Intragenic Rearrangement and Altered RNA Splicing of the Androgen Receptor in a Cell-Based Model of Prostate Cancer Progression

Yingming Li1, Majid Alsagabi2, Danhua Fan3, G. Steven Bova4, Ahmed H. Tewfik2, and Scott M. Dehm1,5

Abstract

Androgen depletion for advanced prostate cancer (PCa) targets activity of the androgen receptor (AR), a steroid receptor transcription factor required for PCa growth. The emergence of lethal castration-resistant PCa (CRPCa) is marked by aberrant reactivation of the AR despite ongoing androgen depletion. Recently, alternative splicing has been described as a mechanism giving rise to COOH-terminally truncated, constitutively active AR isoforms that can support the CRPCa phenotype. However, the pathologic origin of these truncated AR isoforms is unknown. The goal of this study was to investigate alterations in AR expression arising in a cell-based model of PCa progression driven by truncated AR isoform activity. We show that stable, high-level expression of truncated AR isoforms in 22Rv1 CRPCa cells is associated with intragenic rearrangement of an approximately 35-kb AR genomic segment harboring a cluster of previously described alternative AR exons. Analysis of genomic data from clinical specimens indicated that related AR intragenic copy number alterations occurred in CRPCa in the context of AR amplification. Cloning of the break fusion junction in 22Rv1 cells revealed long interspersed nuclear elements (LINE-1) flanking the rearranged segment and a DNA repair signature consistent with microhomology-mediated, break-induced replication. This rearrangement served as a marker for the emergence of a rare subpopulation of CRPCa cells expressing high levels of truncated AR isoforms during PCa progression in vitro. Together, these data provide the first report of AR intragenic rearrangements in CRPCa and an association with pathologic expression of truncated AR isoforms in a cell-based model of PCa progression. Cancer Res; 71(6); 1–9. ©2011 AACR.

Introduction

Prostate cancer (PCa) is the most frequently diagnosed male cancer in the United States and the second leading cause of male cancer deaths (1). Normal prostate tissue requires androgens for healthy function and cellular homeostasis. Androgens exert their cellular action by binding to the androgen receptor (AR), a 110-kDa transcription factor and member of the steroid nuclear receptor family (2). Initially, PCa depends on normal androgenic activation of the AR for ongoing growth and survival and presents as an androgen- and AR-dependent disease. Therefore, androgen depletion is the standard systemic therapy for locally advanced or metastatic PCa (3). The limitation of androgen depletion is that PCa eventually recurs with a lethal, castration-resistant phenotype. Although this stage of the disease seems to be independent of normal androgenic signaling, it is well established that castration-resistant PCa (CRPCa) remains AR-dependent through various mechanisms of aberrant AR activation and the AR remains an important therapeutic target for CRPCa (4).

AR mutations, which can broaden AR ligand specificity, and AR amplification, which can lead to AR protein overexpression, are 2 genomic mechanisms that can support the CRPCa phenotype (5–14). Ligand-independent AR activation has also been described and can occur through enhanced dependence on mitogenic signaling cascades that converge on the AR and associated transcriptional coregulators (15). More recently, alternative splicing was described as a mechanism of aberrant AR activation in CRPCa (16–20). Wild-type AR is a modular protein with an NH2-terminal (NTD) transcriptional activation function-1 (AF-1) domain, a central DNA-binding domain (DBD), and a dual-function COOH-terminal ligand-binding domain/AF-2 domain. Splicing of cryptic exons or exon skipping can yield truncated AR isoforms consisting of the NTD, DBD, and short, variable-length C-terminal extensions (16–20). These truncated AR isoforms are constitutively active and can support various features of the CRPCa phenotype.
such as the androgen-independent activation of AR target genes and androgen-independent growth. Importantly, truncated AR isoforms have been observed in various PCa cell lines, xenografts, and clinical samples, which supports an important role in disease progression (16–20). Alternative splicing is a widespread mechanism for increasing diversity from a single gene (21), and normal regulation of this process is disrupted in pathologic conditions such as cancer (22). The discovery of alternatively spliced AR isoforms has underscored the importance of understanding how AR splicing may be disrupted in CRPCa. This could provide clues to how truncated AR isoforms play a pathologic role at later stages of the disease. Therefore, the purpose of this study was to investigate the mechanisms underlying changes in AR isoform expression in a cell-based model of PCa progression.

Materials and Methods

Cell culture

Benign prostate BPH-1 cells were generously provided by Dr. Haojie Huang (University of Minnesota) and cultured in RPMI 1640 (Invitrogen) with 10% FBS (Invitrogen). The CRPCa 22Rv1 cell line was obtained from American Type Culture Collection and cultured in RPMI 1640 medium with 10% FBS. Androgen-dependent PCa CWR22Pc cells were generously provided by Dr. Marja Nevalainen (Thomas Jefferson University; ref. 23) and cultured in RPMI 1640 supplemented with 10% FBS, 2.5 mmol/L L-glutamine, and 0.8 nmol/L dihydrotestosterone (DHT; Sigma). Cell growth in RPMI 1640 medium containing 10% charcoal-stripped serum (CSS) ± 1 nmol/L DHT was monitored by crystal violet staining. For androgen response experiments, cells were cultured in RPMI 1640 + 10% CSS for 48 hours, treated at f = 0 with 1 nmol/L DHT (Sigma) or vehicle (EtOH), and then harvested at indicated time points. For long-term androgen deprivation, 22Rv1 cells were cultured in RPMI 1640 + 10% CSS for 7 days and then split to fresh plates in RPMI 1640 + 10% CSS. Cells were trypsinized and reseeded in RPMI 1640 + 10% CSS after an additional 10 days to disperse emerging foci of growth. Samples were harvested following 7, 12, 17, 22, 27, and 32 days of culture in RPMI 1640 + 10% CSS.

Western blot

Western blotting of CWR22Pc and 22Rv1 lysates with AR (Santa Cruz N-20), ERK-2 (Santa Cruz D-2), and ARV-7 (Precision Antibody no. AG10008) antibodies was conducted exactly as described (16).

Quantitative real-time RT-PCR

Total cellular RNA was isolated from CWR22Pc and 22Rv1 cells as described (24). RNA was reverse transcribed using a RT kit and an oligo(dT) primer (Roche). Absolute quantitation of AR mRNA species was carried out using forward and reverse primers listed in Supplementary Table S1. Concurrently, quantitative PCR with serial dilutions of plasmids harboring wild-type AR, AR ΔN2, AR ΔN3/ΔN2, AR ΔN1/ΔN2/ΔN3/ΔN1, and AR ΔN1/ΔN2/ΔN3/ΔCE1 cDNAs was carried out using a SYBRGreen FastMx reaction cocktail (PerfeCTa; VWR Life Sciences) and an iCycler instrument (BioRad) exactly as described (16). Threshold cycle (Ct) of amplification values obtained from cDNA standards were used to construct Ct versus cDNA standard copy number standard curves. The values of Ct obtained from real-time reverse transcriptase PCR (RT-PCR) were plotted on these standard curves to derive copy number values for individual AR mRNA isoforms. For relative quantitation, fold expression change relative to glyceraldehyde 3-phosphate dehydrogenase was determined by the comparative Ct method (2−ΔΔCt).

Genomic PCR

Genomic DNA was isolated from BPH-1, CWR22Pc, and 22Rv1 cells, using a NucleoSpin kit (Clontech). Genomic DNA from clinical CRPCa tissues was isolated as described previously (25). PCR primers were designed using the Primer3 program of the MacVector software package and are listed in Supplementary Table S1. For copy number determination, quantitative PCR with serial dilutions of BPH-1 genomic DNA was carried out for each primer pair, using SYBRGreen fastmix and an iCycler instrument. The values of Ct obtained from BPH-1 genomic DNA dilutions were used to construct Ct versus genomic copy number standard curves, with the inference that one BPH-1 genome contains one copy of the X chromosome and therefore one copy of the target region. The values of Ct obtained from test genomic DNA in real-time PCR reactions were plotted on these standard curves to derive genomic copy numbers for each of the PCR target regions. For conventional PCR, genomic DNA was amplified using a Taq Polymerase PCR kit (Qiagen), according to the manufacturer’s protocol. For long-range PCR, genomic DNA was amplified using outward facing primers (Supplementary Table S1) and a LongRange PCR Kit (Qiagen). Cloned PCR products originating from the AR locus were completely sequenced to identify the 22Rv1 AR locus break fusion junction.

Affymetrix Genome-Wide Human SNP Array 6.0 analysis

Affymetrix SNP6.0 profiling of primary PCa (26) and metastatic CRPCa (25) was done in previous studies. Raw data in .CEL format was obtained from the Gene Expression Omnibus web site (accession numbers GSE18333 and GSE14996). Copy numbers were calculated for each probe set, using Partek Genomics Suite 6.4 analysis software with default settings. Briefly, for each probe set, raw intensity was corrected for fragment length and sequence and the geometric means of allele intensity values were scaled to 1 (0 in log2 space). Copy number was calculated from these summarized intensities by normalizing intensity of each individual tumor samples to the mean intensity of the pooled noncancerous samples. Probe-level copy number data were used as input in an algorithm designed to determine the collection of breakpoints that satisfy the maximum likelihood between the input data and the noise-free version. The detailed algorithm is described in the Supplementary Methods section and is available in MATLAB (The MathWorks) upon request.
Results

AR intragenic rearrangement and aberrant AR mRNA splicing

The CWR22Pc cell line was recently established from the CWR22 human PCa xenograft (23). This cell line is androgen dependent for growth, which is in contrast to the CWR22-derived CRPCa 22Rv1 cell line (Fig. 1A; ref. 27). Recent reports have shown that alternatively spliced, truncated AR isoforms support constitutive AR-mediated transcription and androgen-independent proliferation of 22Rv1 cells (16–18). We therefore examined whether these isoforms were also synthesized in CWR22Pc cells. Using different PCR primer sets with different amplification efficiencies to identify the various AR mRNA isoforms precludes the use of the differential 

\[ \frac{C_t}{DDC_t} \]

method for determining relative expression by real-time PCR. We therefore pursued RT-PCR–based absolute quantification (Fig. 1B). As previously reported, full-length AR expression and high-level expression of the AR 1/2b, AR 1/2/3/2b, and AR 1/2/3/CE3 isoforms were observed at the mRNA and protein level in 22Rv1 cells (Fig. 1B and Supplementary Fig. S1). In androgen-dependent CWR22Pc cells, full-length AR expression was predominant but expression of AR 1/2/3/CE3 mRNA and protein was also detectable (Fig. 1B and Supplementary Fig. S2). No substantial change in these AR expression patterns was observed following 24 or 72 hours of treatment with androgens (Fig. 1C and Supplementary Fig. S2). Similarly, AR isoform expression was stable during 10 days of 22Rv1 cell culture in the presence or absence of androgens (Fig. 1D). Together, these data show that both androgen-dependent CWR22Pc and CRPCa 22Rv1 cells can synthesize truncated AR isoforms but 22Rv1 cells can sustain stable, high-level expression.

Because the full-length AR mRNA in 22Rv1 cells has 2 copies of exon 3, resulting in an AR DBD with 3 zinc fingers (28), and because 22Rv1 cells can efficiently synthesize mRNAs with contiguously spliced exons 1, 2, 3, and 2b (16), we hypothesized that a genomic aberration in the 180-kb AR locus at...
Xp11-12 may underlie the stable splicing alterations observed in these cells. We therefore interrogated copy number at distributed loci along the length of the AR gene. Strikingly, in castration-resistant 22Rv1 cells, we observed increased copy number of AR exons 2b, 3, and CE3, suggesting a rearrangement involving this genomic segment (Fig. 2). This aberration was not observed in the androgen-dependent CWR22Pc cell line (Fig. 2). These data therefore suggest that alternative AR isoforms, which support the castration-resistant phenotype of 22Rv1 cells, may arise via enhanced splicing of alternative exons harbored on a rearranged genomic segment in the AR locus.

**AR intragenic copy number alterations in metastatic CRPCa tissues**

To determine whether AR intragenic copy number alterations occurred in human CRPCa, we analyzed high-resolution Affymetrix Genome-Wide Human SNP Array 6.0 (SNP6.0) data derived from clinical primary PCa (n = 44 tissues) from 44 patients and metastatic CRPCa (n = 58 tissues) from 14 patients (25, 26). To localize the boundaries of putative breakpoints, we used a dynamic program that estimates the number and locations of segments adaptively on the basis of probe-level data. This analysis revealed a high incidence of rearrangement in conjunction with AR amplification, only in CRPCa, which to our knowledge is a novel phenomenon that has not been described (Supplementary Fig. S3). Because the outcome of the 22Rv1 AR rearrangement seemed to be a focal copy number increase of a segment between AR exons 2 and 4, resulting in higher dosage of exon 3 and alternative exons relative to AR exon 4 (Fig. 2), we asked whether these phenomena occurred in clinical PCa. Indeed, focal copy number increases were observed between AR exons 2/3 and 3/4 in 12 of 58 (20.7%) metastases from 6 of 14 (42.9%) subjects, which presented as rearrangement of a segment harboring some or all alternative AR exons 2b or CE1-3 (Fig. 3). For most of these CRPCa samples, the outcome was the higher gene content of a segment containing AR exon 3 and alternative exons compared with AR exon 4 (Supplementary Fig. S4). These alterations were not observed in genomic DNA samples from these subjects’ normal tissue (Supplementary Fig. S5). Focal copy number increase of this segment in a CRPCa subject was confirmed using targeted quantitative PCR (Supplementary Fig. S6). SNP6.0 analysis revealed no changes in overall AR copy number or focal alterations in this region in 44 primary PCa samples (Supplementary Fig. S4), indicating that CRPCa patients are more likely to harbor this rearrangement in at least one of their tumors than patients with localized, androgen-dependent PCa (6/14 vs. 0/44; P = 0.0000074, Fisher’s exact test). Overall, these data suggest that the region encompassing AR exon 3 may represent a “hotspot” for intragenic rearrangement in CRPCa.

**AR breakpoint junction boundaries lie within LINE-1 elements in 22Rv1 cells**

To establish with more precision the breakpoint junctions between AR exons 2 and 2b and AR exons CE3 and 4 in 22Rv1 cells, we carried out higher-resolution copy number interrogation (Fig. 4A). Using this approach, we mapped the 5’ breakpoint between AR exons 2 and 2b to a resolution of 4 kb (Fig. 4B). Concurrently, we mapped the 3’ breakpoint between AR exons CE3 and 4 to a resolution of 8 kb (Fig. 4B). Attempts to map the 5’ or 3’ breakpoints with higher resolution yielded real-time PCR products associated with very low Ct values in both reference and test DNA samples, indicating repetitive sequence. Indeed, analysis of public reference genome sequence revealed long interspersed nuclear elements (LINE-1) and low complexity (TA)n repeats and AT-rich sequence in both of these regions (Fig. 4C).

It is common for the endpoints of genomic deletions or insertions to map to repetitive elements such as LINE-1,
although the underlying mechanisms are not fully established (29, 30). One possibility is that extensive homology between LINE-1 elements at breakpoint junctions could lead to deletion on one sister chromatid and duplication on the other via nonallelic homologous recombination (NAHR; ref. 31). Pairwise alignments between the 5′ LINE-1 fragments and the full-length 3′ LINE-1 element identified a longer than 1-kb stretch of 87% sequence identity with one particular 5′ LINE-1 fragment, implicating NAHR as the basis for this rearrangement (Supplementary Fig. S7). Therefore, we conducted long-range PCR by using 2 pairs of outward facing primers to isolate the breakpoint junction in 22Rv1 cells (Fig. 5A and Supplementary Fig. S8). This resulted in long PCR products of 6,723 and 4,762 bp (Supplementary Fig. S8). Sequencing of cloned PCR products revealed that they were identical over the common 4,762 bp and localized the 22Rv1-specific 5′ and 3′ breakpoints to genomic positions 66,889,976 and 66,924,525, respectively (Fig. 5B and C). Analysis of the break fusion junction revealed 27 bp of inserted sequence (Fig. 5C). The origin of this sequence was not apparent by BLASTN and BLAT searches; however, the first 8 bp of this sequence perfectly matched an 8-bp motif at the 5′ breakpoint. Sequence alignments of the cloned break fusion junction and the 5′ and 3′ breakpoints showed virtually no extended homology through this region (Fig. 5D). However, regions of 3-bp microhomology were found at the breakpoints (Fig. 5D). Microhomology at the
breakpoints and inserted sequence at the fusion site argues against NAHR and supports a microhomology-mediated break-induced replication (MMBIR; ref. 32) mechanism of segmental duplication in 22Rv1 cells.

**Emergence of CRPCa cells during long-term CWR22Pc castration**

Using conventional and nested PCR strategies, we confirmed that the AR breakpoint observed in 22Rv1 cells was
indeed restricted to this cell line (Supplementary Fig. S8C and Fig. 6B). Previous studies have shown that androgen-dependent CWR22Pc xenograft tumors initially regress during castration but eventually recur with a CRPCa phenotype (23). To probe the link between AR intragenic rearrangement and CRPCa, we cultured CWR22Pc cells over a 1-month period in androgen-depleted medium. During the first 12 days of culture, no changes in AR protein expression patterns were observed (Fig. 6C). Interestingly, the 22Rv1 breakpoint was detected in CWR22Pc cells by nested PCR after 7 days of castration (Fig. 6D). The sensitivity of this nested PCR approach was determined to be as low as 1 to 2 genomes in limiting dilution assays (Supplementary Fig. S8D), indicating that the subpopulation of cells harboring this rearrangement was very rare. By day 17, discrete proliferative foci were apparent, which coincided with faint expression of truncated AR isoforms (Fig. 6C) and detection of the 22Rv1 breakpoint via conventional PCR (Fig. 6D). On day 17, cells were trypsinized and reseeded to disperse these proliferative foci. By day 22 and onward, androgen-independent cell growth was apparent, as was the expression of truncated AR isoforms. Together, these findings show that AR intragenic rearrangement is linked to high-level truncated AR isoform expression and CRPCa growth in a cell-based model of PCA progression.

Discussion

Recent reports describing the synthesis and function of truncated, constitutively active AR isoforms have provided a novel and conceptually simple mechanism for the resistance of CRPCa cells to androgen depletion (16–19). However, the mere presence of truncated AR isoforms does not correlate perfectly with androgen responsiveness, which highlights the importance of quantitative understanding in this area. This is especially apparent from a recent study showing that AR 1/2/3/CE3 [also termed AR-V7 (18) or AR-3 (17)] increases during progression to CRPCa but is also expressed in benign prostate tissue and hormone naïve PCA (17). Because truncated AR isoforms were originally identified in CRPCa cells derived from the CWR22 model, the recent establishment of an androgen-dependent cell line from CWR22 xenografts has permitted an evaluation of the changes in AR mRNA splicing regulation that may occur during PCA progression in a lineage-related context. One striking difference between androgen-dependent CWR22Pc cells and 22Rv1 CRPCa cells was the expression profile of full-length and alternatively spliced AR mRNAs. Although alternatively spliced AR mRNAs and protein were detectable in both cell lines, we found that 22Rv1 cells had an enhanced capacity to efficiently synthesize AR 1/2/2b, AR 1/2/3/2b, and AR 1/2/3/CE3 mRNAs. We further showed tandem duplication of an approximately 35-kb segment harboring these alternative exons as a likely basis for the deregulation of AR mRNA splicing observed in 22Rv1 cells. Interestingly, a recent study identified 2 additional alternative exons expressed in VCaP cells that are clustered on this segment between AR exons CE1 and CE3 (20). Mechanistically, such a rearrangement could impair normal splicing by lengthening of the already vast distance between the AR transcription start site and AR exon 4, increasing the likelihood of incorporating 1 of the 2 sets of alternative exons preceding exon 4, disrupting the normal genomic organization of cis-acting intronic and exonic splicing elements or any combination of these possibilities. It will be important to elucidate a clear cause–effect mechanism, but technical limitations such as
the size of the AR locus (~180 kb) and even larger aberrant locus in 22Rv1 cells (~215 kb) will have to be addressed.

AR overexpression is common in CRPCa, and AR gene amplification is thought to be a main driver of increased AR protein expression (33, 34). Most prior assessments of AR amplification in PCa tissues employed FISH, which lacks resolution and does not permit accurate copy number assessment along the length of the AR gene (7, 13, 34–36). Our findings indicate that a subset of amplified AR loci in CRPCa harbor intragenic rearrangements similar to 22Rv1, in addition to other alterations, which would clearly lead to a reconfigured AR exon organization for many of these alleles. It will therefore be important to conduct a comprehensive study of the relationship between AR intragenic rearrangements and levels of alternatively spliced AR isoforms in CRPCa to determine whether there is selection for intragenic rearrangement or whether intragenic rearrangement is simply arising as a by-product of AR gene amplification. Long-term castration of CWR22Pc cells suggests that AR intragenic rearrangement and the CRPCa growth phenotype are nearly 1 Mb upstream of the mouse AR locus (20). Together, these findings indicate that tissues displaying AR intragenic rearrangements should be prioritized for further studies of AR splice variants and their importance to PCa prognosis and therapeutic response.

It will also be important to map and sequence the break fusion junctions within the AR locus in individual CRPCa metastases to obtain a more complete understanding of the mechanisms and significance of AR intragenic rearrangements. Our work revealed LINE-1 elements at the 5′ and 3′ ends of the 22Rv1 AR rearrangement. Transposable elements (TE) are implicated in the genesis of rearrangements underlying TE-related genetic diseases, including cancer (29), and often arise through NAHR. However, sequencing the 22Rv1 AR break fusion junction revealed a 27-bp insertion of unknown origin, which opposes a NAHR-based model. Indeed, stressed cancer cells are deficient in NAHR (37) and cancer-specific rearrangements frequently contain insertions ranging from 1 to 154 bp of so-called nontemplate sequence at the break fusion junction (38–40). Therefore, a new model, MMBIR, has recently been proposed to account for this class of break fusion junctions in cancer cells (32).

In summary, our work describes a novel AR intragenic rearrangement in the 22Rv1 model of PCa progression, which is linked to enhanced synthesis of truncated AR isoforms and androgen-independent growth. We further show that similar genomic rearrangements occur in metastatic CRPCa specimens. It will be important in future studies to define whether intragenic AR rearrangements directly cause disrupted AR splicing, because a scenario of AR intralocus breaks leading to enhanced synthesis of truncated AR isoforms indicates that there may be little plasticity in the repertoire of AR isoforms synthesized. This would potentially limit the effectiveness of manipulating "alternative" splicing as a therapy for CRPCa. Nevertheless, a genomic basis for pathologic AR isoform expression may serve as a stable mechanism-based marker for resistance to androgen depletion therapies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank the Minnesota Supercomputing Institute and the Masonic Cancer Center Bioinformatics and Biostatistics Core for providing computing, bioinformatics, statistical, software, and data storage support for this project. We also thank Dr. Haojie Huang (Masonic Cancer Center, University of Minnesota) for critical reading of the manuscript.

Grant Support

This work was supported by a Young Investigator Award from the Prostate Cancer Foundation (S.M. Dehm), DOD New Investigator Award PC094584 (S.M. Dehm), NCI grants CA141011 (S.M. Dehm) and CA105217 (G.S. Bova), and NCI Cancer Center Support grant P30 077598 (G.S. Bova, S.M. Dehm). S. Dehm is a Masonic Scholar of the Masonic Cancer Center, University of Minnesota.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 2, 2010; revised December 3, 2010; accepted December 21, 2010.

Published OnlineFirst January 19, 2011; DOI: 10.1158/0008-5472.CAN-10-1998

References

Intragenic Rearrangement and Altered RNA Splicing of the Androgen Receptor in a Cell-Based Model of Prostate Cancer Progression

Yingming Li, Majid Alsagabi, Danhua Fan, et al.

Cancer Res Published OnlineFirst January 19, 2011.

Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-10-1998

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2011/01/19/0008-5472.CAN-10-1998.DC1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerres.aacrjournals.org/content/early/2011/03/04/0008-5472.CAN-10-1998. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.