Published in final edited form as:

Prostate. 2011 November ; 71(15): 1656–1667. doi:10.1002/pros.21382.

A Snapshot of the Expression Signature of Androgen Receptor Splicing Variants and Their Distinctive Transcriptional Activities

Rong Hu¹, William B. Isaacs^{1,2}, and Jun Luo^{1,2,*}

¹Departments of Urology, The Johns Hopkins University School of Medicine, Baltimore, Maryland

²Departments of Oncology, The Johns Hopkins University School of Medicine, Baltimore, Maryland

Abstract

BACKGROUND—The diversity and complexity of the human androgen receptor (AR) splicing variants are well appreciated but not fully understood. The goal of this study is to generate a comprehensive expression signature of AR variants in castration-resistant prostate cancer (CRPC), and to address the relative importance of the individual variants in conferring the castration-resistant phenotype.

METHODS—A modified RNA amplification method, termed selective linear amplification of sense RNA, was developed to amplify all AR transcripts containing AR exon 3 in CRPC specimens, which were profiled using tiling expression microarrays. Coding sequences for the AR variants were cloned into expression vectors and assessed for their transcriptional activities. Quantitative RT-PCR was used to determine their *in vivo* expression patterns in an expanded set of clinical specimens.

RESULTS—In addition to expression peaks in AR intron 3, a novel AR exon, termed exon 9, was discovered. Exon 9 was spliced into multiple novel AR variants. Different AR splicing variants were functionally distinctive, with some demonstrating constitutive activity while others were conditionally active. Conditionally active AR-Vs may activate AR signaling depending on the cellular context. Importantly, AR variant functions did not appear to depend on the full-length AR.

CONCLUSIONS—This study provided the first unbiased snapshot of the AR variant signature consisting of multiple AR variants with distinctive functional properties, directly in CRPC specimens. Study findings suggest that the aggregate function of multiple AR variants may confer a castration-resistant phenotype independent of the full-length AR.

Keywords

tiling microarray; androgen receptor splicing variant; castration-resistant prostate cancer

INTRODUCTION

The androgen receptor (AR) signaling pathway is the key molecular determinant of castration resistance in prostate cancer, and an established target for prostate cancer drug design [1]. Multiple recent studies [2–6] established that AR signaling can occur in the

complete absence of ligand binding, due to the expression of truncated AR-splicing variants that contain an intact AR N-terminal domain (NTD, encoded by AR exon 1 and DNA-binding domain (DBD, encoded by exons 2 and 3, but lack the ligand-binding domain (LBD, encoded by exons 4–8). The implications of these findings are multifaceted. First, AR signaling mediated by these truncated AR molecules is not targeted by any of the existing therapies, calling for the development of novel agents that target all AR molecules [7]. Second, assessment of AR variant levels and activities may provide indicators or predictors of therapeutic response to various hormone therapy regimens targeting the AR LBD, including those currently under Phase III trials [8,9]. In support of the clinical relevance, both mRNA and protein levels of AR-V7/AR3, one of the most extensively characterized AR variants [2,4], were dramatically elevated in clinical specimens from patients who failed hormone therapies. Interestingly, AR-V7/AR3 was also elevated, though to a less extent, in specimens from untreated (i.e., hormone naïve) patients who progressed after radical prostatectomy (RP), suggesting that AR-V7/AR3 is associated with prostate cancer progression in both castrate and non-castrate conditions.

The structural and functional diversity and complexity of AR splicing variants are well appreciated [2–6]. We reported seven AR variants that originated from splicing of transcribed “intronic” sequences (i.e., cryptic exons) downstream of AR DBD coding exons (i.e., intron 3). The seven AR variants were named numerically, going from AR-V1 to AR-V7. The two AR transcript variants identified by Dehm et al., [3] named AR1/2/2b and AR1/2/3/2b, contained spliced cryptic exons identical to those in our AR-V3 and AR-V4, respectively. The three AR variant transcripts identified by Guo et al., [4] named AR3, AR4, and AR5, contained spliced cryptic exons identical to those in our AR-V7, AR-V1, and AR-V4, respectively. More recently, Watson et al. [6] sequenced AR mRNA transcripts in the VCaP prostate cancer cell line, and identified four new AR cryptic exon splicing variants, named AR-V8 through AR-V11. In addition to the splicing of cryptic exons, constitutively active AR variants can also arise through exon skipping [5,6]. Sun et al. [5] reported the discovery of AR^{v567es}, a constitutively active AR variant encoded by an AR transcript lacking exons 5 through 7. These AR variants differ from each other in their relative abundance in clinical specimens, their functional activities, and their suitability for antibody development [2–6,10].

Despite these recent advances, the diversity and complexity of the AR variant signature are still not fully understood. Importantly, the AR variant transcriptome has not been profiled directly in CRPC specimens. Previous approaches for AR variant discovery either relied on targeted transcript cloning [2–5], or involved sequencing of 3′ rapid amplification of cDNA ends (RACE) products from a cell line [6]. Unbiased expression profiling approaches in CRPC specimens may capture the molecular portrait of the AR variant signature, thereby providing a proper context in analyzing the relative importance of the full-length AR and the AR variant signature in driving castration-resistant progression.

Expression profiling of transcripts from a specific genomic locus poses a few unique challenges in relation to detection sensitivity and accuracy. Since the targeted AR transcripts likely represent a small fraction of the entire transcriptome, it is essential to amplify the AR transcripts to facilitate their detection. PCR-based exponential amplification methods, such as 3′ RACE, are prone to both sequence- and abundance-dependent biases [11] and thus are not ideally suited for expression profiling. RNA linear amplification involves *in vitro* transcription (IVT) of double-stranded cDNA templates catalyzed by the T7 RNA polymerase [12]. IVT is widely used in expression profiling studies because it faithfully preserves the representation of the transcriptome. However, routine RNA amplification methods are designed to facilitate genome-wide analysis of the transcriptome (i.e., not for gene-specific analysis).

In this study, we developed and validated a modified RNA amplification method, termed selective linear amplification of sense RNA (SLASR), for specific, unbiased amplification of the AR transcriptome. Linearly amplified AR transcripts from the CWR22Rv1 cell line and two CRPC specimens were interrogated with expression microarrays consisting of 60-mer probes tiled across the human *AR* gene locus. This study yielded: (1) the first unbiased snapshot of the *AR* expression peaks corresponding to putative cryptic exons in relation to the canonical exons, (2) revealed a novel *AR* exon, exon 9, and (3) resulted in the identification of novel AR splicing variants. The functional diversity of the AR variants was analyzed by characterization of representative AR variants derived from this study and previous studies.

MATERIALS AND METHODS

Cell Lines and Human Prostate Tissue

PC-3, CWR22Rv1, and LNCaP cell lines were purchased from ATCC (Manassas, VA) and maintained in RPMI1640 medium (Invitrogen, Carlsbad, CA) with 10% fetal bovine serum (FBS, Sigma–Aldrich, St. Louis, MO). For androgen deprivation and R1881 treatments, cells were maintained in phenol red-free RPMI 1640 medium supplemented with 10% charcoal stripped FBS (CSS) 24 hr before treatment with R1881 (NEN, Waltham, MA) or ethanol vehicle control. Hormone naïve prostate specimens used in this study were described previously [2]. They were fresh frozen specimens harvested at the time of RP surgeries according to an established procedure [13]. Prior to RNA extraction, cryosections for normal and tumor areas of the RP specimens were prepared following manual trimming of the frozen blocks and histological verification, as described previously [2]. All tumor specimens contained at least 65% tumor. Castration-resistant prostate cancer (CRPC) specimens used in this study were also described previously [2]. They were either autopsy specimens from patients who died from prostate cancer, or transurethral resection of the prostate (TURP) specimens from patients who failed hormone therapies. RNA samples isolated from clinical specimens were stored in -80°C as small aliquots for long-term use.

Selective Enrichment and Amplification of AR Transcripts

RNA quantity and quality were determined by the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Two different RNA amplification methods (SLAAR and SLASR) were used. In selective linear amplification of antisense RNA (SLAAR), we modified the routine genome-wide amplification method provided by the Agilent Whole Genome Expression Microarray system (Agilent Technologies), by using an AR exon 3 primer during second strand cDNA synthesis. In SLASR, first strand cDNA was generated with an oligo(dT) primer. For second strand cDNA synthesis, a fusion primer, consisting of the T7 promoter, a spacer, and AR exon 3 sequence, was used to target all AR transcripts containing AR exon 3. To generate a double-stranded cDNA template containing 5' T7 primers, a third round cDNA synthesis was performed using the oligo(dT) primer. Following cDNA purification, IVT and transcript labeling was performed according to standard guidelines provided by the Agilent Whole Genome Expression Microarray system (Agilent Technologies). For both methods, 2 μg of total input RNA was used. The two methods were compared by semi-quantitative RT-PCR analysis of AR-V7 in unlabeled amplified products produced from equivalent quantities of total RNA by the two methods. To further verify the selective enrichment and amplification of AR transcripts in SLASR, the abundance of both AR-V7 and the full-length AR (AR-FL, unlabeled) derived from SLASR was compared to those from equivalent quantities of non-amplified input RNA. Primer sequences for AR-V7 and full-length AR detection were described previously [2]. Sequences of the primers used for RNA amplification are provided in Table I.

Tiling Microarrays

Tiling DNA microarrays were designed to cover both sense and antisense strands of a 200-kb interval of the X chromosome (chrX:66,680,000–66,880,000) (NCBI HG18 release), corresponding to the human *AR* gene and the immediate vicinity. Genomic DNA sequences were uploaded to the Agilent eArray server for in situ synthesis of 60-mer DNA probes. Probes were spaced 50 bp apart with 10-bp overlap. Following hybridization with labeled *AR* transcripts, the tiling array data were analyzed and viewed with the Affymetrix Integrated Genome Browser [14]. Data analysis was carried out according to published methods, following exclusion of probes with repetitive sequences [15].

Cloning of AR Splicing Variants

Sequences of the primers used for *AR* transcript cloning are provided in Table I. Reverse primers were designed according to the genomic positions of the cryptic exon 5 (P1R) and exon 9 (P2R) as revealed in the tiling microarray data. The reverse primers were paired with forward primers located in the *AR* exon 2 (P1F, P2F) for PCR amplification. The PCR products were purified and subsequently cloned into TopoTA vector (Invitrogen) for sequencing analysis. Based on the confirmed splicing junctions, the open reading frames (ORF) for AR-V9, AR-V12, AR-V13, and AR14 were cloned into pcDNA (Invitrogen) and pEGFP (Clontech Laboratories, Mountain View, CA) expression vectors, respectively.

Luciferase Reporter Assay

Luciferase and cellular localization assays were performed for various *AR* constructs in both LNCaP and PC-3 cells. For luciferase assay, cells were co-transfected with the *AR*-dependent PSAP1 luciferase reporter, an internal control pRL-CMV plasmid, together with various *AR* overexpression pcDNA constructs using Lipofectamine 2000 (Invitrogen), as described [2]. Transiently transfected cells were cultured in phenol red-free RPMI 1640 medium supplemented with 10% CSS for 24 hr, and cultured for another 24 hr in the presence of either 1 nM R1881 (NEN) or ethanol vehicle control, before being harvested for Dual Luciferase Assay (Promega, Madison, WI) or immunoblot analysis using the anti-*AR* (N20) antibody that recognizes all *AR* isoforms (Santa Cruz Biotechnology, Santa Cruz, CA). In experiments involving siRNA knockdown or forced co-expression of the full-length *AR*, siRNA targeting the *AR* exon 7 coding sequence (5'-UCAAGGAACUCGAUCGUAU-3') or a validated *AR*-FL expression construct [2] was co-transfected together with the plasmid constructs described above.

Cellular Localization of the AR Variants

To determine the cellular localization of various *AR* variant proteins, plasmid constructs expressing GFP-tagged *AR* variant proteins were transfected into the target cells individually. Transfected cells were then cultured in androgen-deprived condition (10% CSS) for 24 hr before being treated with either 1 nM R1881 or ethanol vehicle control for another 24 hr. Cells were subsequently fixed with 4% paraformaldehyde, mounted on the slides with mounting medium containing nuclei-staining DAPI (Invitrogen), and examined under a fluorescence microscope.

AR Variant Expression Analysis in Clinical Specimens

AR variant mRNA expression analysis in clinical specimens was performed using quantitative RT-PCR as described previously [2], with the only difference being the primer sequences used for each specific *AR* variant. Only primer pairs with validated amplification specificity and efficiency were used. Primer sequences for each *AR* variant are listed in Table I. Averaged threshold cycle numbers from reactions run in triplicate were used for comparative threshold analysis. For presentation purposes, all expression values were log₂

transformed with median values for the hormone naïve specimens set at 1. The Mann–Whitney U-test was used to evaluate expression differences between different groups of samples. Expression differences were considered significant under $P < 0.05$ (two-sided test).

RESULTS

Amplification of the AR Transcriptome by SLASR

We initially evaluated two different linear amplification protocols, named SLAAR and SLASR, for amplification of the AR transcriptome. To illustrate the unique advantages of the SLASR method, we compare the technical details of the two modified RNA amplification methods (SLAAR and SLASR) in Figure 1A,B. SLAAR involved a minor modification of the routine RNA amplification method to achieve AR-specific amplification. In SLAAR, an AR exon 3-specific primer was used, in place of the random primer in genome-wide linear amplification, during second-strand cDNA synthesis. Similar to the routine RNA amplification method, IVT in SLAAR starts from cDNA ends corresponding to the poly-A tail of the transcripts, yielding amplified antisense RNA (Fig. 1A). In SLASR (Fig. 1B), we designed a fusion primer (Table I) consisting of the core T7 promoter sequence [12], a spacer [16], and a sequence complementary to AR exon 3 for second strand cDNA synthesis. Following a third round of poly(dT) primed cDNA synthesis, cDNA templates were synthesized for T7 RNA polymerase-mediated linear amplification of all transcripts containing AR exon 3, which is essential for encoding a functional AR DBD, and downstream sequences that may participate in AR splicing. SLASR generates amplified sense RNA (Fig. 1B).

AR mRNA was predicted to be 10 kb in length [17,18], with undefined lengths of 3′ untranslated region (UTR). The presence of a potentially long stretch of uncharacterized 3′ UTR of the AR variants may compromise efficient amplification of distant upstream splicing junctions (SLAAR, Fig. 1A). SLASR (Fig. 1B) circumvents 3′ UTR by unidirectional IVT of cDNA templates from the 5′ end of the AR DBD coding transcripts, a direction opposite to routine RNA amplification methods and SLAAR. To verify the efficiency of SLASR in selective amplification and enrichment of AR variants, we first performed quantitative RT-PCR comparing AR transcripts in amplified RNA from SLASR to those in equivalent quantities of non-amplified input RNA. As shown in Figure 2A, both AR-V7 and the full-length AR (AR-FL) demonstrated dramatic enrichment relative to SF3A3, a house-keeping gene established in our previous study [2]. We then compared SLAAR and SLASR to determine their relative amplification efficiency. Compared with SLAAR, SLASR resulted in at least eightfold (more than three cycles) improvement of detection sensitivity supported by semi-quantitative RT-PCR detection of AR-V7 (Fig. 2B) in amplified RNA generated from identical quantities of input total RNA. No difference in AR-V7 abundance between the two methods was detected when the reactions were performed in the absence of reverse transcriptase (RT, Fig. 2B), suggesting that the improved detection of AR-V7 was due to improved IVT efficiency of AR-V7 splicing junction, and not a result of different AR-V7 cDNA levels prior to IVT.

Tiling Microarray Data

AR transcripts amplified using SLASR was profiled using an expression array of 60-mer probes tiled across the entire human *AR* locus. Shown in Figure 3A is the tiling expression microarray results from two CRPC specimens (CRPCI, CRPCII) and the CWR22Rv1 cells. We confirmed transcribed “intronic” sequences corresponding to previously characterized cryptic exons (CE1–CE4) [2]. Two additional prominent genomic expression peaks were detected, one located immediately upstream of CE3, which encodes AR-V7 (more prominent in CWR22Rv1), and the other downstream of exon 8 (more prominent in CRPC

specimens). We named the novel cryptic exon upstream of CE3 as CE5, and the downstream exon as exon 9 (Fig. 3A).

To confirm the participation of these expression peaks in alternative splicing we designed primers to amplify AR variants containing CE5 and exon 9 (Table I). Follow-up cloning and sequencing confirmed a novel splicing junction between exon 3 and CE5 (Fig. 3C), and three additional AR variants containing exon 9 (Fig. 3C). Watson et al. [6] recently discovered AR variants named AR-V8 through AR-V11. The CE5-containing variant, if translated, would contain C-terminal peptides identical to ARV9. In keeping with the nomenclature, we named the CE5-containing variant as AR-V9, and the three AR variants containing exon 9 as AR-V12, AR-V13, and AR-V14. All three exon 9-containing variants were exon-skipping variants. Sun et al. [5] reported an exon-skipping variant named AR^{v567es}. Because the same exons (exons 5 through 7) were skipped in AR^{v567es} and AR-V12, the coding sequences of AR^{v567es} and AR-V12 would be identical. However, the two transcripts differ by their 3' untranslated sequences, with a large portion of exon 8 replaced by exon 9 in AR-V12 (Fig. 3C). This difference between AR^{v567es} and AR-V12 highlighted the complexity of the AR variant signature that may compromise the accurate measurement of transcript abundance and functional activity.

Functional Characterization of the Novel AR Variants

AR-V9 is structurally similar to AR-V1 in that it does not have the basic amino acids characteristic of the bipartite nuclear localization sequence [19]. AR-V12, however, retains the full nuclear localization sequence encoded by part of exons 3 and 4 (Fig. 3C). Therefore, AR-V12 was predicted to be constitutively active, while AR-V9 was predicted to be inactive. Previous studies of various AR deletion constructs suggested that the presence of partially truncated AR LBD sequences inhibited AR activity [19,20]. Based on these findings, it was predicted that AR-V13 and AR-V14 would be functionally inactive because they contained additional exons encoding a partially truncated AR LBD (Fig. 3C). To confirm these predictions and to assess the relative ligand-independent activity of the AR variants, we performed dual-luciferase assay in the AR-negative PC-3 cells for ARV1, AR-V7, AR-V9, AR-V12, and AR-V13. We first confirmed the expression of AR variant proteins in transfected cells using an antibody that recognizes the AR NTD (Fig. 4A). As predicted and consistent with a previous study [5], AR-V12 demonstrated constitutive AR activity at levels comparable with those of AR-V7. In contrast, AR-V9 and AR-V13 were not active, demonstrating activities similar to AR-V1, previously predicted [2] and confirmed [6] to be inactive (Fig. 4C). The different transcriptional activities of the AR variants in PC-3 cells can be explained by their cellular localization (Fig. 4B). All constitutively active AR variants are constitutively nuclear localized, while all inactive AR variants localized to the cytoplasm (Fig. 4B). Addition of R1881 did not change the localization (Fig. 4B) or activity (not shown) of any of the AR variants.

We then assessed the transcriptional activities of the AR variants in the AR-positive LNCaP cells. Surprisingly, AR-V1 and AR-V9, despite their cytoplasmic localization (not shown), demonstrated constitutive activities in LNCaP cells in the absence of R1881 (Fig. 4D). We previously reported diverse functional activities of AR-V1 in different cell types [10]. This result confirmed that AR-V9, similar to AR-V1, is conditionally active depending on the cellular context.

Previous studies suggest that AR-V7 activities require the presence and function of the full-length AR (AR-FL), as knockdown or inhibition (by the anti-androgen MDV3100) of AR-FL abrogated AR-V7-mediated functions [6]. However, AR-V7 was constitutively active in both AR-negative PC-3 and AR-positive LNCaP cells (Fig. 4). Our data raised the possibility that AR-FL may modulate the activity of conditionally active AR-V1 and AR-

V9, but is inconsistent with the hypothesis that the constitutively active AR-V7 also requires AR-FL. To assess whether manipulation of AR-FL levels affect AR variant activity, we assessed the AR variant transcriptional activities in LNCaP cells following AR-FL knockdown. As shown in Figure 5A, AR-FL knockdown completely abolished the ligand-induced AR transcriptional activity, but did not affect AR-V7-mediated activity regardless of the presence or absence of R1881. Similarly, we found no evidence supporting the requirement of AR-FL for the conditionally active AR-V9, as AR-V9 remained strongly active in LNCaP cells following AR knockdown and in the absence of R1881 (Fig. 5A). Next, we assessed whether forced AR-FL expression in PC-3 cells may increase the transcriptional activity of AR-V9. Again, while ligand-dependent AR activities are increased following AR-FL expression, the transcriptional activities mediated by AR-V7 and AR-V9 did not change as a function of full-length AR expression (Fig. 5B). These results suggest that AR-V7 activity is completely independent of the full-length AR, and that the cellular context, rather than the AR status, may modulate the activity of conditionally active AR variants such as AR-V1 and AR-V9.

Expression Studies in Clinical Specimens

In our previous study, transcript levels for both AR-V1 and AR-V7 were found to be significantly elevated in CRPC specimens. Elevated expression of AR-V7, but not AR-V1, was also found in aggressive, but hormone naïve prostate tumors that subsequently progressed after surgical treatment. We assessed expression levels of AR-V9 and AR-V12 in normal ($n = 9$) and tumor specimens ($n = 77$) from hormone naïve prostate cancer patients ($n = 77$), as well as CRPC specimens ($n = 10$). Consistent with their role in conferring the castration resistance phenotype, we found significantly elevated expression levels of AR-V9 and AR-V12 in CRPC specimens (Fig. 6A). In 66 hormone naïve specimens for which complete follow-up data were available, AR-V9 and AR-V12 expression was not significantly associated with biochemical recurrence (not shown). We next assessed the possible change of AR-V9 and AR-V12 expression in tumors categorized by established prognostic factors such as pre-operative PSA (Fig. 6B), post-operative Gleason score (Fig. 6C), or disease stage (Fig. 6D). Neither AR-V9 nor AR-V12 was associated with preoperative PSA values or disease stages (Fig. 6B,D). A significant higher AR-V12 levels was found in cases with Gleason score 8 and above, when compared to those with Gleason 6 and below ($P = 0.017$, Mann–Whitney U -test) or Gleason 7 ($P = 0.038$, Mann–Whitney U -test, Fig. 6C). AR-V9 was not significantly elevated in higher Gleason prostate tumors.

DISCUSSION

In this study, we generated the first unbiased snapshot of the human AR transcriptome in CRPC specimens. We expanded the list of known AR splicing variants and provided additional insight into the pattern of genomic expression peaks in the human AR gene locus. Based on current data, we compiled a list of the known AR variants (Table II) identified so far. These variants were discovered using various approaches. With the exception of a few (AR-V3 and AR-V4) that were only present in cell lines, most of these AR variants were found to be dramatically elevated in CRPC. The canonical GT-AG sequence flanking the introns was found in all splicing junctions unique to each of the 14 AR variants, further confirming the transcript structure of the AR variants identified thus far.

The SLASR method (Fig. 1B) overcame many technical limitations of previous approaches. First, linear amplification does not affect the relative transcript abundance and the expression profile [11], therefore the original profile was faithfully preserved in the final results. Second, the use of a fusion primer allowed selective amplification of AR transcripts, increasing the detection sensitivity and reducing the non-specific signals (Fig. 2). Third, the unique 5′–3′ direction of AR transcript amplification in SLASR ensures that detection of

splicing junctions is not compromised by undefined 3' UTR's that may exist in AR transcripts (Fig. 1A and B). The advantage of the SLASR method is exemplified by the identification of the novel exon 9. A number of previous approaches relied on cloning or sequencing of AR transcripts prepared by 3' RACE. Since 3' RACE involves exponential amplification of poly-A tailed sequences, the method is inevitably biased toward short transcripts, leading to inefficient amplification of exon 9-containing transcripts. The expression signals from exon 9, at times even higher than those from exons 5 to 8 (Fig. 3), suggest its participation in alternative splicing of AR variants. Interestingly, 100% of exon 9-containing AR transcripts (7/7) we have sequenced so far are AR splicing variants (shown as four unique AR variants in Fig. 3C), suggesting a key role of exon 9 in generating AR variants.

A few salient features of the tiling array results are notable. First, we detected the previously described cryptic exon 4, located in intron 2, in CWR22Rv1 cells but not in the CRPC specimens. The detection is due to duplication of AR exon 3 and flanking introns unique to the CWR22Rv1 cell line. As the fusion primer was designed to amplify exon 3 containing sequences, this expression peak corresponds to AR1/2/3/2b [3], or AR-V4 described in our study [2]. Second, no reliable expression peaks were detected in introns 4–7, confirming results from our previous study [2]. Third, previous findings from our study [2] and Watson et al.[6] suggest that although AR variants were dramatically elevated in CRPC specimens, the absolute abundance of individual variants, such as AR-V7, was about two orders of magnitude lower than the full-length AR. While the exceptionally high AR cryptic exon signals in AR intron 3 (e.g., CE3, CE5) would be expected in CWR22Rv1 cells [2,4], the expression peaks in the two CRPC specimens (Fig. 3) were much higher than expected, suggesting a higher AR-V/AR-FL ratio in clinical specimens than previously reported [2,6]. It is possible that the three relatively abundant AR-Vs derived from cryptic exons in AR intron 3, AR-V1, AR-V7, and AR-V9, may account for a small proportion of the intron 3 expression signals, and additional AR variants involving these transcribed intron 3 sequences are yet to be discovered.

Given that alternatively spliced transcripts tend to be degraded through the nonsense-mediated decay (NMD) mechanism [21], it is critical to establish the existence of the corresponding protein product. Thus far, AR-V7 (AR3) remains the only AR variant with a proven protein product that can be detected using variant-specific antibodies [10]. Some AR variants are not suitable for antibody development due to short length of C-terminal variant-specific sequences (Table II). Future studies may establish protein translation by knockdown of specific AR variants followed by examination of the total AR variant content and activity.

Our cell line-based functional analysis of the various AR variants yielded results consistent with the cellular localization patterns predicted from the structural features of the individual variants. However, in the AR-positive LNCaP cells, we found consistent and unexpectedly high transcriptional activities for AR-V1 and AR-V9, which were not active in PC-3 cells and did not appear to localize in the nuclei in either cell types. These results raised the possibility that AR-V1 and AR-V9 may require or modulate the full-length AR activity. However, these conditionally active AR variants did not appear to depend on the full-length AR (Fig. 5). Thus, why AR-V1, AR-V9 are conditionally active is currently not explained. It is possible that AR-V1 and AR-V9, though lacking the canonical nuclear localization sequence, may still enter the nuclei, albeit inefficiently, by engaging the molecular motors and the nuclear import system that is intact in LNCaP cells but deficient in PC-3 cells.

Our observation that AR variants did not rely on the full-length AR for their transcriptional activity contrasts the findings from Watson et al. [6] Since the findings were not directly comparable as different cell lines were used, future studies are required to resolve this

important issue. Continued supply of AR ligands [22], coupled with molecular alterations that sensitize the full-length AR to castrate levels of androgens, has been the prevailing explanation for castration resistance [1,23]. Based on this model, a number of promising therapeutic agents have been developed to treat CRPC [24]. Among those under Phase III clinical trials, Abiraterone targets AR ligand synthesis by inhibiting CYP17, a key steroidogenic enzyme, while MDV3100 targets the LBD of the AR [8,9,25]. Analysis of clinical specimens from these ongoing clinical trials may help to determine the relative roles of the full-length AR and AR variants in driving the castration-resistant phenotype.

CONCLUSIONS

In summary, this study provided the first snapshot of the human AR transcriptome in CRPC specimens and revealed the existence of multiple AR variants that are either constitutively active or conditionally active, supporting the functional importance of the AR variant signature. If the existence of multiple translated AR variants are further confirmed, the aggregate function of AR signature may compensate for the low abundance of individual variants in clinical specimens relative to the full-length AR. AR variants may also demonstrate different expression patterns in CRPC specimens with different clinical parameters, and provide the cancer cells with a repertoire of AR molecules that can be selected for under different treatment conditions and levels of androgen. Due to lack of sufficient number of CRPC specimens, this hypothesis was not tested in this study. It remains formally possible, based on our current data, that more detailed and unbiased analysis will uncover additional AR variants in CRPC specimens.

Acknowledgments

The study is supported by pilot project funding from the NIH/NCI Specialized Program in Research Excellence in Prostate Cancer P50CA58236 (Johns Hopkins University), the Patrick C. Walsh Prostate Cancer Research Fund (to JL as the Carolyn and Bill Stutt Scholar), and the David H. Koch Foundation.

Grant sponsor: NIH/NCI Specialized Program in Research Excellence in Prostate Cancer (Johns Hopkins University); Grant number: P50CA58236; Grant sponsor: Patrick C. Walsh Prostate Cancer Research Fund; Grant sponsor: David H. Koch Foundation.

References

1. Chen Y, Sawyers CL, Scher HI. Targeting the androgen receptor pathway in prostate cancer. *Curr Opin Pharmacol.* 2008; 8(4):440–448. [PubMed: 18674639]
2. Hu R, Dunn TA, Wei S, Isharwal S, Veltri RW, Humphreys E, Han M, Partin AW, Vessella RL, Isaacs WB, Bova GS, Luo J. Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. *Cancer Res.* 2009; 69(1):16–22. [PubMed: 19117982]
3. Dehm SM, Schmidt LJ, Heemers HV, Vessella RL, Tindall DJ. Splicing of a novel androgen receptor exon generates a constitutively active androgen receptor that mediates prostate cancer therapy resistance. *Cancer Res.* 2008; 68(13):5469–5477. [PubMed: 18593950]
4. Guo Z, Yang X, Sun F, Jiang R, Linn DE, Chen H, Kong X, Melamed J, Tepper CG, Kung HJ, Brodie AM, Edwards J, Qiu Y. A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. *Cancer Res.* 2009; 69(6):2305–2313. [PubMed: 19244107]
5. Sun S, Sprenger CC, Vessella RL, Haugk K, Soriano K, Mostaghel EA, Page ST, Coleman IM, Nguyen HM, Sun H, Nelson PS, Plymate SR. Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. *J Clin Invest.* 2010; 120(8):2715–2730. [PubMed: 20644256]
6. Watson PA, Chen YF, Balbas MD, Wongvipat J, Socci ND, Viale A, Kim K, Sawyers CL. Constitutively active androgen receptor splice variants expressed in castration-resistant prostate

- cancer require full-length androgen receptor. *Proc Natl Acad Sci USA*. 2010; 107(39):16759–16765. [PubMed: 20823238]
7. Andersen RJ, Mawji NR, Wang J, Wang G, Haile S, Myung JK, Watt K, Tam T, Yang YC, Banuelos CA, Williams DE, McEwan IJ, Wang Y, Sadar MD. Regression of castrate-recurrent prostate cancer by a small-molecule inhibitor of the amino-terminus domain of the androgen receptor. *Cancer Cell*. 2010; 17(6):535–546. [PubMed: 20541699]
 8. Attard G, Reid AH, A'Hern R, Parker C, Oommen NB, Folkard E, Messiou C, Molife LR, Maier G, Thompson E, Olmos D, Sinha R, Lee G, Dowsett M, Kaye SB, Dearnaley D, Kheoh T, Molina A, de Bono JS. Selective inhibition of CYP17 with abiraterone acetate is highly active in the treatment of castration-resistant prostate cancer. *J Clin Oncol*. 2009; 27(23):3742–3748. [PubMed: 19470933]
 9. Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, Arora V, Wongvipat J, Smith-Jones PM, Yoo D, Kwon A, Wasielewska T, Welsbie D, Chen CD, Higano CS, Beer TM, Hung DT, Scher HI, Jung ME, Sawyers CL. Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science*. 2009; 324(5928):787–790. [PubMed: 19359544]
 10. Hu R, Denmeade S, Luo J. Molecular processes leading to aberrant androgen receptor signaling and castration resistance in prostate cancer. *Exp Rev Endocrinol Metab*. 2010; 5(5):753–764.
 11. Baugh LR, Hill AA, Brown EL, Hunter CP. Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res*. 2001; 29(5):E29. [PubMed: 11222780]
 12. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci USA*. 1990; 87(5):1663–1667. [PubMed: 1689846]
 13. Bova GS, Fox WM, Epstein JI. Methods of radical prostatectomy specimen processing: A novel technique for harvesting fresh prostate cancer tissue and review of processing techniques. *Mod Pathol*. 1993; 6(2):201–207. [PubMed: 8483892]
 14. Nicol JW, Helt GA, Blanchard SG Jr, Raja A, Loraine AE. The Integrated Genome Browser: Free software for distribution and exploration of genome-scale datasets. *Bioinformatics*. 2009; 25(20):2730–2731. [PubMed: 19654113]
 15. Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G, Sementchenko V, Piccolboni A, Bekiranov S, Bailey DK, Ganesh M, Ghosh S, Bell I, Gerhard DS, Gingeras TR. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*. 2005; 308(5725):1149–1154. [PubMed: 15790807]
 16. Kerkhoven RM, Sie D, Nieuwland M, Heimerikx M, De Ronde J, Brugman W, Velds A. The T7-primer is a source of experimental bias and introduces variability between microarray platforms. *PLoS One*. 2008; 3(4):e1980. [PubMed: 18431470]
 17. Janne OA, Shan LX. Structure and function of the androgen receptor. *Ann N Y Acad Sci*. 1991; 626:81–91. [PubMed: 2058979]
 18. Tilley WD, Marcelli M, McPhaul MJ. Expression of the human androgen receptor gene utilizes a common promoter in diverse human tissues and cell lines. *J Biol Chem*. 1990; 265(23):13776–13781. [PubMed: 2380187]
 19. Zhou ZX, Sar M, Simental JA, Lane MV, Wilson EM. A ligand-dependent bipartite nuclear targeting signal in the human androgen receptor. Requirement for the DNA-binding domain and modulation by NH₂-terminal and carboxyl-terminal sequences. *J Biol Chem*. 1994; 269(18):13115–13123. [PubMed: 8175737]
 20. Jenster G, van der Korput HA, van Vroonhoven C, van der Kwast TH, Trapman J, Brinkmann AO. Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization. *Mol Endocrinol*. 1991; 5(10):1396–1404. [PubMed: 1775129]
 21. Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, Maquat LE, Frey BJ, Blencowe BJ. Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes Dev*. 2006; 20(2):153–158. [PubMed: 16418482]
 22. Mostaghel EA, Nelson PS. Intracrine androgen metabolism in prostate cancer progression: Mechanisms of castration resistance and therapeutic implications. *Best Pract Res Clin Endocrinol Metab*. 2008; 22(2):243–258. [PubMed: 18471783]

23. Knudsen KE, Scher HI. Starving the addiction: New opportunities for durable suppression of AR signaling in prostate cancer. *Clin Cancer Res.* 2009; 15(15):4792–4798. [PubMed: 19638458]
24. Chen Y, Clegg NJ, Scher HI. Anti-androgens and androgen-depleting therapies in prostate cancer: New agents for an established target. *Lancet Oncol.* 2009; 10(10):981–991. [PubMed: 19796750]
25. Antonarakis ES, Eisenberger MA. Is abiraterone acetate well tolerated and effective in the treatment of castration-resistant prostate cancer? *Nat Clin Pract Oncol.* 2009; 6(1):12–13. [PubMed: 18957947]

Fig. 1.

Illustration of technical steps in two linear amplification methods for selective amplification of AR transcripts. **A:** Selective amplification of antisense RNA (SLAAR): T7-mediated in vitro transcription (IVT) (i.e., linear amplification) using double strand cDNA templates with T7 primer sequence tagged at the 3' of the target AR transcripts; **(B)** selective amplification of sense RNA (SLASR): IVT using cDNA templates with T7 primer sequences tagged at the 5' of the target AR transcripts. SLAAR is modified from a widely used in transcript profiling method by AR-specific synthesis of second-strand cDNA. The resulting double-stranded cDNA was used as templates for IVT, generating antisense RNA species, from the 3' end of the AR transcripts. In SLASR, three rounds of cDNA synthesis was performed to generate double-stranded AR cDNA templates with the T7 priming site located upstream of AR exon 3. The cDNA templates were used for IVT, generating amplified sense RNA species containing AR exon 3 and downstream sequences of varying lengths.

Fig. 2.

Efficient AR transcript amplification and enrichment by SLASR. **A:** Quantitative RT-PCR results of SF3A3, full-length AR (AR-FL), and AR-V7 comparing RNA amplified using SLASR and equivalent amount of non-amplified input RNA extracted from CWR22Rv1 cells. Transcript levels were normalized to that of SF3A3 in non-amplified RNA, and shown in log2 scale on the Yaxis. **B:** Semi-quantitative RT-PCR comparing AR-V7 in equal amounts of amplified RNA from SLAAR versus SLASR, in the presence or absence of reverse transcription. Expression signals were that of AR-V7 at 28 cycles and 31 cycles.

Fig. 3.

Tiling microarray results and discovery of novel AR variants. **A:** Expression signature of the AR transcripts revealed by tiling expression microarrays. Four separate tracks of the human AR locus were shown, with the top track showing the canonical exon–intron boundaries in relation to the genomic coordinates (NCBI HG18 release), followed by tracks annotated by the tiling expression data from three samples, CRPCI, CRPCII, and CWR22Rv1. Values shown on the Yaxis reflected the full-range signal intensity values (from zero to maximum) for the tiled probes and were not scaled. Probes corresponding to repetitive sequences were excluded. Signal peaks for the cryptic exons and exon 9 were marked by arrows. **B:** Schematic illustration of the human AR gene and genomic coordinates of the novel splicing junctions. Open boxes are canonical exons, and shaded boxes are either cryptic exons (within intron) or the novel exon (exon 9). Genomic coordinates for the splicing junctions are marked (NCBI HG18 release). **C:** Transcript structure of cloned full-length AR-V9, AR-V12, AR-V13, AR-V14, and their translated sequences, with variant-specific sequence in gray. Genomic coordinates corresponding to the novel splicing junctions were marked underneath each AR variant transcript, and the position of the stop codons marked with arrowheads.

Fig. 4.

Transcriptional activity of the AR splicing variants in PC-3 and LNCaP cells. **A:** Western blot showing protein products corresponding to AR-V1, AR-V7, AR-V9, AR-V12, and AR-V13 in PC-3 cells. Expression constructs were transiently transfected into the PC-3 cells. **B:** Localization of different AR variants in PC-3 cells, in the presence or absence of R1881. Blue: DAPI staining. Green: AR variant. **C:** Assessment of constitutive transcriptional activities of the AR variants in PC-3 cells cultured in medium with charcoal-stripped serum in the absence of R1881. **D:** Assessment of transcriptional activities of the AR variants in LNCaP cells cultured in medium with charcoal-stripped serum in the absence of R1881.

Fig. 5. Transcriptional activities of the AR variants do not depend on full-length AR (AR-FL). **A:** Knockdown of AR-FL does not affect the constitutive activity of AR-V7 and minimally affect the conditional activity of AR-V9 in LNCaP cells cultured in the presence or absence of R1881. **B:** Forced expression of AR-FL in PC-3 does not confer conditional activity of AR-V9 in the absence of R1881.

Fig. 6.

Quantitative RT-PCR analysis of AR-V9 and AR-V12 in clinical specimens. **A:** Relative expression levels of AR-V9 and AR-V12 in normal, tumor, and CRPC specimens. PCa, hormone naïve prostate cancer specimens from radical prostatectomies; CRPC, castration-resistant prostate cancer specimens. **B:** Lack of correlation between AR-V9 and AR-V12 expression and pre-operative PSA. **C:** Correlation between AR-V9 and AR-V12 expression and Gleason score. **D:** Lack of correlation between AR-V9 and AR-V12 with disease stage. Correlations in B^D were assessed in hormone naïve specimens with available clinical and pathological data.

TABLE I

Primer Sequences Designed and Used in This Study

Name sequence	Primer	Figures
T7-oligo(dT)	5'-AATTAATACGACTCACTATAGGGAGATTTTTTTTTTTTTTTN	Figure 1
T7-exon3	5'-AATTAATACGACTCACTATAGGGAGATGTCCATCTTGTCGTCTTCGGAAATG	Figure 1
P1/2(F)	5'-TGTCACATATGGAGCTCTCACATGTGG	Figure 3
P1(R)	5'-CATTGTGGCCAACATGACACTTCA	Figure 3
P2(R)	5'-CAGGCACTTCACTGTAGGACAGTT	Figure 3
P3(F) (V9)	5'-CCATCTTGTCGTCTTCGGAAATGTTATGAAGC	Figure 6
P3(R) (V9)	5'-TTAGTTCTACTTCTTAACAACGTGATCCCA	Figure 6
P4(F) (V12)	5'-GCCTTGCTGATTGCGAG	Figure 6
P4(R) (V12)	5'-CATGTGTGACTTGATTAGCAGGTCAAA	Figure 6

TABLE II

Diverse Characteristics of Human Androgen Receptor Splicing Variants

AR-Vs	Splicing junction	Length of variant-specific peptide	Alternative names	Transcriptional activity	In vivo protein detection	References
AR-V1	3/CE1	19	AR4	Conditional	No	[2,4]
AR-V2	3/3/CE1	19	N/A	Unknown	No	[2]
AR-V3	2/CE4	53	AR1/2/2b	Constitutive	No	[2,3]
AR-V4	3/CE4	53	AR1/2/3/2b, AR5	Constitutive	No	[2-4]
AR-V5	3/CE2	1	N/A	Unknown	No	[2]
AR-V6	3/CE2	6	N/A	Unknown	No	[2]
AR-V7	3/CE3	16	AR3	Constitutive	Yes	[2,4]
AR-V8	3/intron 3	10	N/A	Unknown	No	[6]
AR-V9	3/CE5	23	N/A	Conditional	No	This study [6]
AR-V10	3/intron 3	39	N/A	Unknown	No	[6]
AR-V11	3/intron 3	20	N/A	Unknown	No	[6]
AR-V12	4/8/9	10	ARv567es	Constitutive	No	This study [5]
AR-V13	6/9	3	N/A	Unknown	No	This study
AR-V14	7/9	7	N/A	Unknown	No	This study