

Yang X., Chen M.W., Terry S., Vacherot F., Bemis D.L., Capodice J., Kitajewski J., de la Taille A., Benson M.C., Guo Y., Buttyan R. (2006) Complex regulation of human androgen receptor expression by Wnt signaling in prostate cancer cells. *Oncogene*, **25**, 3436-3444.

Yang X., Guo Z., Sun F., Li W., Alfano A., Shimelis H., Chen M., Brodie A.M.H., Chen H., Xiao Z., Veenstra T.D., Qiu Y. (2011) Novel Membrane-associated Androgen Receptor Splice Variant Potentiates Proliferative and Survival Responses in Prostate Cancer Cells. *Journal of Biological Chemistry*, **286**, 36152-36160.

Yoshida M. (1993) HTLV-1 Tax: Regulation of gene expression and disease. *Trends in Microbiology*, **1**, 131-135.

Yoshida T., Kinoshita H., Segawa T., Nakamura E., Inoue T., Shimizu Y., Kamoto T., Ogawa O. (2005) Antiandrogen bicalutamide promotes tumor growth in a novel androgen-dependent prostate cancer xenograft model derived from a bicalutamide-treated patient. *Cancer Research*, **65**, 9611-9616.

Zhang S., Zhau H., Osunkoya A., Iqbal S., Yang X., Fan S., Chen Z., Wang R., Marshall F., Chung L., Wu D. (2010) Vascular endothelial growth factor regulates myeloid cell leukemia-1 expression through neuropilin-1-dependent activation of c-MET signaling in human prostate cancer cells. *Molecular Cancer*, **9**, 9.

Zhang X., Morrissey C., Sun S., Ketchandji M., Nelson P.S., True L.D., Vakar-Lopez F., Vessella R.L., Plymate S.R. (2011) Androgen Receptor Variants Occur Frequently in Castration Resistant Prostate Cancer Metastases. *PLoS ONE*, **6**, e27970.

Zhou Z.X., Sar M., Simental J.A., Lane M.V., Wilson E.M. (1994) A ligand-dependent bipartite nuclear targeting signal in the human androgen receptor. Requirement for the DNA-binding domain and modulation by NH2-terminal and carboxyl-terminal sequences. *Journal of Biological Chemistry*, **269**, 13115-13123.

Zoppi S., Wilson C.M., Harbison M.D., Griffin J.E., Wilson J.D., McPhaul M.J., Marcelli M. (1993) Complete testicular feminization caused by an amino-terminal truncation of the androgen receptor with downstream initiation. *Journal of Clinical Investigation*, **91**, 1105-1112.

Appendix I. Supplementary methods

Table I.1 Real-time RT-PCR primer sequences

NRP1	Forward	CAC CAA GGT GAC CAC TGG AA
	Reverse	CAG CAA TCC CAC CAA GGT TT
BNIP3L	Forward	CAC ACC AGC AGG GAC CAT AG
	Reverse	GTG GAA CTC CTT GGG TGG AA
ACPP	Forward	CTT GCC ACT TGA CGG AAT TG
	Reverse	CTG CAG CCA GGT AGC ATG AG
PLA2G10	Forward	TAC CTG TGT GCC TGC CAA TC
	Reverse	ACC CAC AGT TCC TGC CAG TT
NR4A2	Forward	GAC TCC AAC CCG GCT ATG AC
	Reverse	TGG AGC CAG TCA GGA GAT CA
PMEPA1	Forward	TGC AAC TGC AAA CGC TCT TT
	Reverse	CCA CCA CCA TCA CCA TCA TC
SIGMAR1	Forward	ACC ATC ATC TCT GGC ACC TT
	Reverse	CTC CAC CAT CCA TGT GTT TG
AR	Forward	AGC TCA CCA AGC TCC TGG AC
	Reverse	TTG GGC ACT TGC ACA GAG AT
AR-V1	Forward	CCA TCT TGT CGT CTT CGG AAA TGT TAT GAA GC
	Reverse	CTG TTG TGG ATG AGC AGC TGA GAG TCT
AR-V3	Forward	GTG GAA GCT CAA GGT CTT CTT
	Reverse	TTT CTT CAG TCC CAT TGG TG
AR-V4	Forward	CTA CTC CGG ACC TTA CGG GGA CAT GCG
	Reverse	CTT TTA ATT TGT TCA TTC TGA AAA ATC
AR-V7	Forward	CCA TCT TGT CGT CTT CGG AAA TGT TAT GAA GC
	Reverse	TTT GAA TGA GGC AAG TCA GCC TTT CT
SGK	Forward	AGG AGG ATG GGT CTG AAC GA
	Reverse	AGG AGA AGG GTT GGC ATT CA
LEF1	Forward	ACG TGA AGC CTC AGC ATG AA
	Reverse	AGC AAC GAC ATT CGC TCT CA
PALLD	Forward	AAG ATG CTG GTG CGT GAG AA
	Reverse	CAC AAG CTC CAG GCT GAA TG
MKNK2	Forward	TCC TGA GCC ACA TCC ACA AG
	Reverse	GTG GGC GAT GCC TTT GTT AT

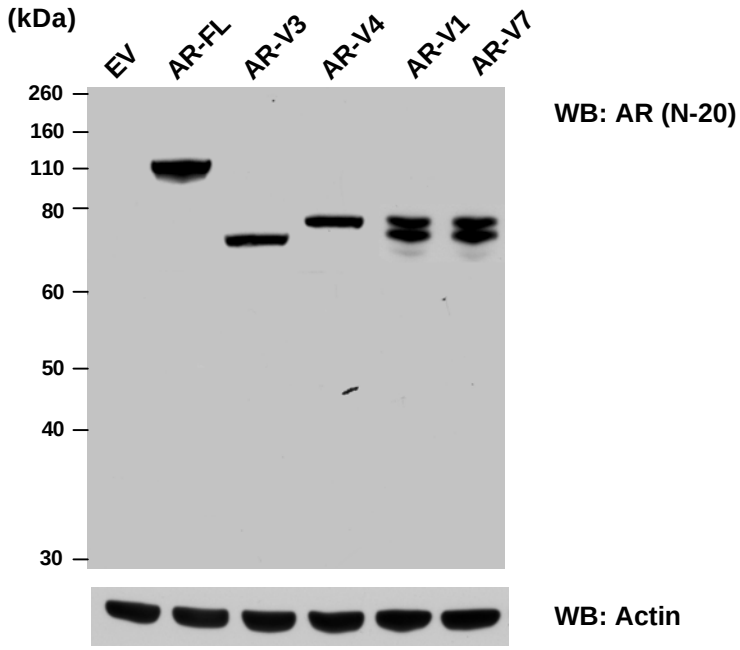
Figure I.1 Complete coding sequence for human AR Exons 1-8

The complete coding sequence for human AR is 2733 nucleotides in length (NCBI Accession: NM_000044.3). Each exon (1-8) is depicted in a different color. The short hairpin target sequences used in this study are also highlighted in exons 1 and 8 respectively.

```
   1 atggaagtgc agttagggct gggaagggtc taccctcggc cgccgtccaa gacctaccga
  61 ggagctttcc agaatctgtt ccagagcgtg cgcgaagtga tccagaaccc gggccccagg
 121 cacccagagg ccgcgagcgc agcacctccc ggcgccagtt tgctgctgct gcagcagcag
 181 cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcaa
 241 gagactagcc ccaggcagca gcagcagcag cagggtgagg atggttctcc ccaagcccat
 301 cgtagaggcc ccacaggcta cctggtcctg gatgaggaac agcaaccttc acagccgcag
 361 tcggccctgg agtgccaccc cgagagaggt tgcgtcccag agcctggagc cgccgtggcc
 421 gccagcaagg ggctgccgca gcagctgcca gcacctccgg acgaggatga ctcagctgcc
 481 ccatccacgt tgtccctgct gggccccact ttccccggct taagcagctg ctccgctgac
 541 cttaaagaca tcctgagcga ggccagcacc atgcaactcc ttcagcaaca gcagcaggaa
 601 gcagtatccg aaggcagcag cagcgggaga gcgagggagg cctcgggggc tcccacttcc
 661 tccaaggaca attacttagg gggcacttcg accatttctg acaacgccaa ggagttgtgt
 721 aaggcagtgt cggtgtccat gggcctgggt gtggaggcgt tggagcatct gagtccaggg
 781 gaacagcttc ggggggattg catgtacgcc ccacttttgg gagttccacc cgctgtgcgt
 841 cccactcctt gtgccccatt ggccgaatgc aaaggttctc tgctagacga cagcgcaggc
 901 aagagcactg aagatactgc tgagtattcc cctttcaagg gaggttacac caaagggcta
 961 gaaggcgaga gcctaggctg ctctggcagc gctgcagcag ggagctccgg gacacttgaa
1021 ctgccgtcta ccctgtctct ctacaagtcc ggagcactgg acgaggcagc tgcgtaccag
1081 agtcgcgact actacaactt tccactggct ctggccggac cgccgccccc tccgccgcct
1141 ccccatcccc acgctcgcat caagctggag aacccgctgg actacggcag cgcctgggcg
1201 gctgcggcgg cgcagtgccg ctatggggac ctggcgagcc tgcatggcgc gggtgcagcg
1261 ggacccggtt ctgggtcacc ctcagccgcc gcttcctcat cctggcacac tctcttcaca
1321 gccgaagaag gccagttgta tggaccgtgt ggtggtggtg ggggtggtgg cggcggcggc
1381 ggcggcggcg gcggcggcgg cggcggcggc ggcggcggcg aggcgggagc tgtagccccc
1441 tacggctaca ctcggccccc tcaggggctg gcgggccagg aaagcgactt caccgcacct
1501 gatgtgtggt accctggcgg catggtgagc agagtgccct atcccagtcc cacttgtgtc
1561 aaaagcgaaa tgggcccctg gatggatagc tactccggac cttacgggga catgcgtttg
1621 gagactgcca gggaccatgt tttgcccatt gactattact ttccacccca gaagacctgc
1681 ctgatctgtg gagatgaagc ttctgggtgt cactatggag ctctcacatg tggaagctgc
1741 aaggtcttct tcaaaagagc cgctgaaggg aaacagaagt acctgtgcgc cagcagaaat
1801 gattgcacta ttgataaatt ccgaaggaaa aattgtccat cttgtcgtct tcggaaatgt
1861 tatgaagcag ggatgactct gggagcccgg aagctgaaga aacttggtaa tctgaaacta
1921 caggaggaag gagaggcttc cagcaccacc agccccactg aggagacaac ccagaagctg
1981 acagtgtcac acattgaagg ctatgaatgt cagcccatct ttctgaatgt cctggaagcc
2041 attgagccag gtgtagtgtg tgctggacac gacaacaacc agcccgactc ctttgcagcc
2101 ttgctctcta gcctcaatga actgggagag agacagcttg tacacgtggt caagtgggcc
2161 aaggccttgc ctggcttccg caacttacac gtggacgacc agatggctgt cattcagtac
2221 tcctggatgg ggctcatggt gtttgccatg ggctggcgat ccttcaccaa tgtcaactcc
2281 aggatgctct acttcgcccc tgatctggtt ttcaatgagt accgcatgca caagtcccgg
2341 atgtacagcc agtgtgtccg aatgaggcac ctctctcaag agtttggatg gctccaaatc
2401 acccccccagg aattcctgtg catgaaagca ctgctactct tcagcattat tccagtggat
2461 gggctgaaaa atcaaaaatt ctttgatgaa cttcgaatga actacatcaa ggaactcgat
2521 cgtatcattg catgcaaaag aaaaaatccc acatcctgct caagacgctt ctaccagctc
2581 accaagctcc tggactccgt gcagcctatt gcgagagagc tgcatcagtt cactttgac
2641 ctgctaatca agtcacacat ggtgagcgtg gactttccgg aaatgatggc agagatcatc
2701 tctgtgcaag tgcccaagat cctttctggg aaagtcaagc ccatctattt ccacacccag
2761 tga
```

Figure II.2 Expression of full-length AR and AR variant plasmids in COS-7 cells

COS-7 cells were transfected with empty vector control (EV), full-length AR (AR-FL), AR-V3, AR-V4, AR-V1 and AR-V7 expression plasmids and harvested at 48 hr. Lysates were subjected to Western immunobloting using a polyclonal antibody targeted to the AR N-terminal domain. The immunoblot was stripped and reproped with an actin antibody to confirm equal protein loading.


(kDa)
EV
AR-FL
AR-V3
AR-V4
AR-V1
AR-V7
260
160
110
80
60
50
40
30
WB: AR (N-20)
WB: Actin

Appendix III. Clinical characteristics of patients

Table III.2 Clinical characteristics of patients

Array ID	Case ID	Age	Race	Pre-Op PSA	Surgery Date	Gleason Pattern	Gleason Score	Stage (pTNM)	GPSM	Date PSA Failure	Date of Last Follow Up	Date of Death
177	536	71	C	41.75	11/28/2001	3	7	T2B,N0+	12		7/10/2007	
178	552	64	C	7	12/19/2001	3	6	T2A,N0+	9		4/19/2007	
179	828	65	C	7.55	9/16/2002	3	7	T2B,N0+	10	11/6/2000	7/8/2003	7/8/2003
180	847	58	C	7.5	10/4/2002	3	7	T2B,N0+	10		8/14/2007	
181	1030	60	C	6.2	5/13/2003	3	7	T2A,N0+	10	4/11/2002	4/20/2007	
182	1036	53	C	1.2	5/28/2003	3	6	T2A,N0-	6		7/30/2007	
183	1041	66	C	3.97	6/4/2003	3	6	T2B,N0-	6		10/2/2007	
184	166	65	C	62	7/17/2000	5	9	T3B,N0+	16		4/22/2007	
185	468	57	C	4.9	3/27/2002	5	6	T2B,N0-	7		2/25/2008	
187	802	71	C	8.2	8/13/2002	5	9	T3B,N0-	12		4/5/2006	
188	960	70	C	9.1	2/5/2003	5	9	T3B,N0-	12		2/26/2007	
189	1179	69	C	4.2	10/17/2003	5	10	T3B,N0+	15		2/11/2003	2/11/2003
190	1190	59	C	8.1	11/4/2003	5	10	T3B,N0+	15		1/14/2008	
191	1593	74	C	49.4	9/9/2004	5	9	T3B,N1+	16		12/29/2007	
274	17	74	C	7.5	11/24/1999	3	6	T2B,N0-	7	6/22/2004	12/17/2007	
275	38	65	C	8.7	1/18/2000	3	6	T2B,N0-	7		7/30/2005	
276	258	70	C	10.8	12/19/2000	3	6	T2B,N0+	10		4/3/2008	
277	302	49	C	4.6	3/13/2001	3	6	T2B,N0+	9	11/20/2003	10/8/2007	
278	510	63	C	7.6	10/30/2001	3	6	T2B,N0-	7		10/8/2002	
279	776	59	U	5.5	7/17/2002	3	6	T2A,N0+	9		11/15/2007	
281	348	58	C	5.9	5/9/2001	4	6	T2B,N0+	9	7/27/2006	4/11/2007	
282	702	62	A	40.8	5/16/2002	4	8	T3B,N0+	15		2/20/2007	
283	730	67	C	15.4	6/11/2002	4	7	T3A,N0-	9		12/12/2007	
284	795	58	U	4.7	8/7/2002	4	8	T3A,N0-	11		5/8/2006	5/8/2006
285	1269	71	C	31.7	12/19/2003	4	9	T3A,N0+	14		10/26/2007	
286	1330	48	C	36.1	3/9/2004	4	8	T3B,N1+	15	11/9/2006	3/25/2008	

287	1447	59	C	6	5/26/2004	5	9	T3A,N0+	12		10/10/2007	
288	1476	68	C	20	6/22/2004	5	8	T3A,N0-	10		11/30/2007	
309	816	54	A	4.9	8/27/2002	3	6	T2B,N0+	9		2/20/2007	
310	840	51	C	26.1	9/27/2002	3	6	T2A,N0-	9		3/9/2007	
311	881	56	U	6.33	11/13/2002	3	7	T2B,N0-	8	12/15/2003	6/4/2007	
313	1060	50	C	4.2	6/18/2003	3	6	T2A,N0-	7	12/15/2003	6/4/2007	
332	1026	71	C	1.5	5/8/2003	4	7	T2A,N0-	7		6/20/2007	
333	1060	50	C	4.2	6/18/2003	3	6	T2A,N0-	7		10/12/2007	
334	1095	65	C	7.9	7/18/2003	4	9	T3B,N0+	12		1/15/2007	
338	1374	57	U	9.8	4/7/2004	4	7	T2B,N0-	8		1/10/2008	
349	1111	63	C	5	8/12/2003	3	6	T2B,N0-	7		7/13/2007	
350	1243	69	C	8	1/8/2004	4	7	T2A,N0-	8		7/13/2007	
351	1493	68	C	18.2	7/1/2004	4	7	T3B,N0-	11		10/24/2007	
406	1125	68	C	6.1	8/22/2003	3	6	T2B,N0-	7		7/12/2007	
407	1131	48	C	2.8	8/28/2003	3	7	T2a,N0-	7		8/2/2007	
408	1171	52	C	4.1	10/9/2003	3	6	T2B,N0+	9		6/10/2007	6/10/2007
410	1298	56	C	3.3	2/3/2004	4	7	T2b,N0-	7		1/15/2007	
412	1191	57	C	11.4	11/4/2003	3	6	T2b,N0+	10	4/12/2005	6/12/2007	
416	1291	57	C	7	1/27/2004	3	7	T2b,N0-	8	5/11/2006	1/24/2008	
418	1321	61	C	4.6	2/26/2004	4	7	T3b,N0+	12		1/21/2008	
419	1523	65	C	10	7/23/2004	4	6	T2b,N0-	7	12/16/2005	8/8/2007	
436	1199	60	C	16.3	11/6/2003	4	7	T3b,N0-	11		10/5/2007	
439	877	49	C	8.4	11/12/2002	3	6	T2a,N0-	7		8/16/2007	
440	832	72	C	11.4	9/20/2002	3	7	T2b,N0-	9	6/29/2004	1/23/2006	1/23/2006
441	843	64	C	17.2	10/2/2002	4	9	T3b,N0-	15		11/26/2007	
443	957	54	C	4.1	2/3/2003	3	6	T2b,N0+	9	9/14/2004	1/7/2008	
444	1067	66	U	3.5	6/20/2003	3	6	T2b,N0-	6	4/15/2005	12/24/2007	
445	1096	47	U	4.22	7/23/2003	3	6	T2b,N0-	7		7/12/2007	
446	1185	74	C	7.9	10/27/2003	4	7	T3b,N0-	12	6/22/2006	9/21/2007	
449	831	62	C	7	9/18/2002	4	7	T2b,N0-	8		7/2/2007	

C caucausian, *A* African American, *U* Unknown, *PSA* prostate specific antigen, *TNM* tumor node metastasis, *GPSM* Gleason, PSA, seminal vesicle and margin status

Table III.3 Clinical characteristics for patients which follow-up is available

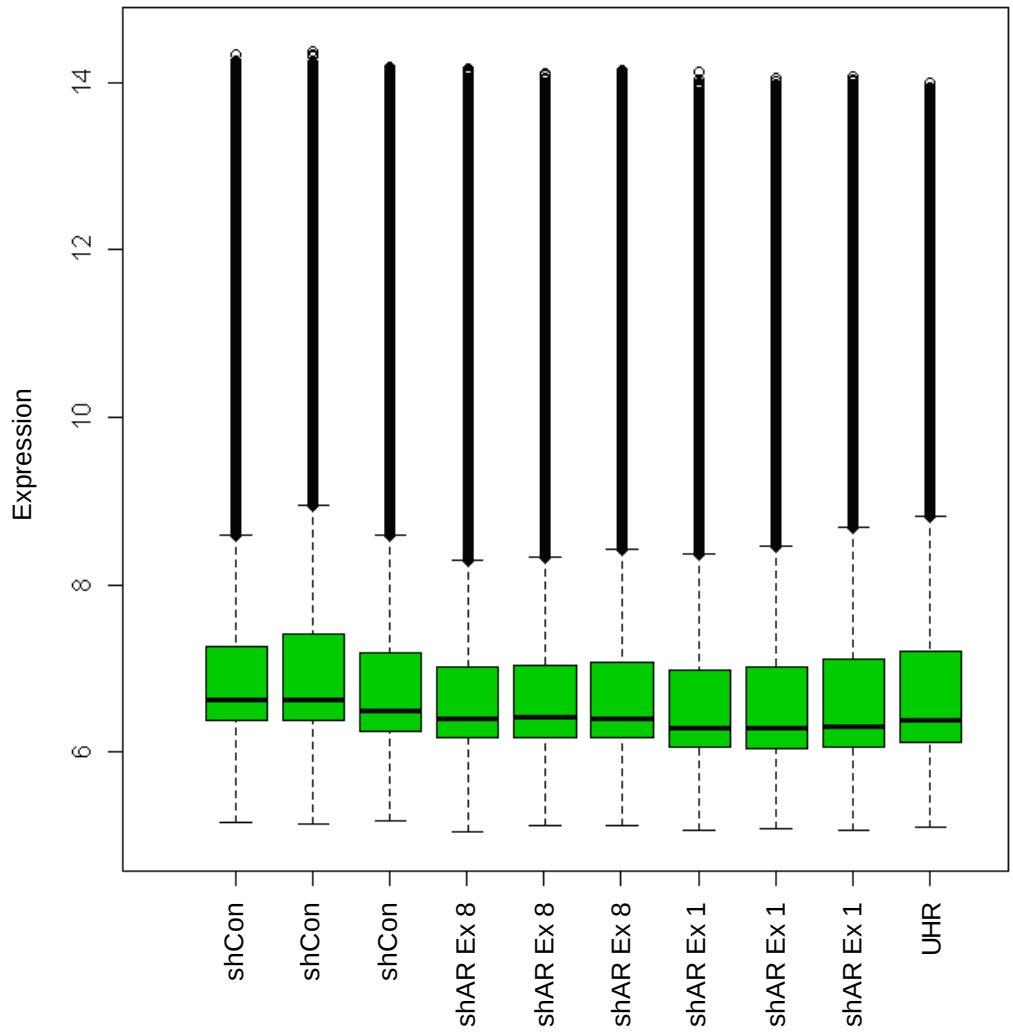
	No PSA Fail (n=41)	PSA Fail (n=15)	Total (n=56)	p value
Age				0.6701
Mean (SD)	61.8 (7.56)	60.5 (8.05)	61.4 (7.64)	
Median	62.0	62.0	62.0	
Q1, Q3	57.0, 68.0	50.0, 68.0	56.5, 68.0	
Range	47.0-74.0	48.0-71.0	47.0-74.0	
Fail or follow-up (months)				< 0.0001
Mean (SD)	51.6 (17.49)	17.6 (13.88)	42.5 (22.42)	
Median	49.1	14.1	45.2	
Q1, Q3	43.6, 61.6	5.9, 23.0	25.5, 55.7	
Range	0.6-91.5	3.6-45.8	0.6-91.5	
Gleason score				0.0710
6	20 (48.8%)	4 (26.7%)	24 (42.9%)	
7	14 (34.1%)	4 (26.7%)	18 (32.1%)	
8-10	7 (17.1%)	7 (46.7%)	14 (25%)	
Gleason Pattern				0.0353
GP3	25 (61%)	4 (26.7%)	29 (51.8%)	
GP4	12 (29.3%)	6 (40%)	18 (32.1%)	
GP5	4 (9.8%)	5 (33.3%)	9 (16.1%)	
Stage (pTNM)				0.0636
T2	30 (73.2%)	7 (46.7%)	37 (66.1%)	
T3	11 (26.8%)	8 (53.3%)	19 (33.9%)	

Figure IV.3 Analysis of 22Rv1 microarray data (pre- and post-normalization)

A. Box plot of raw 22Rv1 microarray expression data from each sample, before data normalization was performed. UHR is the Universal Human Reference.

B. Unsupervised clustering of the microarray expression data following removal of the UHR sample, quantile normalization and noise filtering.

A. Relative gene expression (pre-normalization)



B. Unsupervised clustering (post-normalization)

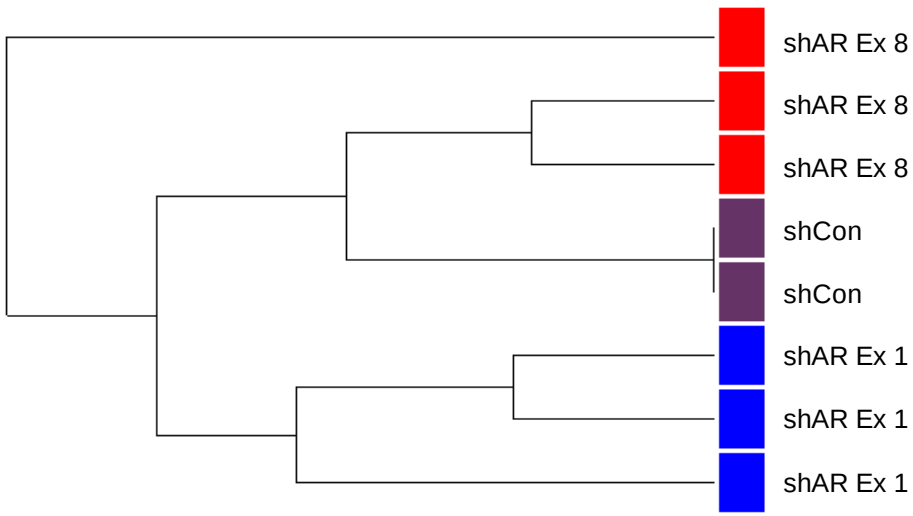
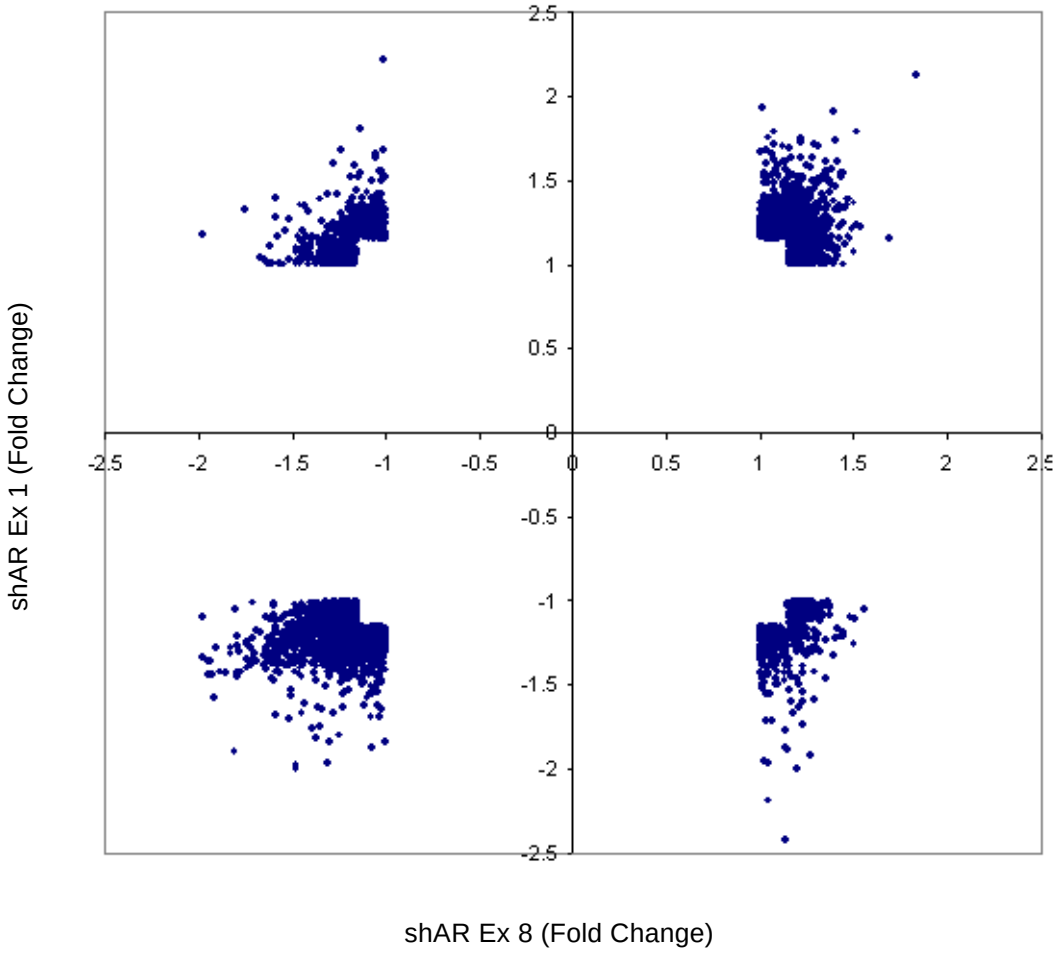


Figure IV.4 Volcano plot of all differentially expressed probes in the 22Rv1 microarray

A volcano plot of all differentially expressed probes (n=4,817) between shAR Ex 1 treatment (y-axis) and shAR Ex 8 treatment (x-axis) identified by microarray in 22Rv1 cells (fold change >1.0; $p < 0.05$). Both axes represent data as Log_2 fold change.


shAR Ex 1 (Fold Change)
shAR Ex 8 (Fold Change)
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2

Appendix IV. Supplementary microarray data

Table IV.4 Full-length AR regulated genes in 22Rv1 cells

Probe Set ID	Target ID	p value	Fold Change
ILMN_1779448	EFHD1	0.00043	1.68848
ILMN_3242315	SNORD3D	0.01609	1.49995
ILMN_1746972	LOC653321	0.22277	1.48095
ILMN_1748124	TSC22D3	0.00990	1.47251
ILMN_1810560	P8	0.01650	1.44967
ILMN_1659273	LOC441408	0.01840	1.43723
ILMN_1672908	TWIST1	0.01453	1.41838
ILMN_1729234	TPP1	0.01726	1.41498
ILMN_2311537	HMGA1	0.01900	1.41418
ILMN_1724194	NPEPL1	0.01916	1.40550
ILMN_1791593	DENND5B	0.07035	1.40243
ILMN_1663444	LIN7B	0.01602	1.39358
ILMN_1691291	PIGS	0.02035	1.39044
ILMN_1665235	CRTAP	0.02497	1.38230
ILMN_1789614	TPT1	0.36997	1.37443
ILMN_2404688	NUPR1	0.02988	1.36785
ILMN_1654468	ACRV1	0.03211	1.36453
ILMN_2256953	CASZ1	0.05959	1.36379
ILMN_3237755	SSPO	0.09182	1.36145
ILMN_1658902	DEAF1	0.04324	1.35742
ILMN_3237665	COX7A2L	0.10634	1.35678
ILMN_1814333	SERPINI1	0.05838	1.35659
ILMN_2210581	B3GAT3	0.04023	1.35546
ILMN_1683277	KIAA0319L	0.04699	1.35476
ILMN_1815306	AP2A1	0.07153	1.35444
ILMN_1754727	GPRASP2	0.04273	1.35254
ILMN_1775736	NPM2	0.05166	1.35175
ILMN_1693630	C16ORF7	0.03133	1.35025
ILMN_1755727	KDM5B	0.06771	1.34992
ILMN_1698213	RBM3	0.13519	1.34879
ILMN_2228873	STARD3NL	0.05619	1.34722
ILMN_1748899	TMED6	0.04499	1.34587
ILMN_1653730	OXCT2	0.05396	1.34372
ILMN_1660864	RHBDL1	0.07681	1.34288
ILMN_1777453	LOC653419	0.03962	1.34077
ILMN_2252408	CNPY4	0.04859	1.33831
ILMN_1764780	SVOP	0.11277	1.33099
ILMN_1713008	ABCC11	0.04412	1.33020
ILMN_2074860	RN7SK	0.06943	1.32962
ILMN_1736481	SECISBP2	0.03843	1.32569
ILMN_1740960	MACROD1	0.05730	1.32351
ILMN_1654772	LOC651511	0.05253	1.32247
ILMN_1666904	SLC17A6	0.04670	1.32206
ILMN_1768575	SFTPD	0.04341	1.32057
ILMN_3234762	RN5S9	0.08875	1.31908
ILMN_1690826	TNKS1BP1	0.15707	1.31867
ILMN_1658758	NPB	0.07059	1.31781
ILMN_1656477	ARSA	0.10634	1.31759

ILMN_1747183	GXYLT1	0.40257	1.31703
ILMN_2389273	FXR1	0.18476	1.31547
ILMN_1784630	KBTBD11	0.07448	1.31505
ILMN_1746403	GTF2IRD2P	0.08322	1.31436
ILMN_1718063	LIPA	0.10892	1.31432
ILMN_1811181	FLJ20444	0.09027	1.31410
ILMN_1713491	VAMP2	0.27263	1.31395
ILMN_1687495	SLC37A1	0.08190	1.31344
ILMN_3246869	SCARNA21	0.07991	1.31248
ILMN_1760109	SLAMF6	0.16243	1.31116
ILMN_1673509	RPL28	0.05227	1.30951
ILMN_1656042	KIAA0319L	0.13434	1.30886
ILMN_1701483	SYP	0.08442	1.30692
ILMN_2185866	PCDHB17	0.12216	1.30532
ILMN_1752526	RNF144B	0.10283	1.30366
ILMN_2407703	SYN1	0.22277	1.30245
ILMN_1799744	GALC	0.13298	1.30184
ILMN_1741131	CHRNB1	0.15098	1.30119
ILMN_1682828	FGL1	0.10453	1.30099
ILMN_3188984	C20ORF199	0.18480	1.30032
ILMN_1655642	FANCI	0.00269	-1.58560
ILMN_1784207	C1ORF128	0.00102	-1.97607
ILMN_1675797	EPDR1	0.00021	-1.79800
ILMN_1717173	ECT2	0.00986	-1.78944
ILMN_2234697	BEX1	0.00021	-1.71265
ILMN_2219712	HMGB2	0.00269	-1.69036
ILMN_1690170	CRABP2	0.00047	-1.63574
ILMN_3288717	LOC392437	0.00329	-1.61600
ILMN_2128741	RDH11	0.00219	-1.60095
ILMN_1698323	PLEKHB2	0.00269	-1.59832
ILMN_3265797	LOC100130561	0.04624	-1.59753
ILMN_1810797	WASF3	0.01198	-1.59587
ILMN_1659027	SLC2A1	0.01693	-1.59381
ILMN_1815184	ASPM	0.05975	-1.59239
ILMN_1657495	MLEC	0.00190	-1.55645
ILMN_2297710	PLEKHB2	0.00197	-1.53628
ILMN_2285996	KIAA0101	0.00489	-1.52929
ILMN_2103685	DEPDC1B	0.00581	-1.52504
ILMN_1773200	CP110	0.12287	-1.51163
ILMN_2409298	NUSAP1	0.00897	-1.50612
ILMN_1681737	TMSB15A	0.00416	-1.49985
ILMN_2067408	CLRN3	0.01170	-1.49916
ILMN_2054392	PPIL1	0.01353	-1.48733
ILMN_1671843	PSRC1	0.01882	-1.48557
ILMN_1739645	ANLN	0.02390	-1.47992
ILMN_1785198	POLE3	0.01290	-1.47704
ILMN_1659364	RFC5	0.00329	-1.47466
ILMN_1698934	CMTM7	0.01171	-1.46782
ILMN_2081988	LANCL1	0.03836	-1.46696
ILMN_1768719	RDH11	0.00791	-1.46484
ILMN_1735093	TIMELESS	0.00581	-1.46397
ILMN_1771697	VRK3	0.00796	-1.46136
ILMN_1749829	DLGAP5	0.04624	-1.46023
ILMN_1703906	HJURP	0.00339	-1.46008

ILMN_2344971	FOXM1	0.01412	-1.46001
ILMN_1654268	HMGB2	0.04104	-1.45648
ILMN_1815924	NUP107	0.02382	-1.45203
ILMN_1768816	TMPO	0.01019	-1.45038
ILMN_2193980	ABCB6	0.01315	-1.44237
ILMN_1669842	CHAF1A	0.01167	-1.44132
ILMN_2168449	DHX15	0.02855	-1.43847
ILMN_1889752	HS.21177	0.01827	-1.43276
ILMN_3306440	TMEM194A	0.02521	-1.42877
ILMN_1713752	SERINC3	0.07198	-1.42729
ILMN_3240003	LOC100133012	0.02935	-1.42695
ILMN_1663195	MCM7	0.00851	-1.42595
ILMN_1658143	RFC3	0.01726	-1.42506
ILMN_1739222	ETV5	0.02373	-1.41605
ILMN_2135984	MASTL	0.01895	-1.41529
ILMN_1735680	TMEM30A	0.00851	-1.41154
ILMN_2143155	KIF11	0.19620	-1.40947
ILMN_1651237	CDT1	0.01871	-1.40822
ILMN_1726388	ACBD7	0.05957	-1.40794
ILMN_1677239	CCDC14	0.14315	-1.40624
ILMN_1728972	FAM64A	0.01811	-1.40536
ILMN_1662932	LCP1	0.02803	-1.40457
ILMN_1665559	CDK2	0.01768	-1.40401
ILMN_2184708	LIN7C	0.03509	-1.40397
ILMN_1798210	E2F7	0.01100	-1.40257
ILMN_1729019	SEPT7	0.11614	-1.40222
ILMN_2362549	ZWINT	0.02661	-1.40196
ILMN_2148012	EPR1	0.02044	-1.40160
ILMN_1711470	UBE2T	0.05740	-1.39525
ILMN_1784860	RFC3	0.03090	-1.39480
ILMN_1666208	C14ORF106	0.20349	-1.39329
ILMN_2183938	LEMD3	0.04324	-1.39215
ILMN_1755834	FEN1	0.02415	-1.39098
ILMN_1765644	COMMD8	0.01833	-1.39041
ILMN_2357577	PRKAA1	0.02661	-1.38920
ILMN_1794539	KIF11	0.21681	-1.38885
ILMN_1759915	ARPC1A	0.01768	-1.38803
ILMN_1698307	DBNL	0.01518	-1.38692
ILMN_2126706	LMNB1	0.03531	-1.38674
ILMN_2389114	FIGNL1	0.03275	-1.38651
ILMN_1661776	CENPJ	0.05906	-1.38494
ILMN_1760201	DNMT1	0.01966	-1.38418
ILMN_1799516	DNAJC9	0.02988	-1.38381
ILMN_1698733	CNIH2	0.07375	-1.38285
ILMN_1696046	SIVA	0.06939	-1.38118
ILMN_1668099	CFC1	0.02657	-1.37811
ILMN_1703279	CXORF57	0.01840	-1.37799
ILMN_1753393	OSGEP	0.04960	-1.37799
ILMN_1789123	PLK4	0.12190	-1.37784
ILMN_1657796	STMN1	0.01895	-1.37764
ILMN_2411190	SMC2	0.16814	-1.37510
ILMN_1811472	KIF23	0.08732	-1.37509
ILMN_1684114	LOC286016	0.03290	-1.37501
ILMN_2370692	ANKZF1	0.02116	-1.37408

ILMN_1710428	CDC2	0.03836	-1.37245
ILMN_2054145	PAK1IP1	0.11134	-1.36983
ILMN_2162860	SLFN11	0.01916	-1.36957
ILMN_1678300	MGC40489	0.02276	-1.36946
ILMN_1724811	PARN	0.08330	-1.36907
ILMN_2139816	GPSM2	0.04780	-1.36722
ILMN_1697652	PLEKHB2	0.06895	-1.36710
ILMN_1704537	PHGDH	0.04751	-1.36644
ILMN_1651966	FGFR3	0.04395	-1.36446
ILMN_1802615	CDK6	0.09671	-1.36406
ILMN_1674380	TRPC1	0.05253	-1.36190
ILMN_1740429	FTL	0.05622	-1.36065
ILMN_1669851	STAG3L4	0.04630	-1.36037
ILMN_2047885	PCDHB9	0.11070	-1.35987
ILMN_1790100	C11ORF82	0.11942	-1.35975
ILMN_1725260	CDC25C	0.04237	-1.35936
ILMN_1796923	LOC81691	0.09130	-1.35587
ILMN_2073307	IL10	0.12137	-1.35554
ILMN_1763359	PEG10	0.19025	-1.35449
ILMN_1776577	DSCC1	0.03957	-1.35259
ILMN_1765701	LOC399942	0.16243	-1.35213
ILMN_3236021	LOC100133923	0.25981	-1.35192
ILMN_1700515	C17ORF58	0.04876	-1.34916
ILMN_1652580	POLD1	0.02432	-1.34844
ILMN_2363392	TNFSF14	0.07020	-1.34805
ILMN_1765829	NUFIP2	0.25178	-1.34731
ILMN_1721868	KPNA2	0.04997	-1.34703
ILMN_1695386	RAD51C	0.03126	-1.34654
ILMN_2169089	C18ORF54	0.07275	-1.34604
ILMN_3220769	LOC729964	0.15489	-1.34512
ILMN_1788701	PSIP1	0.13476	-1.34454
ILMN_1732688	DUT	0.05107	-1.34382
ILMN_1802819	DEPDC1	0.15045	-1.34332
ILMN_2338323	CDC25B	0.06259	-1.34324
ILMN_2374633	ZWILCH	0.13318	-1.34271
ILMN_2210713	COMMD8	0.12533	-1.34128
ILMN_1669479	MEST	0.10961	-1.34019
ILMN_2120273	AP1S2	0.07538	-1.33903
ILMN_1754272	GINS3	0.04946	-1.33876
ILMN_1797307	BUB1B	0.04276	-1.33832
ILMN_1814120	PECR	0.04899	-1.33830
ILMN_1767542	THAP10	0.03431	-1.33824
ILMN_1797813	SUZ12	0.15309	-1.33804
ILMN_1778561	WEE1	0.06503	-1.33803
ILMN_2413330	TMEM107	0.09284	-1.33688
ILMN_1746784	SLAIN1	0.05730	-1.33512
ILMN_1742779	CENPL	0.04993	-1.33342
ILMN_1651828	CCT3	0.05801	-1.33319
ILMN_2174380	KIAA1804	0.06895	-1.33137
ILMN_1804248	FDPS	0.06762	-1.33089
ILMN_1678423	SPA17	0.06457	-1.33028
ILMN_2403247	CMTM7	0.11729	-1.33005
ILMN_2151488	RMI1	0.20790	-1.32910
ILMN_2202423	HELLS	0.14110	-1.32905

ILMN_1801109	NARG1L	0.11131	-1.32902
ILMN_3211935	LOC100132715	0.04899	-1.32774
ILMN_2043452	FANCE	0.05622	-1.32584
ILMN_1738263	PIGU	0.11614	-1.32465
ILMN_2337789	MARCH2	0.04952	-1.32346
ILMN_1810474	UBE2I	0.08777	-1.32333
ILMN_2257432	RAD51	0.08322	-1.32151
ILMN_1720114	GMNN	0.07933	-1.32135
ILMN_2101375	CCDC77	0.06123	-1.32093
ILMN_1747630	DEK	0.31144	-1.32075
ILMN_1805535	VRK3	0.04662	-1.32059
ILMN_1711904	MXD3	0.09288	-1.32032
ILMN_1718136	UQCRHL	0.04662	-1.31918
ILMN_2386100	BUB3	0.06762	-1.31859
ILMN_2075436	PDIA3P	0.10962	-1.31847
ILMN_1725244	HAT1	0.15508	-1.31835
ILMN_1751393	ZNF684	0.13773	-1.31646
ILMN_2097410	DAPP1	0.13064	-1.31561
ILMN_1755862	PFAS	0.07466	-1.31378
ILMN_1778152	FIGNL1	0.21020	-1.31342
ILMN_2190051	CCDC91	0.16445	-1.31334
ILMN_1805828	VRK1	0.18995	-1.31207
ILMN_1684982	PDK4	0.16445	-1.31147
ILMN_2374778	DUT	0.09249	-1.31118
ILMN_1752520	SLFN11	0.09542	-1.31105
ILMN_1659952	MTMR2	0.15390	-1.31045
ILMN_1655622	PRKRIR	0.06712	-1.30993
ILMN_1683441	NCAPD3	0.05957	-1.30966
ILMN_1782488	RNASEH2B	0.12103	-1.30957
ILMN_2214278	ANKRD32	0.15336	-1.30913
ILMN_2113362	ARL6IP1	0.05622	-1.30896
ILMN_2225718	CENPE	0.15826	-1.30800
ILMN_2396948	PSMC3IP	0.22207	-1.30681
ILMN_1706502	EIF2AK2	0.23384	-1.30664
ILMN_1675448	ZFP36L1	0.17033	-1.30590
ILMN_2355738	INCENP	0.20947	-1.30491
ILMN_2412384	CCNE2	0.25156	-1.30406
ILMN_2094061	IMPA2	0.06231	-1.30328
ILMN_1701882	LOC653820	0.11891	-1.30279
ILMN_2184640	NOLC1	0.12190	-1.30187
ILMN_1809285	DCP1A	0.07761	-1.30163
ILMN_1749213	SDF2L1	0.15333	-1.30154
ILMN_1672728	KCTD5	0.09208	-1.30131
ILMN_2096116	HSP90B1	0.17695	-1.30101

Table IV.5 AR variant regulated genes in 22Rv1 cells

Probe Set ID	Target ID	p value	Fold Change
ILMN_1691410	BAMBI	0.00005	1.93859
ILMN_1699574	NRP1	0.00031	1.78735
ILMN_1791912	SIDT2	0.00010	1.75474
ILMN_1723962	LXN	0.00027	1.70616
ILMN_1714108	TP53INP1	0.00358	1.69942
ILMN_1663042	SDC4	0.00052	1.68601
ILMN_3201937	LOC645381	0.00025	1.67983
ILMN_2227817	GSTA3	0.00118	1.66250
ILMN_1704376	GLDN	0.00057	1.65233
ILMN_1702487	SGK	0.00071	1.62452
ILMN_1813704	KIAA1199	0.00150	1.61387
ILMN_1781745	C9ORF152	0.00071	1.60586
ILMN_1762561	PLA2G10	0.00126	1.59862
ILMN_1680652	SELENBP1	0.00139	1.58974
ILMN_1814015	CDH17	0.00405	1.56551
ILMN_2339955	NR4A2	0.00198	1.56339
ILMN_1677607	SC5DL	0.00716	1.55553
ILMN_1758323	ACPP	0.00171	1.54924
ILMN_1699206	FHDC1	0.00177	1.54411
ILMN_1765076	APPL2	0.01007	1.53293
ILMN_1784661	TMEM2	0.00152	1.53137
ILMN_1788961	PPP2R2A	0.00566	1.52855
ILMN_3236656	LOC286367	0.00398	1.52205
ILMN_1767475	CERK	0.00202	1.51810
ILMN_2053103	SLC40A1	0.00209	1.51412
ILMN_1718961	BNIP3L	0.01579	1.51270
ILMN_2396875	IGFBP3	0.00341	1.49448
ILMN_1802096	ABTB1	0.00222	1.49076
ILMN_1681890	DYNLT3	0.00901	1.48977
ILMN_1655469	TSPAN3	0.02231	1.48971
ILMN_1703852	EFNB2	0.00839	1.48955
ILMN_1790549	TSPAN3	0.00271	1.48777
ILMN_1889229	HS.28456	0.01380	1.46847
ILMN_2213136	LEF1	0.01194	1.46289
ILMN_1676358	RALB	0.00961	1.44412
ILMN_1759460	TAF7	0.02475	1.43234
ILMN_1719622	RABEP1	0.03248	1.42944
ILMN_1668619	KIAA1467	0.00792	1.42806
ILMN_2362245	HNRNPH2	0.04530	1.42560
ILMN_1748983	RTN4	0.06256	1.42288
ILMN_2367165	ABTB1	0.01732	1.41755
ILMN_1693401	KLHL28	0.25284	1.41600
ILMN_1813399	ATP2B1	0.01210	1.41454
ILMN_1719695	NFKBIZ	0.01330	1.41127
ILMN_1697420	TINF2	0.10596	1.41077
ILMN_2352036	RTN4	0.05393	1.40922
ILMN_1675612	BLCAP	0.06005	1.40890
ILMN_1734190	TCEAL3	0.01034	1.40767
ILMN_1744006	GFOD2	0.03485	1.40754
ILMN_1772692	DICER1	0.03686	1.40628
ILMN_1782305	NR4A2	0.01988	1.40347

ILMN_1655614	DSP	0.01604	1.40269
ILMN_1746085	IGFBP3	0.05959	1.40095
ILMN_1683044	PPP1R2	0.01572	1.40053
ILMN_1779416	SCUBE2	0.01034	1.39936
ILMN_1752582	RAB5B	0.01101	1.39844
ILMN_1869897	HS.24119	0.16301	1.39770
ILMN_2334760	ARMCX3	0.09791	1.39765
ILMN_1681984	GALNT10	0.03076	1.39661
ILMN_2374770	TAX1BP1	0.09744	1.39528
ILMN_1674817	C1ORF115	0.02444	1.39495
ILMN_2340027	TSPAN4	0.04090	1.38913
ILMN_1752968	LAMB2	0.03924	1.38906
ILMN_1720422	G3BP2	0.07974	1.38883
ILMN_1737163	SH3BGRL3	0.01972	1.38797
ILMN_1794190	CCPG1	0.03088	1.38606
ILMN_1837428	HS.25318	0.33113	1.38600
ILMN_1695606	EFNB3	0.04861	1.38193
ILMN_1752927	KIAA1600	0.02355	1.38162
ILMN_1753342	SAT1	0.07438	1.37971
ILMN_1779228	CDH2	0.06791	1.37889
ILMN_1705107	SDCBP2	0.04902	1.37819
ILMN_1750409	RAB9A	0.02167	1.37798
ILMN_2048607	ANKRD9	0.03796	1.37484
ILMN_2313901	PAM	0.03190	1.37107
ILMN_1737157	GRAMD1A	0.02709	1.37103
ILMN_1697597	KIAA0494	0.04140	1.37098
ILMN_1695423	CD9	0.05399	1.37033
ILMN_1722811	CDKN1B	0.03796	1.36826
ILMN_1757406	HIST1H1C	0.04879	1.36816
ILMN_2340565	ATP2C1	0.04974	1.36809
ILMN_2312906	MIA3	0.08701	1.36801
ILMN_2381758	G3BP2	0.09151	1.36645
ILMN_1770031	ABHD10	0.04972	1.36239
ILMN_2112128	MAPK4	0.04801	1.36207
ILMN_3263694	LOC100128771	0.04149	1.36205
ILMN_2354547	TUSC3	0.08001	1.36076
ILMN_2333687	CD59	0.09151	1.35953
ILMN_2078975	GRM3	0.03143	1.35891
ILMN_2094905	COMMD10	0.10115	1.35684
ILMN_2228732	CCNG2	0.03173	1.35645
ILMN_1673998	SSTR1	0.10755	1.35372
ILMN_1730229	CGNL1	0.03143	1.35307
ILMN_1672660	MBP	0.04484	1.35253
ILMN_1664608	INPP5A	0.04177	1.35045
ILMN_1679185	LEF1	0.08745	1.34908
ILMN_1713764	LOC440928	0.04017	1.34880
ILMN_2322842	PPHLN1	0.13763	1.34878
ILMN_1718924	ETFA	0.05597	1.34812
ILMN_1758311	NET1	0.02885	1.34555
ILMN_1705116	C6ORF85	0.04388	1.34493
ILMN_1807972	MICAL1	0.06573	1.34350
ILMN_1792072	FUT4	0.08610	1.34334
ILMN_1740426	RASD1	0.06506	1.34246
ILMN_1712312	RAB11A	0.06928	1.34168

ILMN_3214389	LOC100133583	0.06270	1.34160
ILMN_2168747	GSTA2	0.03885	1.33555
ILMN_2089329	SPRY2	0.04974	1.33528
ILMN_1753164	IPO8	0.03895	1.33527
ILMN_2311020	DNAJC12	0.05845	1.33440
ILMN_3244029	ANKRD52	0.07827	1.33438
ILMN_1694011	WNT2	0.05860	1.33357
ILMN_1768031	DEDD2	0.06564	1.33175
ILMN_1749478	TCEAL3	0.09470	1.33143
ILMN_1712431	FAM113B	0.04974	1.33126
ILMN_1746928	TMPRSS11D	0.10697	1.33105
ILMN_1740466	FAM46A	0.05329	1.32664
ILMN_1879480	HS.13291	0.04986	1.32656
ILMN_1750158	ACOX1	0.09949	1.32587
ILMN_1683023	PDGFC	0.05329	1.32555
ILMN_1665696	EFNA4	0.12816	1.32532
ILMN_1694810	PANX2	0.07277	1.32432
ILMN_1682775	EDN1	0.04767	1.32429
ILMN_1673950	STBD1	0.08561	1.32340
ILMN_2325763	VCAM1	0.06573	1.32224
ILMN_1812262	DDR1	0.07350	1.32046
ILMN_1776154	COG3	0.06101	1.32034
ILMN_1666924	PINK1	0.06433	1.31802
ILMN_2100689	MAP2K4	0.06095	1.31793
ILMN_1781819	PAPSS1	0.08879	1.31681
ILMN_3233930	LOC390557	0.05905	1.31647
ILMN_3246658	LOC728253	0.39972	1.31635
ILMN_2121408	HBEGF	0.08761	1.31623
ILMN_1661994	ESRRG	0.14097	1.31612
ILMN_1781285	DUSP1	0.05329	1.31535
ILMN_2354391	EGLN2	0.06590	1.31426
ILMN_2094166	CHMP5	0.28992	1.31378
ILMN_1712748	C14ORF129	0.11388	1.31348
ILMN_1809292	IMMP2L	0.06321	1.31298
ILMN_1704154	TNFRSF19	0.05485	1.31194
ILMN_1718285	HOXC8	0.19555	1.31120
ILMN_1826531	HS.213541	0.14660	1.31107
ILMN_1764361	DUSP16	0.05523	1.31063
ILMN_3248472	LOC100134591	0.13026	1.30818
ILMN_1712673	SASH1	0.06350	1.30787
ILMN_2361807	OS9	0.10755	1.30783
ILMN_2328776	MST4	0.12553	1.30776
ILMN_1667030	HSBP1	0.07187	1.30628
ILMN_1728445	IGFBP1	0.06270	1.30627
ILMN_2410929	PAPSS2	0.08905	1.30350
ILMN_2412807	DCTN1	0.05523	1.30337
ILMN_1739297	GALNT4	0.10668	1.30320
ILMN_2150654	ZSWIM4	0.19628	1.30266
ILMN_2352009	ACADVL	0.07966	1.30252
ILMN_1656111	MYLIP	0.07527	1.30240
ILMN_1702247	CCNDBP1	0.05860	1.30057
ILMN_1734276	PMEPA1	0.00000	-2.51808
ILMN_2092536	HSPE1	0.01082	-1.96283
ILMN_1657823	PMEPA1	0.00032	-1.87872

ILMN_1721283	HSPB6	0.00057	-1.84822
ILMN_2347068	MKNK2	0.00031	-1.69261
ILMN_1698732	PALLD	0.00032	-1.68871
ILMN_1717925	SIGMAR1	0.00202	-1.63972
ILMN_3306742	SIGMAR1	0.00577	-1.62853
ILMN_1656285	METTL7A	0.00095	-1.62128
ILMN_1663640	MAOA	0.00177	-1.61862
ILMN_1717248	C9ORF66	0.00363	-1.57531
ILMN_2124802	MT1H	0.00627	-1.55571
ILMN_2332250	ACOT7	0.00258	-1.55160
ILMN_2316540	MRPL11	0.00600	-1.54183
ILMN_1892403	SNORD13	0.06270	-1.52876
ILMN_2203807	MRP63	0.00283	-1.52339
ILMN_1686664	MT2A	0.00730	-1.51398
ILMN_3298582	LOC728873	0.00627	-1.49991
ILMN_1661078	LOC644844	0.00368	-1.49259
ILMN_2205935	SFXN1	0.00405	-1.49050
ILMN_1697802	LOC654244	0.01572	-1.48941
ILMN_2212763	ICAM3	0.00711	-1.48107
ILMN_1690802	TRMT112	0.05115	-1.47706
ILMN_1774823	RPL34	0.00740	-1.47685
ILMN_1869087	HS.40289	0.09548	-1.46587
ILMN_1651557	KDELC2	0.01082	-1.46427
ILMN_1717229	C9ORF21	0.01812	-1.44905
ILMN_1702759	TMX4	0.00956	-1.44686
ILMN_1715401	MT1G	0.01117	-1.44624
ILMN_1789502	GPC4	0.00711	-1.44578
ILMN_1691156	MT1A	0.01034	-1.44393
ILMN_1708081	LCLAT1	0.00724	-1.43883
ILMN_1815552	NTAN1	0.01680	-1.43695
ILMN_1722680	DSEL	0.04090	-1.42972
ILMN_2061405	NUP54	0.01972	-1.42930
ILMN_1745329	PRR14	0.01253	-1.42542
ILMN_1717990	CALD1	0.04177	-1.42448
ILMN_1774375	LOC284422	0.01034	-1.42103
ILMN_2322806	CAST	0.03883	-1.42044
ILMN_1671603	MED30	0.02763	-1.41717
ILMN_1702787	SEMA4A	0.04090	-1.41693
ILMN_1786658	BOLA3	0.01322	-1.41689
ILMN_1803775	HSPE1	0.10755	-1.41423
ILMN_3241021	RNY4	0.17272	-1.41193
ILMN_1742824	SPATA13	0.01812	-1.40577
ILMN_2179652	PRELID1	0.03573	-1.40438
ILMN_2370135	HNRNPU	0.01972	-1.40294
ILMN_3237623	RNY1	0.10032	-1.40284
ILMN_1804357	GNG4	0.01829	-1.40243
ILMN_1675124	DDX17	0.15722	-1.40150
ILMN_1801864	LOC730455	0.02352	-1.39701
ILMN_2088124	TMEM154	0.01318	-1.39556
ILMN_1696911	FTHL8	0.04763	-1.39298
ILMN_1789364	ZNF789	0.04177	-1.39292
ILMN_1746769	PLA2G3	0.02248	-1.39198
ILMN_2116556	LSM5	0.18175	-1.38955
ILMN_1772506	ATP5I	0.02332	-1.38743

ILMN_1782788	CSDA	0.03040	-1.38564
ILMN_3261938	LOC100130154	0.06393	-1.38402
ILMN_3227315	LOC729009	0.06624	-1.38307
ILMN_1728983	NR2C1	0.07585	-1.38193
ILMN_1912333	HS.527657	0.06529	-1.38068
ILMN_1761281	LOC441019	0.05485	-1.37237
ILMN_1866216	HS.444999	0.19628	-1.36794
ILMN_1654457	TXN2	0.08491	-1.36750
ILMN_3305447	LOC729898	0.24142	-1.36601
ILMN_1808707	FSCN1	0.02823	-1.36578
ILMN_2390821	THOC2	0.32531	-1.36282
ILMN_1657153	ACTR3	0.03265	-1.36091
ILMN_2170814	LAMP3	0.07187	-1.35975
ILMN_1671265	ING2	0.03686	-1.35933
ILMN_1680727	GLRX2	0.05485	-1.35488
ILMN_3249006	LOC100133888	0.21525	-1.35441
ILMN_3236112	TP53TG1	0.15757	-1.35422
ILMN_2112811	RPL36A	0.05523	-1.35215
ILMN_2347097	KLK15	0.03445	-1.35187
ILMN_1654060	MKNK2	0.05329	-1.35134
ILMN_1807994	PCNP	0.08049	-1.35026
ILMN_3243069	LOC100133692	0.13260	-1.35002
ILMN_3299047	XAGE1C	0.03794	-1.34997
ILMN_3307926	ADRBK1	0.03016	-1.34915
ILMN_1704431	LOC554203	0.12001	-1.34892
ILMN_2204909	XRCC2	0.06270	-1.34762
ILMN_1667213	DFFA	0.06058	-1.34717
ILMN_2197946	SCG3	0.04509	-1.34700
ILMN_3246042	LOC100134147	0.06804	-1.34610
ILMN_1849399	HS.66049	0.24134	-1.34580
ILMN_3253126	FLJ41484	0.12920	-1.34399
ILMN_3240365	MNX1	0.06251	-1.34388
ILMN_1662640	C20ORF127	0.07780	-1.34275
ILMN_1726138	EI24	0.06375	-1.34118
ILMN_1705231	SLCO2A1	0.04523	-1.34117
ILMN_1730201	DTNA	0.07057	-1.34041
ILMN_2343010	BOLA3	0.03173	-1.33881
ILMN_2082762	SNORD68	0.06346	-1.33723
ILMN_1761922	LOC647037	0.13533	-1.33604
ILMN_2136089	MTE	0.03190	-1.33411
ILMN_3305152	XAGE1A	0.07071	-1.33398
ILMN_1655710	LOC642989	0.15504	-1.33394
ILMN_1795344	GOLPH4	0.12517	-1.33376
ILMN_2173611	MT1E	0.09845	-1.33130
ILMN_1669424	LOC646531	0.04882	-1.33092
ILMN_3279219	LOC100132457	0.06727	-1.33048
ILMN_1872419	HS.544637	0.03362	-1.32953
ILMN_1803559	FLJ39632	0.06346	-1.32877
ILMN_3238859	FAM120AOS	0.05445	-1.32872
ILMN_1785703	LMOD3	0.05860	-1.32796
ILMN_3181480	FLJ36131	0.06095	-1.32740
ILMN_1722945	C6ORF52	0.03938	-1.32681
ILMN_2088410	PSMG2	0.17262	-1.32624
ILMN_2047354	C1ORF97	0.10524	-1.32551

ILMN_1742166	GRWD1	0.12171	-1.32368
ILMN_1675130	NFIC	0.19913	-1.32354
ILMN_1704164	ATP13A2	0.08491	-1.32252
ILMN_2052208	GADD45A	0.17897	-1.32222
ILMN_1670870	ALCAM	0.27034	-1.32181
ILMN_3294134	LOC389765	0.09152	-1.32127
ILMN_2091454	HBM	0.05485	-1.32026
ILMN_3243682	C1ORF93	0.07244	-1.31976
ILMN_1746525	FTHL2	0.22344	-1.31877
ILMN_1720858	C6ORF115	0.07223	-1.31801
ILMN_1692145	ZNF14	0.06624	-1.31657
ILMN_2313889	ZNF682	0.07482	-1.31476
ILMN_1764090	AK3L1	0.08049	-1.31291
ILMN_1746552	WDR17	0.07679	-1.31183
ILMN_2056167	OSTC	0.05329	-1.31120
ILMN_1731948	PCK1	0.08801	-1.31003
ILMN_1663538	CLYBL	0.06958	-1.30979
ILMN_1904980	HS.542993	0.08668	-1.30879
ILMN_2286014	CATSPER2	0.05411	-1.30873
ILMN_1802906	LOC646900	0.16310	-1.30851
ILMN_1814526	ADD3	0.19872	-1.30845
ILMN_3258321	LOC100130932	0.28417	-1.30557
ILMN_1786426	TMEM149	0.07791	-1.30523
ILMN_1666776	KCNQ2	0.08298	-1.30486
ILMN_3248882	KIAA0114	0.07482	-1.30420
ILMN_1740737	DCPS	0.07244	-1.30395
ILMN_1776105	PSPH	0.07125	-1.30323
ILMN_2150294	FKBP14	0.08483	-1.30293
ILMN_2207505	LEP	0.18991	-1.30185
ILMN_3238053	LOC100129211	0.15554	-1.30162

Table IV.6 Functional analysis of networks regulated by full-length AR genes in 22Rv1 cells

Molecules in Network	Focus Molecules	Top Functions
CHAF1A,DBNL,DNA-directed DNA polymerase,DSCC1,Dynamin,EIF2AK2,ERK1/2,FANCE,FANCI,FEN1,HMGB2,I kappa b kinase,KIAA0101,OXCT2,Pias,POLD1,POLE3,PRKRIR,RAD51,RAD51C,RBM3,RFC3,RFC5,Rfc,RMI1 (includes EG:306734),RNF144B,RPA,SLC2A1,SYN1,SYP,TIMELESS,TNFSF14,TSC22D3,UBE2I,UBE2T	28	DNA Replication, Recombination, and Repair, Cancer, Gastrointestinal Disease
26s Proteasome,Akt,alcohol group acceptor phosphotransferase,BUB3 (includes EG:12237),Caspase,CDK2,CDK6,CDT1,DENND5B,DLGAP5,DNMT1,ECT2,Estrogen Receptor,FGFR3,GMNN,HELLS,Histone H1,Histone h4,Hsp90,KDM5B,KIF23,KPNA2,MCM7,NCAPD3,NUPR1,p70 S6k,PP2A,PRKAA1,PSIP1,Rb,SECISBP2,SMC2,SUZ12,VRK1,VRK3	24	Cell Cycle, Cellular Assembly and Organization, Cellular Growth and Proliferation
AP1S2,APC,CCNE2,Cdc2,Cdc25,CDC25B,Cdc25B/C,CDC25C,CDK1,Cdk,CDK1/2,CENPJ,Cyclin A,Cyclin B,Cyclin E,DEK,E2F7,E2f,ETV5,FOXM1,HAT1,Hat,HMGA1,KIF11,Mpf,NFkB (complex),NUSAP1,PAK1IP1,PLK4,SIVA1,STMN1,Thymidine Kinase,Tnf receptor,TWIST1,WEE1	20	Cell Cycle, Cellular Assembly and Organization, Cellular Function and Maintenance
ANKZF1,ANLN,APBA1,ATP10A,BEX1,CCDC77,CCP110,COX7A2L,DLG1,Dst,E2F4,FEN1,HNF4A,ISCU,KCNJ4,KCNJ12,KIAA1804,LIN7B,LIN7C,MARCH2,MKI67,MTHFD1,MTMR2,NPAT,NR2E3,PIGS,PPARG,PPFIA2,Ppfia3,PPIL1,RARA,RDH11,RFC3,TMEM30A,YWHAB	17	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry
ANKRD32,ARL6IP1,ATF7IP,C11orf82,CASP7,CCND3,CLRN3,CNOT6L,CTSH,DUT,FIGNL1,GMNN,HJURP,IFNA2,IgG,MIS18BP1,MKI67,MPHOSPH6,NAMPT,NBN,NMNAT1,PARN,PEG10,PLEKHB2,PMS2,PRKRIR,PSRC1,PYHIN1,SMARCA4,TMSB15A,TNKS1BP1,TOB1,TP53 (includes EG:22059),UBE2T,XPC	17	Cancer, Reproductive System Disease, DNA Replication, Recombination, and Repair
Ap1,BCR,Beta Tubulin,CCT3,CD3,DAPP1,F Actin,FDPS,FTL,Hsp70,HSP90B1,IFN Beta,Iga,Igm,IL1,IL10,IL12 (complex),Immunoglobulin,INCENP,Interferon alpha,LCP1,LDL,LEMD3,LIPA,LMNB1,Nfat (family),PHGDH,PI3K (complex),SFTPD,STAT5a/b,Tgf beta,TMPO,TPT1 (includes EG:100043703),Tubulin,ZFP36L1	16	Cellular Development, Hematopoiesis, Cellular Growth and Proliferation
ARPC1A,BUB1B,CaMKII,CASZ1,CENPE,CRABP2,Creb,DHX15,ERK,FSH,GPSM2,HISTONE,Histone h3,Insulin,Integrin,Lh,LOC81691,Mapk,NOLC1,NPM2,P38 MAPK,Pdgf (complex),PDGF BB,PDK4,Pka,Pkc(s),PSMC3IP,Ras,SEPT7,Sos,TCR,Tnf,TPP1 (includes EG:1200),Vegf,WASF3	15	Cell Cycle, Cell Signaling, Cellular Function and Maintenance
ABCB6,ACRV1,ADCYAP1 (includes EG:11516),beta-estradiol,CASP7,CASP9 (includes EG:100140945),CCND3,CD24,CLDN7,COX6A2,CTSH,DEPDC1,DEPDC1B,EGF (includes EG:13645),EGR3,FGFR2,FGL1,GINS3,hemoglobin,HNF1A,IGF1R,KBTBD11,KIF20A,MEST,MLEC,MT1X,PDIA3,PECR,PIK3R3,S100G,SLC37A1,TGFBR3,THAP10,TMED6,TWIST1	14	Cancer, Cell Death, Reproductive System Disease
ADCYAP1 (includes EG:11516),AP2A1,APEX1,APRT,BOP1/LOC727967,COMMD8,CRTAP,EPDR1,FAM64A,GXYLT1,IL4 (includes EG:16189),miR-146a/miR-146b/miR-146b-5p,NGFR,OSGEP,peptidylprolyl isomerase,PFAS,PHB2,PIN1,PPIB,PTPN6,PXN,RARA,RELA,RPL28,SDF2L1,SERINC3,SLAIN1,SLAMF6,SPN,THRAP3,Tlr,TOLLIP,TRAF6,UBQLN4,WDR48	13	DNA Replication, Recombination, and Repair, Cell Death, Gene Expression
ANP32A,APAF1,APP,C18orf54,Calmodulin,CALR,CASP7,CASP9 (includes EG:100140945),CAST,CCDC14,CCDC91,Ck2,CSNK1E,DEAF1,EFHD1,EWSR1,GALC,GSK3B,IgG,MAGEA11,MXD3,NGFR,NPDC1,PER1,phosphatidylserine,PRAM1,Prss2 (includes others),PSEN1,RHO,RNA polymerase II,SERPINI1,SPA17,TRPC1,UQCRHL,VAMP2	12	Cell Death, Neurological Disease, DNA Replication, Recombination, and Repair
ARL4A,ARSA,ASPM,CALR,CASP7,CAST,CKAP5,CSNK1E,CSNK1G2,CTSH,DCP1A,FAS,FMR1,FOS,FSCN1,FXR1,GSC,IgG,IMPA2,ITGA7,KCTD2,KCTD5,LANCL1,NAT6,NUFIP2,NUP107,PHLDA2,RARA,RARG,ST8SIA1,TGFB1 (includes EG:21803),ZFP36,ZW10,ZWILCH,ZWINT	11	Protein Synthesis, Dermatological Diseases and Conditions, Inflammatory Disease

ADCYAP1 (includes EG:11516),ARRB1,BDNF,BRF2,Cg,CHRM1,CHRNB1,CNPY4,cyclic AMP,Cytochrome c,DNAJC9,DR1,DYNLL1,GPRASP2,HCAR2,HIST1H4A (includes others),Jnk,L-glutamic acid,LOC146880,MACROD1,MASTL,NAA16,NFYB,NGFR,NPB,NPBWR1,NPBWR2,PER1,PER2,RARA,RN7SK,SLC17A6,SLC17A,TBP,Ubiquitin	11	Cell Signaling, Molecular Transport, Nucleic Acid Metabolism
CNIH2,GRIA2	1	Cell Morphology, Cellular Assembly and Organization, Cellular Compromise
CDC73 (includes EG:214498),PITHD1	1	Endocrine System Disorders, Genetic Disorder, Cancer
CIAO1,RPAP1,SLFN11	1	Vitamin and Mineral Metabolism, Gene Expression, Cardiovascular Disease
B3GAT3,CHST10,galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase	1	Carbohydrate Metabolism, Molecular Transport, Small Molecule Biochemistry

Table IV.7 Functional analysis of networks regulated by AR variant genes in 22Rv1 cells

Molecules in Network	Focus Molecules	Top Functions
Akt,Alpha tubulin,ATP2C1,Calpain,CAST,CD9,Collagen type I,DICER1,DTNA,EFNB2,Erm,F Actin,Fgfr,Focal adhesion kinase,FSCN1,Growth hormone,GSTA2,ICAM3,Integrin,MAOA,Mapk,MBP,MST4,NET1 (includes EG:10276),Notch,PALLD,Pkc(s),PMEPA1,RALB,RTN4,SLC40A1,Sos,SPRY2,TSPAN3,TSPAN4	21	Cellular Assembly and Organization, Cellular Function and Maintenance, Cell Morphology
ALCAM,Alpha catenin,CCNG2,Cdc2,CDKN1B,CSDA,Cyclin A,DFFA,E2f,EGLN2,ESRRG,Estrogen Receptor,FAM46A,GADD45A,Hdac,HISTONE,IGFBP3,IGFBP1 (includes EG:16006),Immunoglobulin,JINK1/2,MT1E,MT1H,MT2A,NFIC,NFkB (complex),NFKBIZ,NR2C1,PCK1 (includes EG:18534),Rxr,SDC4,SDCBP2,TAX1BP1,TNFRSF19,TSH,VitaminD3-VDR-RXR	22	Gene Expression, Cell Death, Cell Cycle
BAMBI,CD59 (includes EG:25407),Cytochrome c,DCTN1,DDR1,DNAJC12,DUSP1,HDL,hemoglobin,Hsp27,Hsp70,HSP,HSPB6,HSPE1,IL1,IL12 (complex),LDL,MAP2K4,MKNK2,MT1A,NADPH oxidase,Nos,NR4A2,P38 MAPK,PI3K (complex),PINK1,Pro-inflammatory Cytokine,RABEP1,Sapk,SAT1,Sod,Tgf beta,TINF2,TXN2,VCAM1	18	Free Radical Scavenging, Cell Signaling, Small Molecule Biochemistry
ACPP,Actin,ACTR3,Ap1,BNIP3L,CD3,CERK,Creb,DDX17,DUSP16,ERK,FAM160B1,FSH,GLRX2, HBEGF,hCG,Histone h3,Histone h4,Ige,ING2,LEF1,LEP,Lh,MAPK4,Mek,MRPL11 (includes EG:293666),MT1G,Nfat (family),NRP1 (includes EG:18186),PRELID1,Proinsulin,RAB9A,SGK1,TCR,Vegf	19	Drug Metabolism, Lipid Metabolism, Molecular Transport
ADD3,ADRB,ATP2B1,Calcineurin protein(s),CALD1,CaMKII,CDH2,Collagen Alpha1,Collagen(s),Cpla2,EDN1,ERK1/2,FUT4,HSBP1,INPP5A,KCNQ2,LAMB2,MAP2K1/2,NMD A Receptor,p70 S6k,Pak,Pdgf (complex),PDGF BB,PDGFC,PLA2,PLA2G3,PLA2G10,Potassium Channel,PP2A,PPP1R2,Ppp2c,PPP2R2A,Rap1,RASD1,SIGMAR1	17	Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport
ABHD10,ATF4,CACNA1I,CATSPER1,CATSPER2,CATSPER3,CATSPER4,CATSPER,CATSPERB, CATSPERG,CBX5,EPB41,GOLIM4,GRWD1,HNF4A,Importin beta,IPO8,KLHL28,KPNB1,LCN2,MDFI,MRP63,NEK7,NUP54,NUP62,NUTF2,PANX2,PCNP,PPR C1,PRR14,SRP19,TMEM146,TMEM154,TRMT112,USF1	14	Cellular Movement, Reproductive System Development and Function, Cellular Assembly and Organization
APOE,AR,BMP1,C6orf115,CCND1,CCNDBP1,CCPG1,CD44 (includes EG:100330801),CDKN1A,CDKN2A,CYP1A1,D-glucose,dihydrotestosterone,DLG4,ERBB2,FGF1,GLDN,HRAS,LINC00467,LXN,MED30,NFKBIA, OSTC,PMEPA1,RB1,SASH1,SFXN1,SH3BGRL3,SIDT2,SMAD7,Sos,STBD1,Tgf beta,TSC22D1	14	Tissue Development, DNA Replication, Recombination, and Repair, Cellular Growth and Proliferation
APOE,AR,ATP5I,beta-estradiol,CBFB,CCND1,CD44 (includes EG:100330801),CDKN1A,CDKN2A,CRK,dihydrotestosterone,ERBB2,FOS,HOXC8,HRAS,IGFLR1,J UN,KDELC2,MIA3,MICAL1,miR-30c/miR-30a/miR-30d (includes others),NEDD4,NR2F6,PAPSS1,PAPSS2,PPHLN1,RAB27B,RAB5B,RB1,SMARCB1,Sos,Tgf beta,THOC2,TMX4,VIM	13	Cell Cycle, Tissue Development, Cancer

ACADVL,AK4,ANKRD52,APOE,AR,ATP13A2,CD44 (includes EG:100330801),CDKN1A,CDKN2A,CHMP5,cholesterol,COMMD10,CYCS,dihydrotestosterone,EGF (includes EG:13645),ERBB2,FGF1,G3BP2,GSTA3,HRAS,LCLAT1,LEP,lipid,LSM5 (includes EG:23658),MAPK9,MNX1,NFKBIA,PLK1,RB1,RELA,RPL36A/hCG 1787519,Sos,SREBF1,Tgf beta,UXT	13	Cell Death, Cell Cycle, DNA Replication, Recombination, and Repair
APC/APC2,APPL2,AR,ATXN7L3B,beta-estradiol,BRCA1,CCND1,CDH17,CDKN1A,CDKN2A,CGNL1,CTNNB1,ERBB2,FOS,HIST1H1C,HNRNPH2,HRAS,hydrogen peroxide,NFKBIA,PPP1CA,RB1,RNY4,RNY1 (includes EG:19872),SC5DL,SMAD7,SPATA13,SSB,Tgf beta,TMEM2,TP53 (includes EG:22059),TROVE2,TUSC3,WNT3A,XRCC2,YWHAZ	13	Cellular Growth and Proliferation, Cancer, Tissue Morphology
ABTB1,APOE,AR,BCL2,beta-estradiol,CCND1,CD44 (includes EG:100330801),CDK11A/CDK11B,CDKN1A,CDKN2A,DCPS,DEDD2,dihydrotestosterone,DYNLT3,ERBB2,FOS,HRAS,KLK15,LAMP3,METTL7A,miR-146a/miR-146b/miR-146b-5p,MMP13,NEDD4,NFKBIA,NR3C1,PSMG2,RAB11A,RB1,SATB1,SCG3,SCUBE2,Tgf beta,TNF,Trypsin,WNT3A	12	Cell Cycle, Cellular Growth and Proliferation, Tissue Morphology
ACOT7,Alp,APOE,AR,ARMCX3,BLCAP,BMP6,C14orf129,CCND1,CD44 (includes EG:100330801),CDKN1A,CLTC,CLYBL,dihydrotestosterone,DLG4,ERBB2,ESR1,FGF1,FOS,GALNT4,GALNT10,GRB2,GSK3B,HRAS,MUC1,MYLIP,NFKBIA,OS9,PMEPA1,RB1,SLCO2A1,Sos,Tgf beta,WNT3A,ZSWIM4	12	Cellular Growth and Proliferation, Cell Cycle, DNA Replication, Recombination, and Repair
APOE,AR,BMP6,CD44 (includes EG:100330801),CDKN1A,CDKN2A,COG3 (includes EG:33562),EFNB3,ERBB2,ETFA,ETS1,FGF2,FKBP14,FOS,GPC4,HRAS,IL5,KIAA1199,MAX,NFKBIA,OPRD1,PSPH,RAC1,RB1,RPL34,SEMA4A,SMAD7,Sos,TCEAL3,Tgf beta,TGFB1 (includes EG:21803),VEGFA,WDR17,XBP1 (includes EG:140614),ZNF14	12	Cellular Growth and Proliferation, Cellular Development, DNA Replication, Recombination, and Repair
26s Proteasome,ACOX1,ADCY,ADRBK1,AMPK,Calmodulin,Caspase,Ck2,DSP,G protein beta gamma,G-protein beta,G-protein gamma,GNG4,Gpcr,GRM3,HNRNPU,IgG,IKK (complex),Insulin,Jnk,Mmp,NEDD4,PAM,Pka,PRKAA,Ras,Ras homolog,RNA polymerase II,SELENBP1,SSTR1,Tubulin,Ubiquitin,WNT2,Wnt,WNT3A	10	Free Radical Scavenging, Carbohydrate Metabolism, Cell Death
AR,beta-estradiol,CCND1,CDKN1A,CDKN2A,CSNK2A1,EFNA4,EI24,ERBB2,FDXR,FOS,GTF2B,HBEGF,ID1,LEF1,MAP2K3,MAP2K4,MAPK12,mir-155,miR-20a/miR-106b/miR-17-5p (includes others),OSM,RB1,RRM2B,SMAD7,SMARCA4,TAF7,TGFB1 (includes EG:21803),TP63,TP53 (includes EG:22059),TP53AIP1,TP53INP1,TRPV6,UBE3A,VDR	7	Cell Death, Cellular Development, Cellular Growth and Proliferation
NUFIP1,SNORD13	1	RNA Post-Transcriptional Modification, Cancer, Gastrointestinal Disease
TP53TG1,ZNF217	1	Cellular Development, Cardiovascular System Development and Function, Cellular Function and Maintenance
GRAMD1A,RPS6KB1	1	Cell Cycle, Nervous System Development and Function, Cellular Development

NR5A1,TMPRSS11D	1	Cellular Movement, Embryonic Development, Endocrine System Development and Function
C1orf93,CH-OH group:NAD or NADP oxidoreductase,prostaglandin-F synthase	1	Lipid Metabolism, Small Molecule Biochemistry, Connective Tissue Disorders

Table IV.8 Expression patterns of probes in common or unique in GP3, GP4, GP6 and LN metastatic specimens

Unique				Overlap										
GP3 (32)	**GP4 (27)**	**GP5 (43)**	**Met (38)**	**GP3-4 (36)**	**GP3-5 (7)**	**GP4-5 (8)**	**GP3-Met (12)**	**GP4-Met (12)**	**GP5-Met (9)**	**GP3-4-5 (9)**	**GP3-4-Met (20)**	**GP3-5-Met (5)**	**GP4-5-Met (22)**	**GP3-4-5-Met (58)**
ABHD10	BAMBI	ABTB1	ACTR3	ADD3	CAST	BNIP3L	ANKRD9	ACADVL	CALD1	CCNG2	CALD1	CGNL1	ACOX1	ACPP
ADD3	CATSPER2	ANKRD9	ADRBK1	BLCAP	GALNT10	DDX17	CALD1	ADRBK1	CCPG1	CD59	CD59	DUSP16	ALCAM	ALCAM
ADRBK1	CCNG2	APPL2	AK3L1	CD9	HSBP1	KLK15	DUSP16	ATP2C1	DCPS	COMMD10	CD9	MYLIP	ATP2B1	ARMCX3
ATP2C1	CDH2	CALD1	ANKRD52	CDKN1B	INPP5A	MAPK4	GFOD2	BNIP3L	LEF1	GALNT10	DUSP1	NET1	CALD1	ATP2B1
CHMP5	DDX17	CDH2	CCPG1	CERK	MRP63	NFIC	IGFBP3	CCNG2	SAT1	MBP	GFOD2	PINK1	FAM120AOS	ATP5I
COMMD10	DUSP16	COG3	DDX17	CHMP5	RABEP1	RALB	MBP	CCPG1	SEMA4A	MED30	HSPB6		GNG4	BOLA3
CSDA	FKBP14	CSDA	DUSP1	EFNB3	ZNF789	STBD1	NET1	DFFA	TRMT112	MYLIP	MBP		HSPB6	CALD1
DDR1	FLJ41484	DICER1	EFNA4	FAM46A			RASD1	FKBP14		RTN4	PAPSS1		KIAA0494	CAST
DSP	FUT4	DNAJC12	FAM113B	HBM			RPL34	GALNT4			PLA2G10		MAOA	CCNG2
DUSP16	HNRNPH2	DUSP1	FSCN1	KLK15			THOC2	GNG4			SC5DL		MAP2K4	CD9
DYNLT3	HNRNPU	EGLN2	GLRX2	LOC644844			WDR17	MKNK2			SFXN1		MICAL1	EFNB3
EDN1	HOXC8	EI24	GNG4	LXN				TAX1BP1			TINF2		MT1E	ETFA
ESRRG	IMMP2L	FAM120AOS	GRM3	MBP							TMEM2		MT1G	FHDC1
GRAMD1A	KIAA0494	FLJ39632	HBEGF	MKNK2							TP53INP1		MT1H	FLJ39632
KCNQ2	LOC284422	GPC4	HNRNPU	PAM									NRP1	GRAMD1A
KIAA0494	MAOA	HBEGF	KIAA0494	PPHLN1									PPHLN1	HNRNPU
LMOD3	MBP	HNRNPU	KIAA1199	PPP1R2									SSTR1	HSPE1
NFKBIZ	NUP54	IGFBP1	LOC286367	RALB									TAF7	ICAM3
PCNP	PAM	ING2	LOC389765	RTN4									TAX1BP1	KDELC2
RABEP1	PANX2	INPP5A	NR4A2	SCUBE2										KIAA0114
THOC2	PAPSS2	KCNQ2	PAPSS2	SDC4										LAMB2
TP53INP1	SASH1	KLHL28	PRR14	SLC40A1										LAMP3
ZSWIM4	SFXN1	KLK15	RABEP1	THOC2										LCLAT1
	SSTR1	LSM5	SASH1	TMEM154										LSM5
	TMX4	MAP2K4	SC5DL	TMPRSS11D										METTL7A
		METTL7A	TMEM149	TP53TG1										MNX1
		MRP63	TUSC3	TSPAN4										MT2A
		MT1G	ZNF789	ZNF14										MYLIP
		PMEPA1												NRP1
		PPP1R2												NTAN1
		RAB5B												OSTC
		RAB9A												PALLD
		SAT1												PDGFC
		SDCBP2												PRELID1
		SELENBP1												RAB11A
		TNFRSF19												RPL36A
		TSPAN3												SH3BGRL3
														SIDT2
														SPRY2
														TCEAL3
														TINF2
														TMX4
														TNFRSF19
														ZNF789

Table IV.9 Genes correlated with disease progression

Probe Set ID	r	p value	Gene Symbol	Gene Title
210095_s_at	0.735032	9.58E-13	IGFBP3	insulin-like growth factor binding protein 3
210510_s_at	0.712515	9.46E-12	NRP1	neuropilin 1
212298_at	0.647209	2.46E-09	NRP1	neuropilin 1
212143_s_at	0.63848	4.69E-09	IGFBP3	insulin-like growth factor binding protein 3
217774_s_at	0.59159	1.08E-07	TRMT112	tRNA methyltransferase 11-2 homolog (S. cerevisiae)
227962_at	-0.58743	1.39E-07	ACOX1	acyl-Coenzyme A oxidase 1, palmitoyl
202071_at	0.57197	3.48E-07	SDC4	syndecan 4
200977_s_at	0.560916	6.52E-07	TAX1BP1	Tax1 (human T-cell leukemia virus type I) binding protein 1
223467_at	-0.54723	1.37E-06	RASD1	RAS, dexamethasone-induced 1
235591_at	0.546435	1.43E-06	SSTR1	somatostatin receptor 1
230485_at	-0.51399	7.36E-06	LOC644844	hypothetical protein LOC644844
210933_s_at	0.509907	8.93E-06	FSCN1	fascin homolog 1, actin-bundling protein (Strongylocentrotus purpuratus)
201023_at	0.501177	1.34E-05	TAF7	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 55kDa
205569_at	0.49149	2.08E-05	LAMP3	lysosomal-associated membrane protein 3
228298_at	0.48225	3.12E-05	FAM113B	family with sequence similarity 113, member B
201034_at	0.479803	3.47E-05	ADD3	adducin 3 (gamma)
214152_at	-0.47836	3.69E-05	CCPG1	cell cycle progression 1
200985_s_at	0.473419	4.56E-05	CD59	CD59 molecule, complement regulatory protein
203303_at	0.469932	5.27E-05	DYNLT3	dynein, light chain, Tctex-type 3
219197_s_at	0.46781	5.76E-05	SCUBE2	signal peptide, CUB domain, EGF-like 2
204011_at	-0.46746	5.84E-05	SPRY2	sprouty homolog 2 (Drosophila)
210592_s_at	0.461345	7.51E-05	SAT1	spermidine/spermine N1-acetyltransferase 1
210968_s_at	0.461121	7.58E-05	RTN4	reticulon 4
222802_at	0.459428	8.11E-05	EDN1	endothelin 1
231915_at	-0.45135	0.000112	ZSWIM4	zinc finger, SWIM-type containing 4
214151_s_at	-0.44866	0.000124	CCPG1	cell cycle progression 1
216855_s_at	-0.44835	0.000126	HNRNPU	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)
1554868_s_at	0.436748	0.000196	PCNP	PEST proteolytic signal containing nuclear protein
201032_at	0.434525	0.000213	BLCAP	bladder cancer associated protein
233565_s_at	0.43063	0.000247	SDCBP2	syndecan binding protein (syntenin) 2
226304_at	-0.42855	0.000266	HSPB6	heat shock protein, alpha-crystallin-related, B6
201753_s_at	0.422573	0.000331	ADD3	adducin 3 (gamma)
218995_s_at	0.420024	0.000363	EDN1	endothelin 1
238063_at	0.419351	0.000372	TMEM154	transmembrane protein 154
212942_s_at	0.416965	0.000404	KIAA1199	KIAA1199
221462_x_at	-0.41477	0.000437	KLK15	kallikrein-related peptidase 15
203986_at	0.412478	0.000473	STBD1	starch binding domain 1
201776_s_at	0.403422	0.000647	KIAA0494	KIAA0494
226996_at	0.403279	0.00065	LCLAT1	lysocardiolipin acyltransferase 1
200984_s_at	0.398772	0.000756	CD59	CD59 molecule, complement regulatory protein
200976_s_at	0.398405	0.000766	TAX1BP1	Tax1 (human T-cell leukemia virus type I) binding protein 1
236324_at	-0.39822	0.00077	MBP	Myelin basic protein
214629_x_at	0.397752	0.000783	RTN4	reticulon 4
203265_s_at	-0.39545	0.000845	MAP2K4	mitogen-activated protein kinase kinase 4
205882_x_at	0.394061	0.000885	ADD3	adducin 3 (gamma)
211509_s_at	0.39382	0.000892	RTN4	reticulon 4
201752_s_at	0.392918	0.000918	ADD3	adducin 3 (gamma)
211486_s_at	-0.39198	0.000947	KCNQ2	potassium voltage-gated channel, KQT-like subfamily, member 2
1555867_at	-0.39031	0.001	GNG4	guanine nucleotide binding protein (G protein), gamma 4
202166_s_at	0.389879	0.001014	PPP1R2	protein phosphatase 1, regulatory (inhibitor) subunit 2
233477_at	-0.38331	0.001253	KLK15	kallikrein-related peptidase 15
223218_s_at	0.377703	0.001496	NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, ze
216280_s_at	-0.3765	0.001554	DICER1	dicer 1, ribonuclease type III

201951_at	0.376133	0.001572	ALCAM	activated leukocyte cell adhesion molecule
223450_s_at	-0.3744	0.001659	COG3	component of oligomeric golgi complex 3
217816_s_at	0.372564	0.001756	PCNP	PEST proteolytic signal containing nuclear protein
203440_at	0.371585	0.001809	CDH2	cadherin 2, type 1, N-cadherin (neuronal)
213319_s_at	-0.3704	0.001876	CSDA	Cold shock domain protein A
201952_at	0.369158	0.001949	ALCAM	activated leukocyte cell adhesion molecule
228634_s_at	-0.36708	0.002076	CSDA	Cold shock domain protein A
236492_at	-0.36647	0.002115	PPP2R2A	protein phosphatase 2 (formerly 2A), regulatory subunit B, alpha isoform
231832_at	-0.36592	0.00215	GALNT4	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase
221808_at	0.365609	0.00217	RAB9A	RAB9A, member RAS oncogene family
209966_x_at	-0.36516	0.0022	ESRRG	estrogen-related receptor gamma
201615_x_at	-0.36383	0.002289	CALD1	caldesmon 1
201931_at	0.36155	0.002451	ETFA	electron-transfer-flavoprotein, alpha polypeptide
230815_at	-0.35943	0.00261	LOC389765	similar to KIF27C
200594_x_at	0.359361	0.002615	HNRNPU	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)
228257_at	0.358491	0.002684	ANKRD52	ankyrin repeat domain 52
221156_x_at	-0.35409	0.003053	CCPG1	cell cycle progression 1
215728_s_at	-0.35305	0.003147	ACOT7	acyl-CoA thioesterase 7
209227_at	-0.34746	0.003695	TUSC3	tumor suppressor candidate 3
204745_x_at	-0.34509	0.003951	MT1G	metallothionein 1G
212994_at	-0.34342	0.004142	THOC2	THO complex 2
200983_x_at	0.341971	0.004314	CD59	CD59 molecule, complement regulatory protein
224090_s_at	-0.33999	0.004558	TNFRSF19	tumor necrosis factor receptor superfamily, member 19
221511_x_at	-0.33628	0.005051	CCPG1	cell cycle progression 1
228938_at	-0.33604	0.005085	MBP	Myelin basic protein
203725_at	0.334261	0.005338	GADD45A	growth arrest and DNA-damage-inducible, alpha
213988_s_at	0.333364	0.00547	SAT1	spermidine/spermine N1-acetyltransferase 1
226599_at	0.331916	0.005689	FHDC1	FH2 domain containing 1
206461_x_at	-0.33045	0.00592	MT1H	metallothionein 1H
234072_at	-0.3289	0.006172	SEMA4A	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cyt
212859_x_at	-0.3288	0.006189	MT1E	metallothionein 1E
224232_s_at	0.326858	0.006518	PRELID1	PRELI domain containing 1
201132_at	0.326246	0.006625	HNRNPH2	heterogeneous nuclear ribonucleoprotein H2 (H')
207323_s_at	-0.3258	0.006704	MBP	myelin basic protein
200972_at	0.325715	0.006719	TSPAN3	tetraspanin 3
226442_at	-0.32457	0.006925	ABTB1	ankyrin repeat and BTB (POZ) domain containing 1
205184_at	-0.32345	0.007134	GNG4	guanine nucleotide binding protein (G protein), gamma 4
1561365_at	0.323329	0.007157	NRP1	Neuropilin 1
204347_at	-0.31974	0.007862	AK3L1	adenylate kinase 3-like 1
1554757_a_at	-0.31717	0.008404	INPP5A	inositol polyphosphate-5-phosphatase, 40kDa
201778_s_at	0.316558	0.008538	KIAA0494	KIAA0494
237989_at	-0.31648	0.008556	IGFBP1	Insulin-like growth factor binding protein 1
1554448_at	-0.3117	0.009667	LOC554203	alanyl-tRNA synthetase domain containing 1 pseudogene
220374_at	-0.30905	0.010336	KLHL28	kelch-like 28 (Drosophila)
226628_at	-0.30782	0.010661	THOC2	THO complex 2
205737_at	-0.30716	0.01084	KCNQ2	potassium voltage-gated channel, KQT-like subfamily, member 2
225395_s_at	0.300547	0.012764	FAM120AOS	family with sequence similarity 120A opposite strand
203455_s_at	0.299571	0.013072	SAT1	spermidine/spermine N1-acetyltransferase 1
204348_s_at	-0.29896	0.013269	AK3L1	adenylate kinase 3-like 1
225342_at	-0.29822	0.013508	AK3L1	adenylate kinase 3-like 1
215198_s_at	-0.298	0.013581	CALD1	caldesmon 1
201564_s_at	0.296958	0.013928	FSCN1	fascin homolog 1, actin-bundling protein (Strongylocentrotus purpuratus)
225092_at	-0.29688	0.013955	RABEP1	rabaptin, RAB GTPase binding effector protein 1
218256_s_at	0.295163	0.014543	NUP54	nucleoporin 54kDa
234459_at	-0.29433	0.014838	PPHLN1	periphilin 1
244819_x_at	-0.29259	0.015465	PSPH	phosphoserine phosphatase

218218_at	0.291981	0.015692	APPL2	adaptor protein, phosphotyrosine interaction, PH domain and leucine zipper conta
221269_s_at	0.290411	0.016287	SH3BGRL3	SH3 domain binding glutamic acid-rich protein like 3
214076_at	-0.28845	0.017056	GFOD2	glucose-fructose oxidoreductase domain containing 2
1007_s_at	0.288432	0.017065	DDR1	discoidin domain receptor tyrosine kinase 1
234495_at	-0.28784	0.017303	KLK15	kallikrein-related peptidase 15
41644_at	-0.2878	0.01732	SASH1	SAM and SH3 domain containing 1
1555765_a_at	-0.28684	0.017713	GNG4	guanine nucleotide binding protein (G protein), gamma 4
204389_at	0.286311	0.017933	MAOA	monoamine oxidase A
238440_at	-0.28558	0.018243	CLYBL	citrate lyase beta like
209043_at	-0.28321	0.019271	PAPSS1	3'-phosphoadenosine 5'-phosphosulfate synthase 1
204388_s_at	0.28315	0.019298	MAOA	monoamine oxidase A
221766_s_at	0.282783	0.019462	FAM46A	family with sequence similarity 46, member A
223518_at	0.281947	0.01984	DFFA	DNA fragmentation factor, 45kDa, alpha polypeptide
219356_s_at	0.281906	0.019859	CHMP5	chromatin modifying protein 5
221558_s_at	0.280805	0.020367	LEF1	lymphoid enhancer-binding factor 1
232770_at	-0.28068	0.020424	TUSC3	tumor suppressor candidate 3
200606_at	0.279764	0.020857	DSP	desmoplakin
1555571_at	-0.27956	0.020954	IMMP2L	IMP2 inner mitochondrial membrane peptidase-like (S. cerevisiae)
225528_at	0.279193	0.02113	IPO8	importin 8
212463_at	0.278989	0.021228	CD59	CD59 molecule, complement regulatory protein
211747_s_at	0.277501	0.021958	LSM5	LSM5 homolog, U6 small nuclear RNA associated (S. cerevisiae)
211780_x_at	0.277195	0.02211	DCTN1	dynactin 1 (p150, glued homolog, Drosophila)
210508_s_at	-0.27675	0.022336	KCNQ2	potassium voltage-gated channel, KQT-like subfamily, member 2
225408_at	-0.27405	0.023729	MBP	myelin basic protein
244852_at	-0.27334	0.02411	DSEL	dermatan sulfate epimerase-like
214552_s_at	-0.27267	0.024475	RABEP1	rabaptin, RAB GTPase binding effector protein 1
1554447_at	-0.27259	0.024519	LOC554203	alanyl-tRNA synthetase domain containing 1 pseudogene
209263_x_at	0.272585	0.024519	TSPAN4	tetraspanin 4
223032_x_at	0.272381	0.024631	PRELID1	PRELI domain containing 1
227787_s_at	0.272381	0.024631	MED30	mediator complex subunit 30
226022_at	-0.2722	0.024732	SASH1	SAM and SH3 domain containing 1
233317_at	-0.2711	0.025344	CD9	CD9 molecule
207981_s_at	-0.27104	0.025378	ESRRG	estrogen-related receptor gamma
213236_at	-0.26869	0.026725	SASH1	SAM and SH3 domain containing 1
223217_s_at	0.267976	0.027147	NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, ze
209892_at	-0.26702	0.027722	FUT4	fucosyltransferase 4 (alpha (1,3) fucosyltransferase, myeloid-specific)
223827_at	-0.26679	0.027859	TNFRSF19	tumor necrosis factor receptor superfamily, member 19
222156_x_at	-0.26553	0.028636	CCPG1	cell cycle progression 1
214614_at	0.264243	0.029445	MNX1	motor neuron and pancreas homeobox 1
223130_s_at	0.262122	0.030823	MYLIP	myosin regulatory light chain interacting protein
201692_at	-0.25978	0.032409	SIGMAR1	sigma non-opioid intracellular receptor 1
208779_x_at	0.259736	0.032437	DDR1	discoidin domain receptor tyrosine kinase 1
218633_x_at	-0.25786	0.033755	ABHD10	abhydrolase domain containing 10
212185_x_at	-0.25692	0.034431	MT2A	metallothionein 2A
231002_s_at	-0.25468	0.036092	RABEP1	Rabaptin, RAB GTPase binding effector protein 1
240336_at	-0.25411	0.036526	HBM	hemoglobin, mu
1559397_s_at	-0.25364	0.036885	PRR14	proline rich 14
213062_at	0.252149	0.038045	NTAN1	N-terminal asparagine amidase
239683_at	-0.25082	0.039103	CLYBL	citrate lyase beta like
200973_s_at	0.250803	0.03912	TSPAN3	tetraspanin 3
209112_at	0.248213	0.041258	CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)
212930_at	0.244766	0.044253	ATP2B1	ATPase, Ca++ transporting, plasma membrane 1
223044_at	0.243909	0.045025	SLC40A1	solute carrier family 40 (iron-regulated transporter), member 1
222076_at	-0.24371	0.04521	HBEGF	Heparin-binding EGF-like growth factor
212741_at	0.242135	0.046659	MAOA	monoamine oxidase A
221478_at	-0.24101	0.047716	BNIP3L	BCL2/adenovirus E1B 19kDa interacting protein 3-like

Table IV.10 Genes correlated with biochemical failure

Probe Set ID	Hazard Ratio	Confidence Interval (lower)	Confidence Interval (upper)	p value	Symbol
209966_x_at	0.485719	0.319565	0.738263	0.000723	ESRRG
210136_at	0.547273	0.373804	0.801241	0.00194	MBP
235591_at	1.41324	1.11931	1.78437	0.003644	SSTR1
227962_at	0.254348	0.094822	0.682259	0.006538	ACOX1
238063_at	2.203	1.23038	3.94445	0.007869	TMEM154
201830_s_at	4.82693	1.45045	16.0635	0.010282	NET1
225407_at	0.364757	0.167045	0.796478	0.011371	MBP
217816_s_at	30.4823	2.10568	441.269	0.012207	PCNP
1564010_at	10.6992	1.65287	69.2566	0.01287	CAST
203266_s_at	0.265341	0.091261	0.771483	0.014833	MAP2K4
210095_s_at	1.60194	1.09171	2.35063	0.016019	IGFBP3
227787_s_at	2.49346	1.17427	5.29466	0.017401	MED30
202669_s_at	0.687256	0.500259	0.944153	0.020631	EFNB2
219197_s_at	1.23648	1.03185	1.48168	0.021463	SCUBE2
201581_at	0.151552	0.029352	0.782498	0.024269	TMX4
213786_at	4.54925	1.1808	17.5268	0.0277	TAX1BP1
209043_at	0.265622	0.081281	0.868034	0.028218	PAPSS1
213102_at	5.01589	1.18734	21.1896	0.028266	ACTR3
1554685_a_at	0.001775	5.97E-06	0.527409	0.029242	KIAA1199
207981_s_at	0.623612	0.407665	0.953951	0.029453	ESRRG
207323_s_at	0.074424	0.007144	0.775285	0.02979	MBP
225092_at	0.370252	0.149991	0.913965	0.031151	RABEP1
203455_s_at	22.7502	1.29699	399.056	0.032522	SAT1
210091_s_at	0.064413	0.005174	0.801939	0.033043	DTNA
219690_at	0.349506	0.132852	0.919479	0.033163	TMEM149
215716_s_at	2.44695	1.06235	5.63616	0.035546	ATP2B1
1566513_a_at	0.00255	9.28E-06	0.700641	0.037145	GNG4
223518_at	3.25123	1.0589	9.9825	0.039399	DFFA
227786_at	2.82925	1.04167	7.68445	0.041342	MED30
208289_s_at	10.8024	1.07809	108.238	0.042976	EI24
213988_s_at	2.71771	1.02911	7.17701	0.0436	SAT1
207357_s_at	0.447834	0.202306	0.991345	0.047542	GALNT10
210510_s_at	1.32176	1.00291	1.74196	0.047629	NRP1
209264_s_at	15.8926	1.01819	248.063	0.048512	TSPAN4