Diverse *AR* Gene Rearrangements Mediate Resistance to Androgen Receptor Inhibitors in Metastatic Prostate Cancer

Yingming Li[1], Rendong Yang[1,2], Christine Henzler[3], Yeung Ho[1], Courtney Passow[4], Benjamin Auch[4], Suzanne Carreira[5], Daniel Nava Rodrigues[5], Claudia Bertan[5], Tae Hyun Hwang[6], David A. Quigley[7,8], Ha X. Dang[9,10], Colm Morrissey[11], Michael Fraser[12], Stephen R. Plymate[13,14], Christopher A. Maher[9, 10], Felix Y. Feng[7,15], Johann De Bono[5,16], Scott M. Dehm[1,17,18,*]

[1] Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota
[2] The Hormel Institute, University of Minnesota, Austin, Minnesota
[3] Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota
[4] University of Minnesota Genomics Center, University of Minnesota, Minneapolis, Minnesota
[5] The Institute for Cancer Research, London, United Kingdom
[6] Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio
[7] Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, California
[8] Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California
[9] McDonnell Genome Institute, Washington University in St. Louis, St. Louis, Missouri
[10] Department of Internal Medicine, Washington University in St. Louis, St. Louis, Missouri
[11] Department of Urology, University of Washington, Seattle, Washington
[12] Computational Biology Program, Ontario Institute for Cancer Research, Toronto, Ontario
[13] Division of Gerontology, Geriatric Medicine, University of Washington, Seattle, Washington
[14] Geriatric Research Education and Clinical Centers, VA Puget Sound Health Care System, Seattle, Washington
[15] Department of Radiation Oncology, University of California San Francisco (UCSF), San Francisco, California
[16] The Royal Marsden NHS Foundation Trust, London, United Kingdom
[17] Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota
[18] Department of Urology, University of Minnesota, Minneapolis, Minnesota

Running Title: AR Gene Rearrangements in Prostate Cancer

Keywords: Androgen receptor, prostate cancer, enzalutamide, genomics, structural rearrangements

Financial Support: Movember Foundation/Prostate Cancer Foundation Challenge Award (S.M.D.), NIH R01CA174777 (S.M.D.), U.S. Department of Defense Prostate Cancer Research Program Transformative Impact Award W81XWH-13-2-0093 (to S.R.P. and S.M.D). The University of Washington Prostate Cancer Donor Rapid Autopsy Program was supported by the Department of Defense Prostate Cancer Biorepository Network (PCBN) (W81XWH-14-2-0183), the Pacific Northwest Prostate Cancer SPORE (P50CA97186), the PO1 NIH grant (PO1 CA163227), and the Institute for Prostate Cancer Research (IPCR).

*Corresponding Author: Scott M. Dehm, Masonic Cancer Center, University of Minnesota, MCRB 560D, Mayo Mail Code 806, 420 Delaware Street SE, Minneapolis, MN 55455. Phone: 612-625-1504; Fax: 612-625-4915; E-mail: dehm@umn.edu

Conflict of Interest Disclosure Statement: The authors declare no potential conflicts of interest

Word Count: 134-Statement of Translational Relevance, 206-Abstract, 377-Introduction, 1812-Materials & Methods, 2614-Results, 963-Discussion, 1363-Figure Legends
Number of Figures: 6

Statement of Translational Relevance

Castration-resistant prostate cancer (CRPC) is the lethal manifestation of the disease that occurs when tumor cells have developed resistance to therapies that inhibit transcriptional activity of the androgen receptor. This study establishes *AR* as one of the most frequently rearranged genes in CRPC metastases. We find two main patterns of *AR* rearrangement: those concurrent with *AR* amplification and those occurring as the only detectable *AR* alteration. We demonstrate that the latter can cause enzalutamide resistance by promoting expression of constitutively active AR variants (AR-Vs). AR-Vs resulting from *AR* rearrangements resemble AR-V7, a biomarker of endocrine therapy resistance in CRPC. However, these AR-Vs are molecularly distinct from AR-V7 and undetectable by AR-V7 assays. Therefore, *AR* rearrangements could capture a broad spectrum of patients that are unlikely to receive benefit from endocrine therapies that target AR.

Abstract

PURPOSE: Prostate cancer is the second leading cause of male cancer deaths. Castration-resistant prostate cancer (CRPC) is a lethal stage of the disease that emerges when endocrine therapies are no longer effective at suppressing activity of the androgen receptor (AR) transcription factor. The purpose of this study was to identify genomic mechanisms that contribute to development and progression of CRPC.

EXPERIMENTAL DESIGN: We used whole genome and targeted DNA sequencing approaches to identify mechanisms underlying CRPC in an aggregate cohort of 272 prostate cancer patients. We analyzed structural rearrangements at the genome-wide level and carried out a detailed structural rearrangement analysis of the *AR* locus. We used genome engineering to perform experimental modeling of *AR* gene rearrangements and long-read RNA-sequencing to analyze effects on expression of AR and truncated AR variants (AR-Vs).

RESULTS: *AR* was among the most frequently rearranged genes in CRPC tumors. *AR* gene rearrangements promoted expression of diverse AR-V species. *AR* gene rearrangements occurring in the context of *AR* amplification correlated with AR overexpression. Cell lines with experimentally-derived *AR* gene rearrangements displayed high expression of tumor-specific AR-Vs and were resistant to endocrine therapies, including the AR antagonist enzalutamide.

CONCLUSIONS: *AR* gene rearrangements are an important mechanism of resistance to endocrine therapies in CRPC.

Introduction

Androgen receptor (AR) is a ligand-activated transcription factor that regulates expression of genes critical for homeostasis of the normal prostate, as well as growth and progression of prostate cancer. Accordingly, AR is the key target for systemic therapy of prostate cancer patients (1, 2). Therapeutic inhibition of AR is achieved by blocking production of the androgen ligands that bind and activate AR (testosterone and dihydrotestosterone), or by competitive antagonists that bind and repress AR. While these therapies extend patient survival, long-term efficacy is limited by inevitable evolution and progression to an aggressive disease manifestation termed castration-resistant prostate cancer (CRPC) (3).

Evolutionary changes in CRPC cells include *AR* gene amplification that provides transcriptional hypersensitivity of AR to castrate levels of androgens, mutations in the *AR* gene that broaden the spectrum of activating ligands for AR, and splicing alterations that promote expression of AR variant 7 (AR-V7) (4). AR-V7 can function as a constitutively active, ligand-independent transcription factor (5, 6), but preferentially heterodimerizes with full-length AR to transcriptionally activate canonical AR targets and also transcriptionally repress tumor suppressor genes (7-9). Collectively, CRPC cells with these AR alterations have been shown to remain dependent on full-length AR. This knowledge has driven the clinical development and regulatory approvals of many second-generation drugs that provide a more complete blockade of androgen production or bind the AR ligand binding domain with higher affinity (10-13). Drugs in this second-generation category include the androgen synthesis inhibitor abiraterone and the AR antagonist enzalutamide.

Despite widespread clinical use of abiraterone and enzalutamide, CRPC remains a uniformly fatal disease. Clinical data such as rising PSA (an AR-regulated gene) are often used as a surrogate for resistance to these AR-targeted therapies, supporting the concept that CRPC remains AR-dependent in the majority of patients (14). To understand additional mechanisms that promote AR activity in the context of clinical CRPC, we analyzed whole genome and targeted DNA sequencing (DNA-seq) data from an aggregate cohort of 272 prostate cancer patients with variable exposure to abiraterone and enzalutamide. By assessing genomic structural rearrangements that would disrupt the linear architecture of coding genes, we identified *AR* as one of the most frequently rearranged genes in CRPC patients' tumor genomes. We investigated the mechanisms by which these diverse *AR* gene rearrangement events can promote CRPC progression.

Materials and Methods

Prostate Cancer Tissues

A cohort of 41 patients with metastatic CRPC had tumor tissue collected from metastatic lesions by biopsy or surgery at The Institute of Cancer Research and Royal Marsden Hospital (ICR/RMH) between 2009-2015 (26 patients), or warm autopsy under the aegis of the Prostate Cancer Donor Program at the University of Washington as described (15 patients) (15). All patients treated at the ICR/RMH had provided written informed consent and were enrolled in institutional protocols approved by the Royal Marsden NHS Foundation Trust Hospital (London, United Kingdom) ethics review committee (reference 04/Q0801/60). For the 26 patients treated at ICR/RMH, a single tissue sample was obtained from 22 CRPC patients, 2 tissue samples from 2 independent metastatic sites were obtained from 3 CRPC patients, and 3 tissue samples from 3 independent metastatic sites were obtained from 1 CRPC patient. For 5 of these 41 CRPC patients, archival prostatectomy tissue was available from their prior surgeries at ICR/RMH. For the 15 rapid autopsy subjects, 2 tissue samples were collected from 2 independent metastatic sites per patient as described (15). A cohort of 101 patients with CPRC that had a metastatic biopsy collected through the Stand Up 2 Cancer/Prostate Cancer Foundation-funded West Coast Prostate Cancer Dream Team project has been described (16). A cohort of 130 patients that had localized prostate cancer tissue collected by surgery has been described (17). Tissue details including enzalutamide/abiraterone treatment status of CRPC patients prior to or during tissue collection are described in **Supplementary Data 1**. All studies were conducted in accordance with the Declaration of Helsinki.

DNA-seq

Whole genome DNA sequencing data from 101 metastatic biopsies collected from the 101 patient CRPC cohort was obtained from (16). Whole genome DNA sequencing data from 130 prostatectomy specimens was obtained from the International Cancer Genome Consortium (ICGC) study PRAD-CA (17). AR-targeted DNA sequencing data from 30 metastatic tissues collected by autopsy of 15 CRPC patients was obtained from dbGaP phs001223.v1.p1 (15). For the 41 tissue specimens obtained from 26 patients at ICR/RMH, DNA was isolated and submitted to the University of Minnesota Genomics Center for DNA-seq library preparation and hybrid capture with a custom SureSelect (Agilent) bait library as described (15). Post-capture sequencing libraries were pooled and sequenced with an Illumina HiSeq 2500 using 2X150bp settings and rapid run chemistry.

DNA-seq Data Analysis

DNA rearrangements (structural variants), somatic mutations, and copy number variants were called from whole genome sequencing data in the 101 patient CRPC cohort and 130 patient prostatectomy cohort as described (16, 17). DNA somatic mutations, and copy number variants were called from AR-targeted sequencing data in the 41 patient CRPC cohort as described (15). *AR* rearrangements (structural variants) were considered true positives based on identification by two independent structural variant callers (LUMPY and Delly), quantification of variant allele fraction using SHEAR, and containing at least 10 supporting split reads and 10 supporting

paired-end reads as described (15). BAM files of mapped reads were visualized in Integrative Genomics Viewer (IGV) (18).

Genomic PCR

PCR primers were designed to specific genomic regions flanking breakpoints of *AR* gene rearrangements. PCR primers are listed in **Supplementary Table 1**. Genomic DNA from CRPC tumors was subjected to whole genome amplification (WGA) using the REPLI-g Amplification kit (Qiagen) and purified using a QIAquick Nucleotide Removal Kit (Qiagen). WGA DNA or DNA isolated directly from cell lines was used for PCR with specific primer pairs using AccuStart II PCR SuperMix (Quanta Biosciences) as per the manufacturer's recommendations.

Cell lines

LNCaP cells were obtained from American Type Culture Collection (ATCC, #CRL-1740) The R1-AD1 cell line (CWR-R1, androgen-dependent 1, referred to as "deletion-negative clone 1" in the original publication) (19) is a subline derived from single-cell cloning of the CWR-R1 cell line. R1-AD1 cells are androgen responsive and contain a structurally normal copy of the AR gene (19, 20). Cells were maintained in RPMI 1640 medium supplemented with 10% FBS in a 5% CO_2 incubator at 37^{O}C. Cell line authentication and mycoplasma monitoring are described in the Supplementary Methods section.

Plasmids

The lentiGuide-Puro plasmid was a gift from Feng Zhang (Addgene plasmid # 52963). The pLenti-Cas9-GFP plasmid was a gift from David Sabatini (Addgene plasmid # 86145). Pairs of synthetic oligonucleotides (**Supplementary Table 2**) encoding gRNA sequences were annealed, phosphorylated, and ligated with lentiGuide-Puro vector that had been digested with *BsmB*I to remove a ~2kb filler.

CRISPR/Cas9 Transfection

R1-AD1 and LNCaP cells were co-transfected with 3μg pLenti-Cas9-GFP plasmid and 1.5μg each of two separate gRNA-containing lentiGuide-Puro plasmids by electroporation using a ECM 830 Square Wave Electroporator (BTX). DNA mixtures were incubated with approximately 4X10^{6} cells in 400μL of tissue culture medium in a 4mm gap-width electroporation cuvette (BTX) for 5 min at room temperature, and subjected to 250V, 10ms (R1-AD1) or 200V, 2X10ms (LNCaP) electrical pulses prior a 15 minute recovery at room temperature and plating.

Clonal Enrichment Assays with CRISPR/Cas9-Transfected Cells

Electroporated R1-AD1 and LNCaP cells were treated with 6 μg/mL puromycin for 4 days, and allowed to recover in puromycin-free medium for 2 days. Cells were trypsinized and an aliquot was stored at -20^{o}C to represent a baseline (day 0) sample. Trypsinized cells were re-seeded on 6-well plates ("2D conditions") at 500,000 cells/well in regular growth medium containing 10μM enzalutamide or 0.1% v/v DMSO as vehicle control, or medium containing charcoal-stripped (steroid-depleted) serum. Alternatively, cells were seeded for soft agar growth assays ("3D conditions") by mixing 10,000 cells with 1mL of 0.35% agar in phenol red free growth

medium and plating on 6-well plates that had been pre-filled with 1mL of 0.7% agar in growth medium containing 10μM enzalutamide or 0.1% v/v DMSO as vehicle control, or medium containing charcoal-stripped (steroid-depleted) serum. Cells grown in 2D on plates were re-fed 1-2 times per week and harvested after 7 or 14 days of culture. Cells grown in 3D in soft agar were re-fed weekly and harvested after 28 days of culture. Genomic DNA was isolated from plated cells using a NucleoSpin Tissue kit (Takara). Cells grown in 3D in soft agar were collected by heating the agar with colonies at 90^{o}C for 1 min, pelleting cells 11,000Xg for 1 min, removing the agar while it was still in liquid form, and washing the pellet 3 times with 90^{o}C PBS followed by pelleting at 11,000Xg to remove any residual agar. Genomic DNA was isolated from these pelleted colonies using a NucleoSpin Tissue kit.

Quantitative PCR Analysis of CRISPR/Cas9-Transfected Cells

Quantitative PCR was performed using PerfeCTa SYBR Green fastMix (Quanta Biosciences), with 80ng gDNA used as template and the same primers used for endpoint PCR detection of tumor-specific AR gene rearrangements (**Supplementary Table 1**). PCR reactions were analyzed using a CFX Connect Real-Time System (BioRad). Fold change in mRNA expression levels was calculated by the comparative Ct method, using the formula $2^{-(\Delta\Delta Ct)}$ where AR exon 1 PCR was used as calibrator using primers (forward: 5'-TGGATGGATAGCTACTCCGG; reverse-5'- TTTACCCTGCTGAGCTCTCC). All experiments were performed as 3 separate biological replicates, each performed in technical duplicate (n=6). Two-tailed unpaired t-tests were used to calculate P-values.

Preparation of Cell Lysates and Western Blot Analysis

At experimental endpoints, cells were harvested in 1X Laemmli buffer. Alternatively, insoluble nuclear fractions were prepared from cells as described (21) and boiled in 1X Laemmli buffer. Lysates in Laemmli buffer were subjected to western blot as described (22) using primary antibodies (AR SP107, abcam; AR-441, Santa Cruz, Actin C4, Santa Cruz; Histone H3 ab32356, abcam) diluted 1:1000 and secondary antibodies diluted 1:10,000.

AR 3'RACE

Total RNA (50ng) was used for AR 3'RACE reactions using a 5'/3'RACE kit, 2nd generation (Roche) as described (23). The final AR 3'RACE products were purified using a QIAquick PCR Purification Kit (Qiagen).

PacBio Isoform Sequencing

AR 3'RACE products were submitted to the University of Minnesota Genomics Center and converted to barcoded SMRTbell libraries using the PacBio Barcoded Adapters for Multiplex SMRT Sequencing protocol as per the manufacturer's recommendations (Pacific Biosciences). Barcoded SMRTbell libraries were pooled and prepared for diffusion loading on a Pacific Biosciences Sequel and sequenced using Sequel 2.5 chemistry (R1-X-11 cells) or Sequel 3.0 chemistry (clinical specimens).

Analysis of AR 3'RACE/PacBio Sequencing Data

Circular consensus reads requiring a minimum of three full passes were generated from raw unaligned BAM files, then de-multiplexed using Pacific Bioscience's lima software (v.1.9; https://github.com/PacificBiosciences/barcoding) and adapters that included the barcode plus the 5' and 3' RACE primers as input. IsoSeq 3.1 was used to generate high quality reads for alignment to hg19 with minimap2 (24). Pacific Bioscience's Cupcake ToFU supporting scripts (https://github.com/Magdoll/cDNA_Cupcake/wiki/Cupcake-ToFU) were used to collapse reads into transcripts, identify fusion transcripts and quantify the abundance of both. Non-fusion transcripts were summarized across samples using TAMA (https://github.com/GenomeRIK/tama/wiki). Fusion transcripts were summarized across samples using a Cupcake ToFU script. The exon structure of resulting summarized transcripts, and heatmaps and plots were generated in R version 3.5.0 using custom scripts.

Generation of R1-X-11 Cell Line

R1-AD1 cells were transfected with 3ug pLenti-Cas9-GFP and 1.5ug of each gRNA and then selected with 6ug /ml puromycin for 4 days. Cells were plated by limiting dilution in 96 well plates in RPMI 1640 medium supplemented with 10% charcoal-stripped (steroid depleted) serum. Wells containing single colonies were visualized by microscopy after 2-3 weeks and transferred to 48-well plates by trypsinization. Genomic DNA from individual clones was screened by PCR primers targeting the break junction to detect the translocation (chrX forward: 5'-CCTCTGATGCTTGGTTTTCC; chr11 reverse: 5'-GCCCAAACTTACCCAAGCTA) and PCR primers flanking the chromosome X gRNA target site in AR intron 3 (chrX forward: 5'-TGTAACAGCACCAACAGGCA-3; chrX reverse: 5'-GTACCCTTGGAAAGTGCGGA-3). Cell line R1-X-11 was identified based on a positive PCR signal for the translocation event, and negative PCR signal for retention of the wild-type intron 3 segment.

RT-PCR

Total RNA was extracted using the ReliaPrep RNA Cell Miniprep system (Promega), and 1ug total RNA was used for cDNA synthesis using cDNA qScript SuperMix (Quanta Biosciences). cDNA (1uL) was used for PCR detection of a splice junction between AR exon 3 and fusion exon 89 in chromosome 11 using primers AR exon 3 forward and chr11 exon f89 reverse (**Supplementary Table 1**). PCR products were analyzed by agarose gel electrophoresis and Sanger sequencing. For quantitative PCR, gene expression of *FASN*, *FKBP5*, and *ABCC4* was measured using primers listed in **Supplementary Table 1** and GAPDH as calibrator using the relative quantitation calculation $2^{-\Delta\Delta Ct}$.

siRNA Transfection

Cells were transfected with 40 pmoles of siRNA targeting the fusion exon 89 in chromosome 11 using electroporation as described (20). siRNA-1 had sense sequence: 5'-AAACAGAGCUCUAUCAACAUU and antisense sequence 5'-UGUUGAUAGAGCUCUGUUUUU. siRNA-2 had sense sequence: 5'-UUACCUAUCUGGAGGGUCAUU and antisense sequence: 5' UGACCCUCCAGAUAGGUAAUU.

Growth Assays

Cells were seeded at initial densities of $1x10^4$ cells/well in 48 well plates. Seeded cells were cultured in RPMI 1640 medium containing 10% fetal bovine serum supplemented with 0.1% v/v DMSO (vehicle control) or 30uM enzalutamide. Crystal violet assays were performed 24h after plating to represent Day 0. Additional wells were stained after 2, 4, or 6 days of growth.

Results

The *AR* Gene is Frequently Disrupted by Rearrangements in Metastatic Castration-Resistant Prostate Cancer

Structural rearrangements that accumulate in intragenic regions of cancer cells have potential to alter or interrupt the normal linear organization of exons that encode individual genes (**Supplementary Fig. 1A**). To identify genes that are frequently disrupted by these types of genomic rearrangements in advanced prostate cancer, we analyzed whole genome DNA-seq data from 101 prospectively-collected metastatic tumor biopsies (**Supplementary Data 1**) (16). For inclusion, a gene was required to have at least 1 genomic rearrangement breakpoint located within its gene body. We found that 3,377 genes were recurrently affected by a genomic rearrangement breakpoint in at least 2 CRPC samples (**Supplementary Fig. 1B**) and 27 genes were recurrently affected in at least 15 CRPC samples (**Supplementary Fig. 1C**). Inspection of these frequently-rearranged genes revealed that many were large (>1Mb) and therefore might have a high chance of accumulating passenger rearrangements from genomic instability (**Supplementary Fig. 1C**). We corrected for this by normalizing gene rearrangement frequency by gene size (**Supplementary Fig. 1D and Supplementary Data 2**). This normalization highlighted genes known to be affected at high frequency by structural rearrangements in prostate cancer, including oncogenic *TMPRSS2:ERG* fusions (25), and tumor suppressors *PTEN* and *TP53* (16).

Normalized data also highlighted a high frequency of structural rearrangements in the *AR* gene (23/101 tumors) (**Supplementary Fig. 1D**). *AR* gene rearrangements were observed in a previous targeted DNA-seq study of rapid autopsy subjects that died of CRPC (15). However, *AR* gene rearrangements have not been described in unbiased DNA-seq studies of CRPC genomes because these prior studies utilized whole exome sequencing (26-28), which is blind to structural rearrangements with breakpoints in intronic or intergenic regions. Therefore, the relevance of *AR* gene rearrangements to clinical CRPC is not well-established. To investigate *AR* gene rearrangements, we first evaluated their presence in the context of other known *AR* gene alterations including somatic point mutations as well as amplification of the *AR* gene body and/or amplification of an enhancer located upstream of *AR* (**Fig. 1A**). In a cohort of 130 localized prostate cancers analyzed by whole genome DNA-seq (17), no *AR* alterations were detected (**Fig. 1B**). Conversely, the *AR* gene displayed at least one alteration in 88/101 metastatic tumors (**Fig. 1B** and **Supplementary Data 3**). AR gene rearrangements were concurrent with *AR* gene and upstream enhancer amplification in 20/101 metastatic tumors, but also occurred as the only detectable alteration in 3/101 metastatic tumors.

AR Genomic Structural Rearrangements are Clonally Heterogeneous and Occur in Patients after Exposure to First- and Second-Generation AR-Targeted Therapies

Mapping *AR* gene rearrangement breakpoints revealed a heterogeneous landscape of deletion, inversion, tandem duplication, and translocation events (**Fig. 1C** and **Supplementary Table 3**).

In the 3 tumors where an *AR* gene rearrangement was the only *AR* alteration detected, the deletion and duplication events were evident from visual inspection of read coverage plots (**Fig. 1D**). Tumors that were positive for *AR* gene rearrangements had a higher burden of genomic inversions as well as cumulative burden of all genomic structural variations (deletions, inversions, tandem duplications) compared with tumors that were negative for *AR* gene rearrangements (**Supplementary Fig. 2**). *AR* gene rearrangements were independent of alterations in other prostate cancer driver genes, with the exception of *PTEN* where positive enrichment was observed (odds ratio = 3.63) (**Supplementary Fig. 3**). We also observed a slightly higher frequency of *AR* gene rearrangements in tumors from patients that had been treated with one or both of the potent second-generation AR signaling inhibitors abiraterone or enzalutamide compared with patients that were naïve to these agents and had only been treated with first-generation endocrine therapies (**Figs. 2A and B, Supplementary Data 1**).

A slightly higher frequency of *AR* gene rearrangements in abiraterone and/or enzalutamide-treated patients was also observed in a separate cohort of metastatic tumors from 41 patients where we analyzed the *AR* gene by deep targeted DNA-seq (hereafter referred to as the *AR*-targeted cohort, **Figs. 2C and D**, **Supplementary Data 1, Supplementary Data 4**, **Supplementary Tables 4-5**). Patients with *AR* gene rearrangements were observed at a higher frequency overall in this *AR*-targeted cohort (16/41, 39% of patients) compared with the whole genome DNA-seq cohort (23/101, 23% of patients). This higher frequency of *AR* rearrangement-positive patients could be due to higher detection sensitivity from the greater sequencing depth provided by *AR*-targeted sequencing, or higher probability of sampling an *AR* rearrangement-positive lesion since 19 patients in the 41-patient cohort had more than one metastatic site analyzed (**Supplementary Data 1**). The *AR* gene rearrangement events in the *AR*-targeted cohort were highly heterogeneous, similar to the 101-patient whole genome DNA-seq cohort (**Figure 2E**). All *AR* gene structural rearrangements discovered in the *AR*-targeted cohort were validated orthogonally using PCR and Sanger sequencing (**Supplementary Fig. 4**) (15).

For 5 of the *AR* gene rearrangement-positive patients in the *AR*-targeted cohort, archival prostatectomy tissue was available to enable longitudinal assessments of *AR* gene rearrangements using PCR (**Fig. 2F**). Consistent with the general observation that *AR* gene rearrangements are specific to CRPC-stage tumors (**Fig. 2B**), there was no evidence for pre-existence of cell clones harboring these specific *AR* gene rearrangements in the patients' primary tumors (**Fig. 2G-K**).

AR Genomic Rearrangements Provide a Clonal Growth Advantage to Cells under Conditions of AR-Targeted Therapy

The heterogeneous landscape of rearrangement type and intra-*AR* breakpoint location across the patients that were positive for *AR* gene rearrangements in these 2 CRPC cohorts argues against a deterministic model wherein AR-targeted endocrine therapies induce specific types of recurrent *AR* gene rearrangements at susceptible genomic sites. Further, previous studies

showing that AR-targeted endocrine therapies inhibit DNA repair by non-homologous end joining (NHEJ) (29-31) is incongruent with a deterministic model because signatures of NHEJ repair occurred at every *AR* gene rearrangement breakpoint (**Supplementary Fig. 4**). Rather, the origin of *AR* gene rearrangements is more consistent with a clonal evolution model wherein AR-targeted therapies provide selective pressure for outgrowth of rare sub-clones that harbor rearranged *AR* gene architectures.

To test this clonal evolution model, we used the CRISPR/Cas9 system to induce rearrangements in the *AR* gene in prostate cancer cell lines by targeting two DNA double-strand breaks to genomic locations that corresponded to the breakpoints observed for *AR* gene rearrangements in patients C9_a, V5300, and V5301 (**Figs. 3A and B**). We used R1-AD1 (32) and LNCaP prostate cancer cell lines for these experiments because they display androgen-sensitive growth and harbor one copy (R1-AD1) or two copies (LNCaP) of the *AR* gene on the X chromosome. We performed targeted PCR with genomic DNA to identify and track the growth of CRISPR/Cas9-engineered clones over time under different hormonal conditions that model AR-targeted therapies. Guide RNAs (gRNAs) that targeted CRISPR/Cas9-mediated DNA double strand breaks to *AR* intron 3 and chromosome 11 (chr11) induced translocation events in a subset of cells, which reflected the translocation observed in a metastasis from patient C9_a (**Fig. 3C**). When transfected cells were grown in 2-dimensional culture for 7 or 14 days, or 3-dimensional culture for 28 days, cells harboring the *AR*-chr11 translocation displayed clonal enrichment relative to the overall cell population as measured by quantitative PCR, only under conditions that modeled AR-targeted endocrine therapy (castration or enzalutamide, **Figs. 3D-F**). When we evaluated clonal evolution of cells that were engineered to harbor the two separate deletions observed in patient V5300 (**Figs. 3G-J, Supplementary Fig. 5**) or the two separate deletions observed in patient V5301 (**Figs. 5K-N, Supplementary Fig. 6**), we observed similar relative growth advantages of cells harboring targeted *AR* rearrangements under conditions of AR-targeted therapy.

Tumors Harboring *AR* Gene Rearrangements Without other *AR* Alterations Express Abundant AR Variant Transcripts that are Molecularly Distinct from AR-V7

To understand the mechanisms by which *AR* gene rearrangements could promote a clonal growth advantage under conditions that modeled AR-targeted therapies, we examined the effects of these *AR* gene rearrangements on AR mRNA expression. For this, we performed 3' rapid amplification of cDNA ends (RACE) with RNA isolated from these tumors using a forward primer anchored in *AR* exon 1. The 3'RACE reactions were subjected to long-read single molecule real time (SMRT) isoform sequencing (Iso-Seq) to simultaneously identify and quantify complete, intact AR mRNA transcripts that were expressed in these tumors. In a metastasis from patient C9_a, full-length *AR* mRNA was a major species (**Figs. 4A and B, Supplementary Table 6**). We also detected several mRNA fragments that likely originated from mis-priming of the oligo(dT) 3'RACE reverse primer at A-rich sequences in exons 7 and 2 of AR mRNA. Interestingly, we also detected expression of a minor AR variant (AR-V) mRNA species, which consisted of contiguously-spliced *AR* exons 1-3 with a 3' terminal exon derived from chr11. Presumably, this fusion transcript originated from the *AR*-chr11 translocation in this metastasis.

In a previous study, we found that metastatic tumors in patient C6_a expressed high mRNA and protein levels of ARv567es (also referred to as AR-V12) which arises from splicing of *AR* exons 1-4 and exon 8 (15). ARv567es/AR-V12 is a constitutively active form of the AR that lacks the C-terminal ligand binding domain and can drive CRPC phenotypes in experimental models (32, 33). In patient V5300, AR 3'RACE/SMRT-IsoSeq identified full-length *AR* mRNA as the major isoform, as well as high expression of ARv567e/V12 (**Figs. 4C and D**). In patient V5301, ARv567es/AR-V12 was the major isoform (**Figs. 4E and F**). To ask whether these patterns of altered AR expression were also reflected in CRISPR/Cas9 clonal evolution models, we performed western blot analysis of lysates from cells that had been transfected with Cas9 and gRNAs designed to model one of the deletions in tumor V5300 and for which enzalutamide selection drove high clonal enrichment of deletion-positive cells (**Supplementary Fig. 5D**). In this model, 1- and 2-weeks of culture under the condition of enzalutamide selection promoted high expression of AR-Vs, predominantly an AR-V species with a molecular weight that was consistent with ARv567es/AR-V12 (**Fig. 4G**).

Patient V4002 represented another patient in the 41-patient *AR*-targeted cohort that harbored an *AR* gene rearrangement as the only detectable *AR* alteration **(Fig. 2E**) at a high clonal enrichment (37%) in a lymph node metastasis (**Supplementary Data 4**). This structural rearrangement was a ~60 Mb tandem duplication within the 153 Mb X chromosome, which would not be amenable for rapid modeling using CRISPR/Cas9. In this patient's metastasis, AR 3'RACE/SMRT-IsoSeq revealed that full-length AR mRNA was a major species, with high expression of minor species that contained contiguously-spliced AR exons 1-3 and 3' terminal exons derived from the 5' end of the 60 Mb tandem duplication segment (**Figs. 4H and I**). Interestingly, AR-V7 and AR-V9 were also expressed in this metastasis, but at levels that were much lower than the novel AR fusion transcripts. Collectively, these data demonstrate that *AR* gene rearrangements occurring in the absence of other known *AR* gene alterations promote expression of diverse AR variant mRNA species that display splicing alterations after *AR* exons 1-3. Notably, these transcripts and their predicted protein products are molecularly distinct from AR-V7 (**Fig. 4J**).

Tumors Harboring *AR* Gene Rearrangements Concurrent with AR Amplification Express High Levels of AR and a Diversity of Minor AR mRNA Species

To address whether this relationship between *AR* gene rearrangements and high-level AR variant expression extended to tumors that concurrently displayed *AR* gene amplification, we performed AR 3'RACE/SMRT-IsoSeq with 9 additional tumor tissues from 6 patients in the 41-patient AR-targeted cohort (**Supplementary Fig. 7**). In tumors that harbored *AR* gene rearrangements concurrent with *AR* gene amplification, expression of AR-V4, AR-V7, and AR-V9 was frequent, and these AR-V species were often co-expressed. Additional novel AR-V species were identified, but their fractional contributions to overall AR expression were similar to or less than AR-V7, AR-V4, and AR-V9. An exception was a novel AR-V transcript that contained contiguously-spliced AR exons 1-3 and a novel 3' terminal exon that contained cryptic exon 1 (CE1) but did not utilize annotated CE1 splice sites. This AR-V transcript was the most

abundant AR-V species in a tumor from patient V5246, accounting for approximately 10% of overall AR expression (**Supplementary Fig. 7**).

It is challenging to reconstruct the architectures of complex rearranged and amplified genomic loci such as *AR* in heterogeneous tumors from short-read DNA-seq data. Therefore, we were unable to predict whether and how the AR expression patterns observed by AR 3'RACE/SMRT-IsoSeq could be explained by underlying *AR* gene structures in these patients with amplified and rearranged *AR* (**Supplementary Fig. 7**). Adding further complexity, whole genome DNA-seq data from the 101-patient cohort indicated that these amplified and rearranged *AR* gene architectures always occurred in the context of amplification of the upstream *AR* enhancer (**Fig. 1B**). Given that AR mRNA levels are highest in tumors harboring amplification of the *AR* gene body and upstream enhancer (16), we asked whether the presence of *AR* gene rearrangements concurrent with *AR* gene body and enhancer amplification affected AR expression levels. In the 101-patient whole genome DNA-seq cohort, AR expression was higher in tumors that harbored *AR* gene rearrangements and amplification compared with tumors that harbored *AR* amplification alone (**Supplementary Fig. 8**). Thus, tumors harboring complex amplified and rearranged *AR* gene structures express a diversity of AR mRNA species, and also express high levels of AR overall.

Prostate Cancer Cells with Engineered *AR* Gene Rearrangements Express AR Variants that Drive Resistance to AR-Targeted Therapies

To further refine the mechanisms by which *AR* gene rearrangements contribute to prostate cancer progression, we focused on patient C9_a. A metastatic tumor in patient C9_a harbored a chromosomal translocation fusing *AR* intron 3 to a region of chromosome 11, which provided a clonal growth advantage when modeled with CRISPR/Cas9 (**Figs. 3C-F**). We performed single cell cloning and PCR screening of CRISPR/Cas9-transfected R1-AD1 cells to isolate a cell line clone that harbored this *AR*-chr11 translocation event (**Figs. 5A-C**). Western blot analysis of this clone (named R1-X-11) demonstrated a loss of full-length AR protein expression, and gained expression of a truncated AR protein species (**Fig. 5D**). Analysis of AR mRNA transcripts expressed in R1-X-11 cells using AR 3'RACE/PacBio (**Figs. 5E and F**) and RT-PCR (**Figs. 5G and H**) demonstrated utilization of identical splice junctions and expression of the identical AR-V fusion transcript that was discovered in patient C9_a. Two separate siRNAs designed to target the fusion exon f89 from chromosome 11 knocked down expression of the truncated AR protein species in R1-X-11 cells, confirming this truncated protein was a novel AR-V originating from translation of an AR/chromosome 11 fusion transcript (**Fig. 5I**).

R1-X-11 cells displayed enzalutamide-resistant growth, whereas the parental R1-AD1 cell line was enzalutamide-sensitive (**Fig. 6A**). Further, expression of AR target genes *FASN*, *FKBP5*, and *ABCC4* were insensitive to enzalutamide in R1-X-11 cells, but inhibited by enzalutamide in R1-AD1 cells (**Fig. 6B**). To explore the mechanism for enzalutamide-insensitive cell growth and expression of AR target genes in R1-X-11 cells, we investigated whether the novel truncated AR-V protein was functioning as a transcription factor. In R1-AD1 cells grown in androgen-replete medium, binding of full-length AR to chromatin was inhibited by enzalutamide (**Fig. 6C**),

which is a known mechanism of enzalutamide action (34). However, chromatin binding of the AR-V species expressed in R1-X-11 cells was insensitive to enzalutamide (**Fig. 6C**). Knock-down of the AR-V species in R1-X-11 cells using 2 separate siRNAs targeting the fusion exon f89 from chromosome 11 inhibited expression of *FASN*, *FKBP5*, and *ABCC4*, confirming transcriptional activity (**Fig. 6D**). Importantly, growth of R1-X-11 cells was reduced by the siRNAs targeting this AR-V, whereas growth of parental R1-AD1 cells was unaffected by these siRNAs (**Fig. 6E**). These data from a genetically engineered cell line model demonstrate a causal role for a translocation fusing *AR* intron 3 to chromosome 11 in a metastatic tumor, and an enzalutamide-resistant growth phenotype driven by an enzalutamide-insensitive AR-V transcription factor lacking the AR ligand binding domain.

Discussion

This investigation of an aggregate cohort of 272 prostate cancer patients has demonstrated that *AR* gene rearrangements are a common feature of CRPC-stage tumors. Compared to other genes, *AR* is among the most frequently rearranged genes in the CRPC genome, with *AR* gene rearrangements detectable in 23% of CRPC patients analyzed by whole genome DNA-seq and 39% of CRPC patients analyzed by deep *AR*-targeted DNA-seq. The highest frequencies of *AR* gene rearrangements occurred in patients treated with potent second-generation AR inhibitors abiraterone and enzalutamide. *AR* gene rearrangements were not detectable in primary prostate cancer, even in primary tumors from patients that ultimately progressed to CRPC and displayed *AR* gene rearrangements in their metastatic tumors. Overall, these findings support a model wherein *AR* gene rearrangements emerge during treatment with front-line hormonal therapies, and become more frequent in prevalence with subsequent lines of AR-targeted therapy. In this regard, *AR* gene rearrangements display similar evolutionary trajectories to *AR* gene amplification and mutation events, which are well-defined mechanisms of resistance first reported in CRPC patients over 2 decades ago (35, 36).

The first *AR* gene rearrangement reported in prostate cancer was discovered in the CRPC 22Rv1 cell line, consisting of a 35kb tandem duplication encompassing AR exon 3 (37). Subsequent studies with the CWR-R1 cell line, which originated from the same patient as 22Rv1, demonstrated a 50kb deletion within AR intron 1 (19). Additional *AR* gene rearrangements were characterized in several patient-derived xenograft models, which consisted of deletions or inversions impacting the genomic segment containing AR exons 5-7 (32, 38). Investigation of a 15 patient rapid autopsy cohort documented the first *AR* gene rearrangements in CRPC tissues (15). Across these prior studies with CRPC models and specimens, as well as the current study, over 90 different *AR* gene rearrangements have been defined, consisting of deletions, duplications, inversions, and translocations, with breakpoints distributed along the length of the *AR* gene body. No specific *AR* gene rearrangement events have been recurrent in terms of rearrangement type and exact breakpoint coordinates.

This extreme heterogeneity of *AR* gene rearrangements across patients presents a challenge for clinical utility of testing for these events. This heterogeneity also poses a challenge for distinguishing whether these are driver or passenger events in prostate cancer progression. To address this, in the current study we grouped *AR* gene rearrangements into two distinct classes. The first class of *AR* gene rearrangements occurred in tumors that lacked evidence for any other *AR* gene alteration. This class of *AR* gene rearrangements occurred in 8 of the aggregate 142 CRPC patients (5.6%) analyzed in this study. Analysis of representative *AR* gene rearrangements from this class using short-term CRISPR/Cas9 modeling, long-read AR isoform sequencing of tumors, and interrogation of an experimentally-derived cell line, demonstrated that these events drive high expression of tumor-unique AR variant species that are required for cell growth. The second, more frequent, class of *AR* gene rearrangements were concurrent with *AR* gene amplification, and were detectable in 31 of the aggregate 142 CRPC patients (21.8%) analyzed in this study. The exact architectures of these *AR* gene rearrangements, and the number of *AR* gene copies affected by them, is difficult to decipher due to challenges

inherent in inferring structures of amplified and rearranged loci in heterogeneous tumors from short-read DNA seq data. Nevertheless, analysis of these tumors using long-read AR isoform sequencing demonstrated a diversity of AR isoforms expressed in these tumors, including frequent detection of AR-V7. Additionally, the AR levels in these tumors with amplification and rearrangement of the AR gene were the highest among all other groups studied. Understanding whether causal relationships underlie these observations will require a more precise knowledge of the structures of amplified/rearranged *AR* gene architectures than can be ascertained from short-read DNA-seq data.

The finding of novel tumor-specific AR-Vs expressed as major AR species exclusively in CRPC tumors contrasts with the broad expression of AR-V7 as a minor AR species that is observed across primary prostate cancer as well as CRPC patients (26, 39, 40). Detection of AR-V7 mRNA or protein in patient circulating tumor cells correlates with clinical variables such as overall survival (41, 42). Therefore, a positive signal for AR-V7 in circulating tumor cells is being used as a predictive biomarker of resistance to AR-targeted therapies in CRPC patients (43). Our long-read AR isoform analysis of tumors that harbored *AR* gene rearrangements but no other identifiable AR gene alterations revealed that most were AR-V7-negative, but positive for high expression of novel tumor-specific AR-Vs. The unique 3'ends of these AR-V mRNA species, and the unique COOH-termini of their corresponding proteins, would not be detectable by the primer sets used to detect the AR exon3/CE3 splice junction, or antibodies used to detect the novel C-terminus of AR-V7 protein (**Fig. 4J**). As a result, current clinical assays for AR-V7 detection will miss these tumors where resistance to AR-targeted therapies is driven by the alternative AR-Vs that arise from underlying *AR* gene rearrangements. Given that the breakpoints from AR gene rearrangements can be detected in circulating tumor DNA from plasma of CRPC patients (44, 45), prospective studies are warranted to test the utility of *AR* gene rearrangements as biomarkers, especially for cases where no other AR alterations are detectable. In more complex cases where *AR* gene rearrangements and amplification are concurrent, knowledge that these tumors express a diversity of AR species and also have the highest levels of AR expression overall could be important for biomarker development and clinical trial design.

In summary, our study has documented widespread *AR* gene rearrangements in a large cohort of CRPC-stage tumors, and elucidated important roles for diverse *AR* gene rearrangement events as drivers of AR-V-dependence and resistance to AR-targeted therapies. Prospective monitoring of patients for the presence or emergence of AR gene rearrangements is likely to have clinical utility.

Acknowledgements

We thank the patients and their families. We acknowledge Celestia Higano, Evan Yu, Elahe Mostaghel, Heather Cheng, Pete Nelson, Bruce Montgomery, Mike Schweizer, Daniel Lin, Eva Corey, Funda Vakar-Lopez, Lawrence True, and the rapid autopsy teams for their contributions to the University of Washington Medical Center Prostate Cancer Donor Rapid Autopsy Program.

References

1. Attard G, Richards J, de Bono JS. New strategies in metastatic prostate cancer: targeting the androgen receptor signaling pathway. Clin Cancer Res 2011;17:1649-57.
2. Ryan CJ, Tindall DJ. Androgen receptor rediscovered: the new biology and targeting the androgen receptor therapeutically. J Clin Oncol 2011;29:3651-8.
3. Logothetis CJ, Gallick GE, Maity SN, Kim J, Aparicio A, Efstathiou E, et al. Molecular classification of prostate cancer progression: foundation for marker-driven treatment of prostate cancer. Cancer discovery 2013;3:849-61.
4. Watson PA, Arora VK, Sawyers CL. Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. Nat Rev Cancer 2015;15:701-11.
5. Guo Z, Yang X, Sun F, Jiang R, Linn DE, Chen H, et al. A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. Cancer Res 2009;69:2305-13.
6. Hu R, Dunn TA, Wei S, Isharwal S, Veltri RW, Humphreys E, et al. Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. Cancer Res 2009;69:16-22.
7. Cato L, de Tribolet-Hardy J, Lee I, Rottenberg JT, Coleman I, Melchers D, et al. ARv7 Represses Tumor-Suppressor Genes in Castration-Resistant Prostate Cancer. Cancer Cell 2019;35:401-13 e6.
8. Watson PA, Chen YF, Balbas MD, Wongvipat J, Socci ND, Viale A, et al. Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. Proc Natl Acad Sci U S A 2010;107:16759-65.
9. Xu D, Zhan Y, Qi Y, Cao B, Bai S, Xu W, et al. Androgen Receptor Splice Variants Dimerize to Transactivate Target Genes. Cancer Res 2015;75:3663-71.
10. de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and increased survival in metastatic prostate cancer. N Engl J Med 2011;364:1995-2005.
11. Fizazi K, Shore N, Tammela TL, Ulys A, Vjaters E, Polyakov S, et al. Darolutamide in Nonmetastatic, Castration-Resistant Prostate Cancer. N Engl J Med 2019;380:1235-46.
12. Scher HI, Fizazi K, Saad F, Taplin ME, Sternberg CN, Miller MD, et al. Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. N Engl J Med 2012;367:1187-97.
13. Smith MR, Saad F, Chowdhury S, Oudard S, Hadaschik BA, Graff JN, et al. Apalutamide Treatment and Metastasis-free Survival in Prostate Cancer. N Engl J Med 2018;378:1408-18.
14. Scher HI, Morris MJ, Basch E, Heller G. End points and outcomes in castration-resistant prostate cancer: from clinical trials to clinical practice. J Clin Oncol 2011;29:3695-704.
15. Henzler C, Li Y, Yang R, McBride T, Ho Y, Sprenger C, et al. Truncation and constitutive activation of the androgen receptor by diverse genomic rearrangements in prostate cancer. Nature communications 2016;7:13668.
16. Quigley DA, Dang HX, Zhao SG, Lloyd P, Aggarwal R, Alumkal JJ, et al. Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. Cell 2018;174:758-69 e9.
17. Fraser M, Sabelnykova VY, Yamaguchi TN, Heisler LE, Livingstone J, Huang V, et al. Genomic hallmarks of localized, non-indolent prostate cancer. Nature 2017;541:359-64.

18. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol 2011;29:24-6.
19. Li Y, Hwang TH, Oseth L, Hauge A, Vessella RL, Schmechel SC, et al. AR intragenic deletions linked to androgen receptor splice variant expression and activity in models of prostate cancer progression. Oncogene 2012;31:4759-67.
20. Li Y, Chan SC, Brand LJ, Hwang TH, Silverstein KA, Dehm SM. Androgen receptor splice variants mediate enzalutamide resistance in castration-resistant prostate cancer cell lines. Cancer Res 2013;73:483-9.
21. Dalal K, Che M, Que NS, Sharma A, Yang R, Lallous N, et al. Bypassing Drug Resistance Mechanisms of Prostate Cancer with Small Molecules that Target Androgen Receptor-Chromatin Interactions. Mol Cancer Ther 2017;16:2281-91.
22. 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;68:5469-77.
23. Kohli M, Ho Y, Hillman DW, Van Etten JL, Henzler C, Yang R, et al. Androgen receptor variant AR-V9 is co-expressed with AR-V7 in prostate cancer metastases and predicts abiraterone resistance. Clin Cancer Res 2017;23:4704-15.
24. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094-100.
25. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644-8.
26. Abida W, Cyrta J, Heller G, Prandi D, Armenia J, Coleman I, et al. Genomic correlates of clinical outcome in advanced prostate cancer. Proc Natl Acad Sci U S A 2019;116:11428-36.
27. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med 2016.
28. Kumar A, Coleman I, Morrissey C, Zhang X, True LD, Gulati R, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. Nat Med 2016;22:369-78.
29. Goodwin JF, Schiewer MJ, Dean JL, Schrecengost RS, de Leeuw R, Han S, et al. A hormone-DNA repair circuit governs the response to genotoxic insult. Cancer discovery 2013;3:1254-71.
30. Polkinghorn WR, Parker JS, Lee MX, Kass EM, Spratt DE, Iaquinta PJ, et al. Androgen receptor signaling regulates DNA repair in prostate cancers. Cancer discovery 2013;3:1245-53.
31. Schiewer MJ, Knudsen KE. Linking DNA Damage and Hormone Signaling Pathways in Cancer. Trends Endocrinol Metab 2016;27:216-25.
32. Nyquist MD, Li Y, Hwang TH, Manlove LS, Vessella RL, Silverstein KA, et al. TALEN-engineered AR gene rearrangements reveal endocrine uncoupling of androgen receptor in prostate cancer. Proc Natl Acad Sci U S A 2013;110:17492-7.
33. Sun S, Sprenger CC, Vessella RL, Haugk K, Soriano K, Mostaghel EA, et al. Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. J Clin Invest 2010;120:2715-30.

34. Clegg NJ, Wongvipat J, Joseph JD, Tran C, Ouk S, Dilhas A, et al. ARN-509: a novel antiandrogen for prostate cancer treatment. Cancer Res 2012;72:1494-503.
35. Taplin ME, Bubley GJ, Shuster TD, Frantz ME, Spooner AE, Ogata GK, et al. Mutation of the androgen-receptor gene in metastatic androgen-independent prostate cancer. N Engl J Med 1995;332:1393-8.
36. Visakorpi T, Hyytinen E, Koivisto P, Tanner M, Keinanen R, Palmberg C, et al. In vivo amplification of the androgen receptor gene and progression of human prostate cancer. Nat Genet 1995;9:401-6.
37. Li Y, Alsagabi M, Fan D, Bova GS, Tewfik AH, Dehm SM. Intragenic rearrangement and altered RNA splicing of the androgen receptor in a cell-based model of prostate cancer progression. Cancer Res 2011;71:2108-17.
38. Lawrence MG, Obinata D, Sandhu S, Selth LA, Wong SQ, Porter LH, et al. Patient-derived Models of Abiraterone- and Enzalutamide-resistant Prostate Cancer Reveal Sensitivity to Ribosome-directed Therapy. Eur Urol 2018;74:562-72.
39. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. Cell 2015;161:1215-28.
40. Sharp A, Coleman I, Yuan W, Sprenger C, Dolling D, Nava Rodrigues D, et al. Androgen receptor splice variant-7 expression emerges with castration resistance in prostate cancer. J Clin Invest 2018.
41. Antonarakis ES, Lu C, Wang H, Luber B, Nakazawa M, Roeser JC, et al. AR-V7 and Resistance to Enzalutamide and Abiraterone in Prostate Cancer. N Engl J Med 2014;371:1028-38.
42. Scher HI, Graf RP, Schreiber NA, Jayaram A, Winquist E, McLaughlin B, et al. Assessment of the Validity of Nuclear-Localized Androgen Receptor Splice Variant 7 in Circulating Tumor Cells as a Predictive Biomarker for Castration-Resistant Prostate Cancer. JAMA oncology 2018;4:1179-86.
43. Armstrong AJ, Halabi S, Luo J, Nanus DM, Giannakakou P, Szmulewitz RZ, et al. Prospective Multicenter Validation of Androgen Receptor Splice Variant 7 and Hormone Therapy Resistance in High-Risk Castration-Resistant Prostate Cancer: The PROPHECY Study. J Clin Oncol 2019;37:1120-9.
44. Annala M, Vandekerkhove G, Khalaf D, Taavitsainen S, Beja K, Warner EW, et al. Circulating Tumor DNA Genomics Correlate with Resistance to Abiraterone and Enzalutamide in Prostate Cancer. Cancer discovery 2018;8:444-57.
45. De Laere B, van Dam PJ, Whitington T, Mayrhofer M, Diaz EH, Van den Eynden G, et al. Comprehensive Profiling of the Androgen Receptor in Liquid Biopsies from Castration-resistant Prostate Cancer Reveals Novel Intra-AR Structural Variation and Splice Variant Expression Patterns. Eur Urol 2017.

Figure Legends

Figure 1.
Diverse *AR* gene rearrangements are frequent in CRPC and detectable by whole genome DNA-seq. **A,** Schematic of the *AR* gene and classes of somatic alterations that occur in prostate cancer genomes. SNV, single nucleotide variant; CN, copy number; enh, enhancer. **B,** Frequency of *AR* gene alterations detected by whole genome DNA-seq of 130 primary prostate cancers or 101 metastatic CRPC specimens. Bars represent tumors with 1, 2, or more than 2 (3+) concurrent alterations in *AR*. Oncoprints on the right illustrate the type of *AR* alterations observed in each tumor sample. Each column represents an individual tumor. **C,** Map of *AR* gene rearrangements and breakpoint locations (triangles) within the *AR* gene body. Genome coordinates are genome build hg38. Locations of *AR* exons 1-8 are shown as black boxes. **D,** Coverage plots of DNA-seq reads in 3 tumors where an *AR* gene rearrangement was the only *AR* alteration detected by whole genome DNA-seq.

Figure 2.
AR gene rearrangements are enriched in prostate cancer patients treated with AR-targeted endocrine therapies. **A,** Frequency of *AR* gene rearrangements in the 101-patient whole genome DNA-seq (WGS) cohort based on prior exposure of patients to abiraterone (abi) or enzalutamide (enz). **B,** Oncoprint illustrating *AR* alterations occurring in *AR* gene rearrangement-positive patients based on prior abiraterone or enzalutamide exposure. **C,** Frequency of *AR* gene rearrangements in a 41-patient *AR*-targeted DNA-seq cohort based on prior exposure of patients to abi or enz. **D,** Oncoprint illustrating *AR* alterations occurring in *AR* gene rearrangement-positive patients based on prior abi or enz exposure. **E,** Map of *AR* gene rearrangements and breakpoint locations (triangles) within the AR gene body discovered in a 41-patient *AR*-targeted DNA-seq cohort. Genome coordinates are genome build hg38. Patients are labeled based on whether tumor samples were obtained by autopsy (_a), biopsy (_b), or surgery (_s). Locations of *AR* exons 1-8 are shown as black boxes. **F,** Schematic of primer designs and PCR products for detection of genomic breakpoints arising from deletions, duplications, and inversions. Ref., reference genome; rearr., rearrangement. **G-K,** PCR products from patient-matched longitudinal samples of primary prostate cancer and metastatic CRPC (mCRPC).

Figure 3.
Cells with engineered *AR* gene rearrangements have a clonal growth advantage under conditions of AR-targeted therapy. **A,** Schematic of experimental strategy for inducing and tracking clonal evolution of AR gene rearrangements using CRISPR/Cas9. Transfected cells were selected with puromycin (puro), and then cultured in medium supplemented with androgen-replete fetal bovine serum (FBS), charcoal-stripped, androgen-depleted FBS (CSS), enzalutamide (ENZ), or vehicle control (DMSO). **B,** AR gene rearrangements modeled using CRISPR/Cas9. **C,** Translocation-spanning PCR of DNA isolated from R1-AD1 cells that were transfected with gRNA and Cas9 plasmids as indicated. Quantitative PCR was used to track clonal enrichment or de-enrichment of translocations relative to the overall population of cells grown in **D,** 2-dimensional conditions or **E,** 3-dimensional (soft agar) conditions **F,** LNCaP cells

were transfected as indicated and assayed as in (D). **G-H,** Deletion 1 from patient V5300_b, **I-J,** Deletion 2 from patient V5300_b, **K-L,** Deletion 1 from patient V5301_b, and **M-N,** Deletion 2 from patient V5301_b were modeled in R1-AD1 cells and monitored for clonal enrichment by breakpoint-spanning PCR as described for panels C and E. Gray bars represent mean +/- 95% confidence interval. Individual data points from biological replicate experiments are shown as black filled circles. P-values were determined using unpaired 2-sided t-tests.

Figure 4

Tumors harboring *AR* gene rearrangements as the only detectable *AR* gene alteration display high expression of AR variant mRNAs. **A,** Schematic of the translocation between AR and chromosome 11 in patient C9_a. **B,** Exon composition and quantification of AR transcripts isolated from a translocation-positive metastasis from patient C9_a cells using 3' rapid amplification of cDNA ends (RACE) using a forward primer anchored in AR exon 1. Individual pixels represent discrete exons contained in individual AR transcripts. Pixel colors indicate whether that exon was spliced via annotated splice sites at the 5' and/or 3' ends of known exons. Read counts represent the number of single molecule transcripts that matched the indicated splicing pattern. AR transcripts were inspected manually for predicted translation, and annotated based on a previous nomenclature system. AR transcripts that had not been identified previously were classified as novel. **C,** Schematic of deletions in patient V5300_b and **D,** quantification of transcripts expressed in this metastasis as described in (B). **E,** Schematic of deletions in patient V5301_b and **F,** quantification of transcripts expressed in this metastasis as described in (B). **G,** Western blot of lysates from R1-AD1 cells transfected with gRNA and Cas9 plasmids as indicated and cultured in medium supplemented with the AR antagonist enzalutamide (ENZ), or vehicle control (DMSO) for 7 or 14 days as indicated. AR expression was determined using a pan-AR antibody that recognizes the AR N-terminal domain. Actin is a loading control. **H,** Schematic of deletions in patient V4002_b and **I,** quantification of transcripts expressed in this metastasis as described in (B). **J,** Comparison of AR-V7 mRNA 3' terminus and AR-V7 protein C-terminus with AR mRNA variants discovered in prostate cancer metastases harboring AR gene rearrangements.

Figure 5.

An *AR*-chromosome 11 translocation causes expression of a tumor-specific AR variant. **A,** CRISPR/Cas9 engineering strategy to generate an *AR*-chr11 translocation that models patient C9_a. Locations of PCR primers used for screening single cell clones are indicated. **B,** PCR-based characterization of parental R1-AD1 cells and a single cell clone (R1-X-11) derived by CRISPR/Cas9 engineering. **C,** Sanger sequencing of the PCR product from (B). **D,** Western blot of lysates from indicated cell lines. AR expression was determined using a pan-AR antibody that recognizes the AR N-terminal domain. Actin is a loading control. **E,** Schematic of the translocation between *AR* and chromosome 11, and location of the novel fusion exon expressed in patient C9_a. **F,** Exon composition and quantification of AR transcripts isolated from R1-X-11 cells using 3' rapid amplification of cDNA ends (RACE) with a forward primer anchored in AR exon 1. Individual pixels represent discrete exons contained in individual AR transcripts. Pixel colors indicate whether that exon was spliced via annotated splice sites at the 5' and/or 3' ends of known exons. Read counts represent the number of single molecule

transcripts that matched the indicated splicing pattern. AR transcripts were inspected manually for predicted translation, and annotated based on a previous nomenclature system. AR transcripts that had not been identified previously were classified as novel. **G,** Schematic of primers and siRNAs designed to study the C9a-AR-Vfusion1 transcript. **H,** RT-PCR with RNA isolated from indicated tumor and cell lines using primers illustrated in (G). **I,** Indicated cell lines were transfected with control (CTRL) siRNA or siRNAs illustrated in (G). Lysates were analyzed by Western blot as in (D).

Figure 6.

A tumor-specific AR variant caused by *AR*-chromosome 11 translocation drives enzalutamide resistance. **A,** Growth of R1-AD1 and R1-X-11 cells was analyzed in culture medium containing enzalutamide (30μM) or DMSO as vehicle control. Bold black and red lines are mean +/- 95% CI from 3 independent experiments performed in biological quadruplicate (n=12). Light gray and red lines are the individual replicates. Significance was tested using unpaired 2-sided t-tests. **B,** Expression of AR target genes (*FASN*, *FKBP5*, and *ABCC4*) was tested by RT-PCR in R1-AD1 and R1-X-11 cells grown in culture medium containing enzalutamide (ENZ) or vehicle (DMSO) for 24h as in (A). Data are mean +/- 95% CI from 3 independent experiments performed in technical triplicate (n=9). Individual data points are shown. P-values were determined using unpaired 2-sided t-tests. **C,** Western blot of insoluble nuclear (chromatin) extracts from indicated cell lines. AR expression was determined using a pan-AR antibody that recognizes the AR N-terminal domain. Histone H3 is a loading control. **D,** R1-X-11 cells were transfected with control (CTRL) siRNA or siRNAs targeting exon f89 of the C9-a AR-Vfusion1 transcript and expression of AR target genes was tested by RT-PCR as in (B). Data are mean +/- 95% CI from 3 independent experiments performed in technical duplicate (n=6). Significance was tested using unpaired 2-sided t-tests. **E,** R1-AD1 and R1-X-11 cell lines were transfected with siRNAs as in (D) and growth was analyzed as in (A). Data are mean +/- 95% CI from 3 independent experiments performed in biological quadruplicate (n=12). P-values were determined using unpaired 2-sided t-tests.

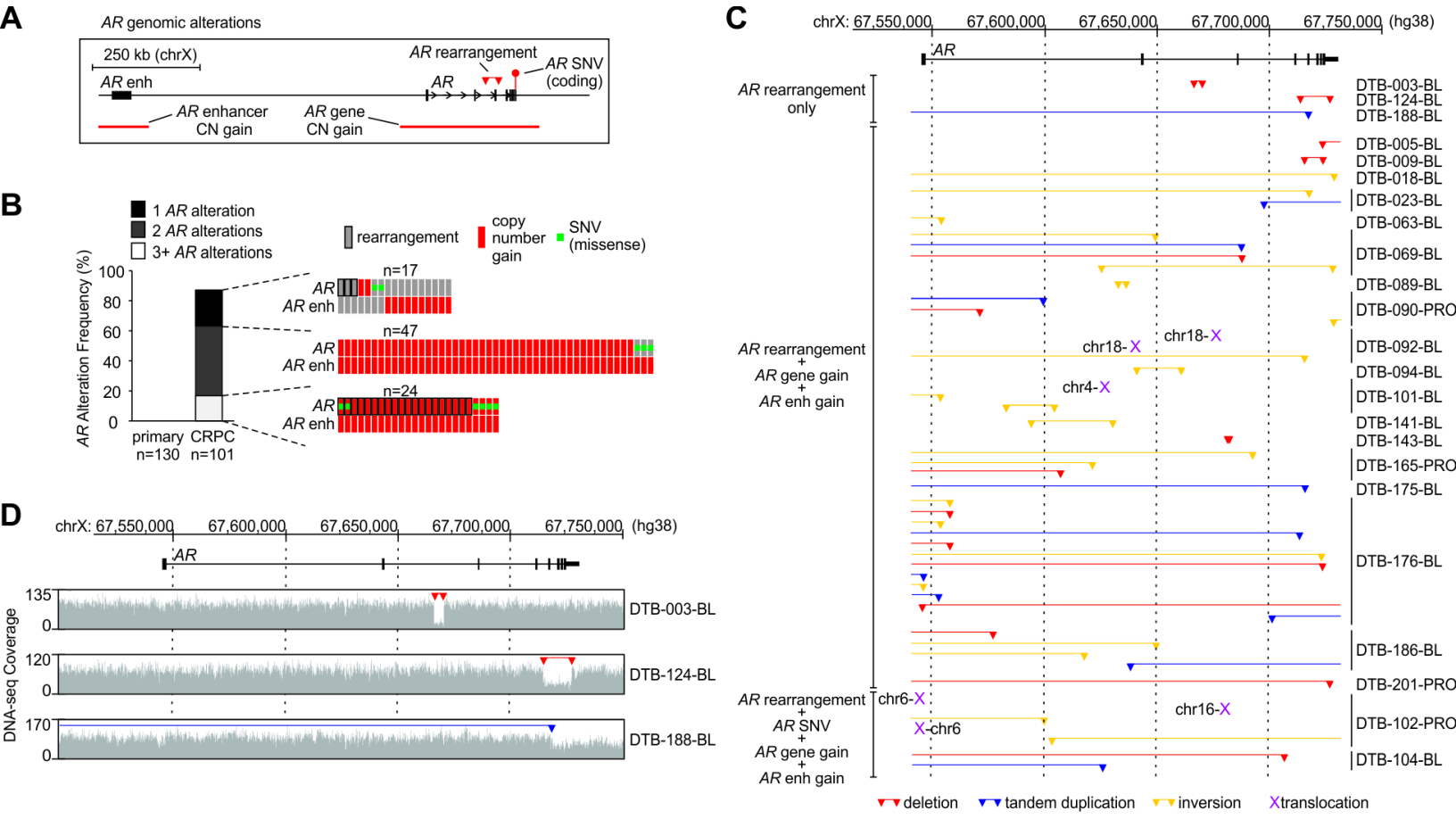
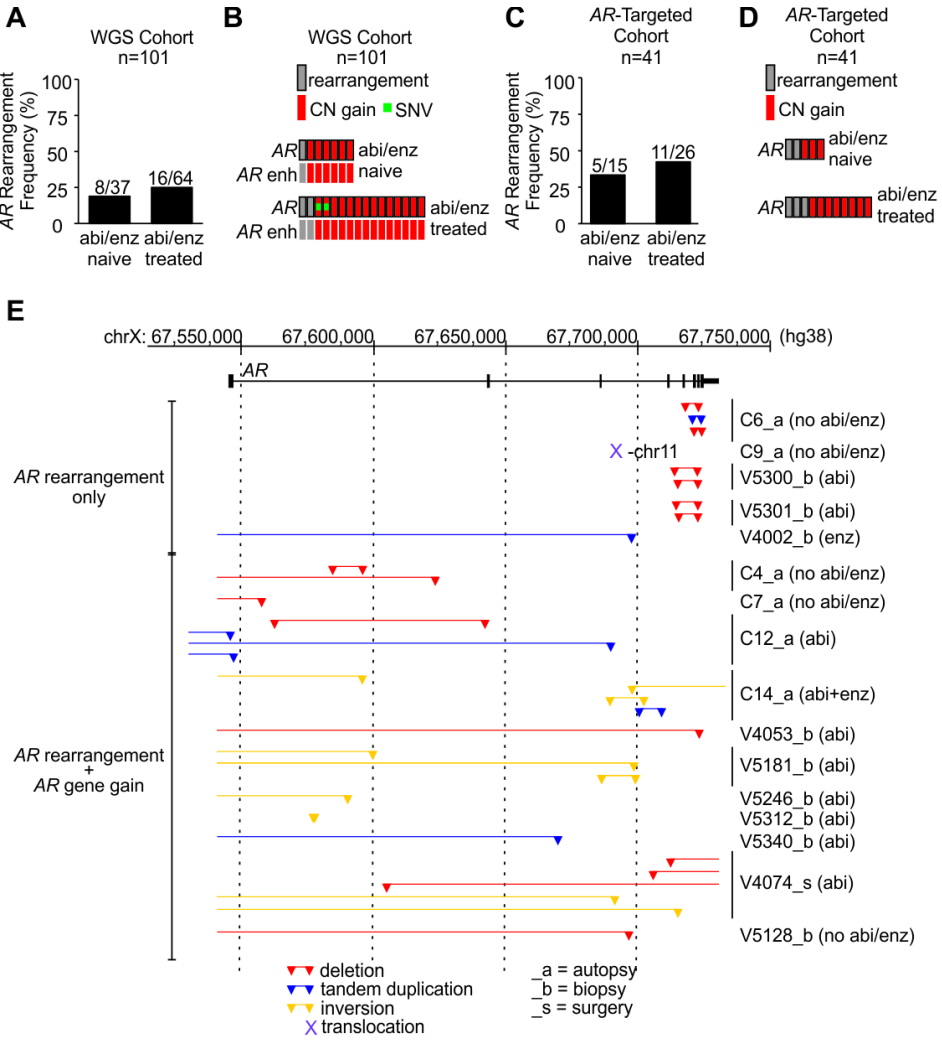


FIGURE 1

Author Manuscript Published OnlineFirst on January 13, 2020; DOI: 10.1158/1078-0432.CCR-19-3023
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.



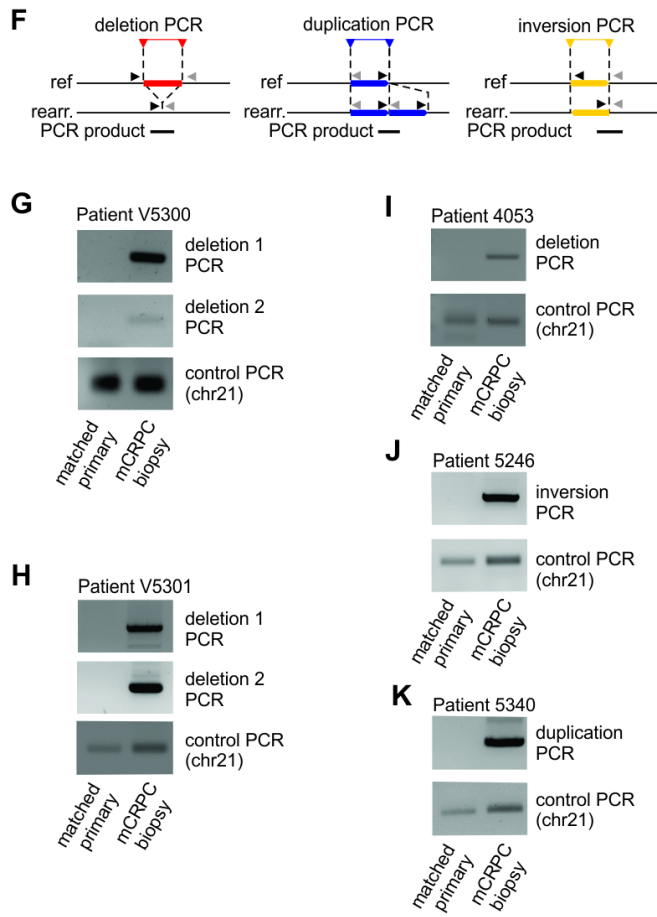


FIGURE 2

Author Manuscript Published OnlineFirst on January 13, 2020; DOI: 10.1158/1078-0432.CCR-19-3023
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

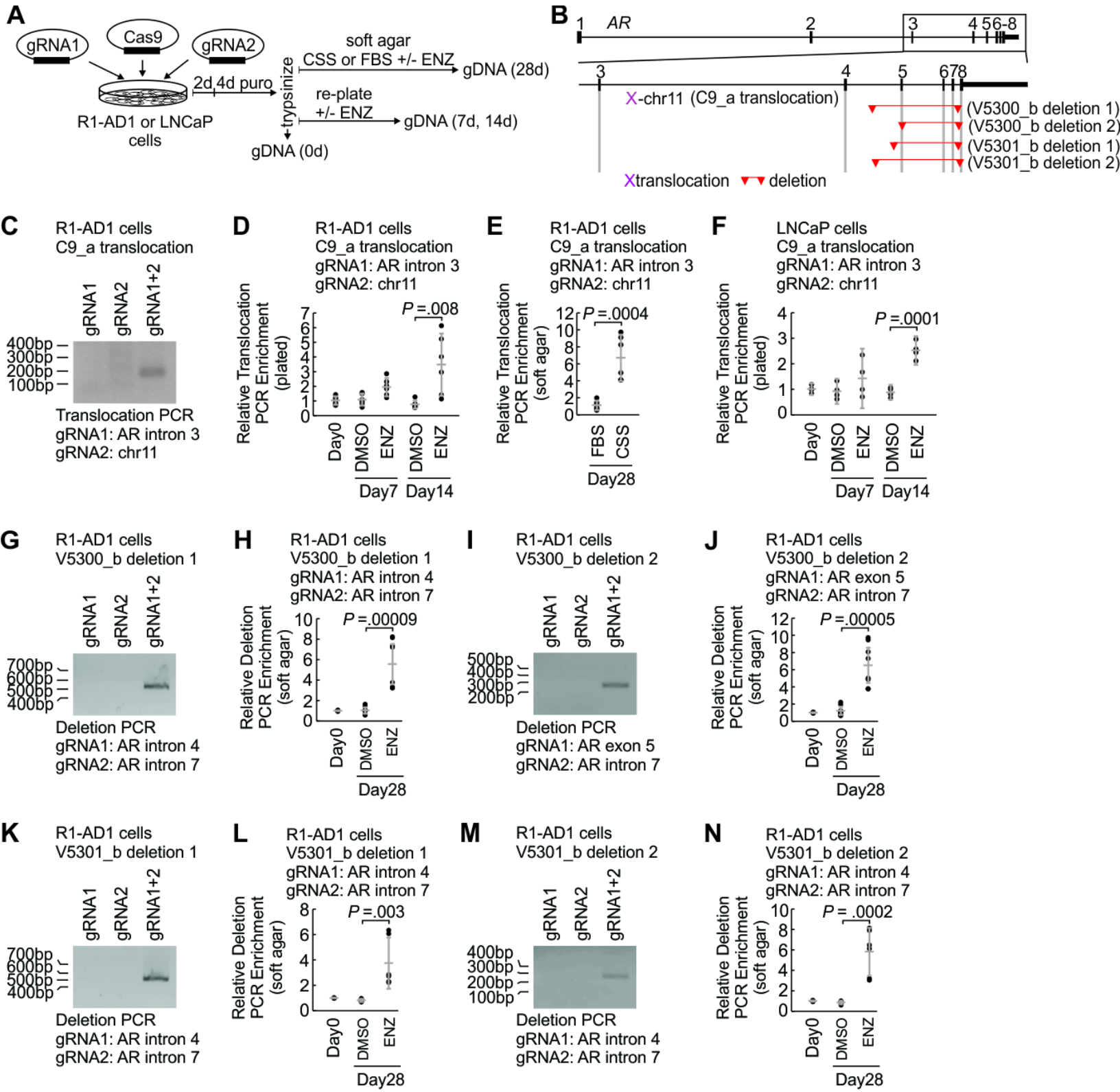


FIGURE 3

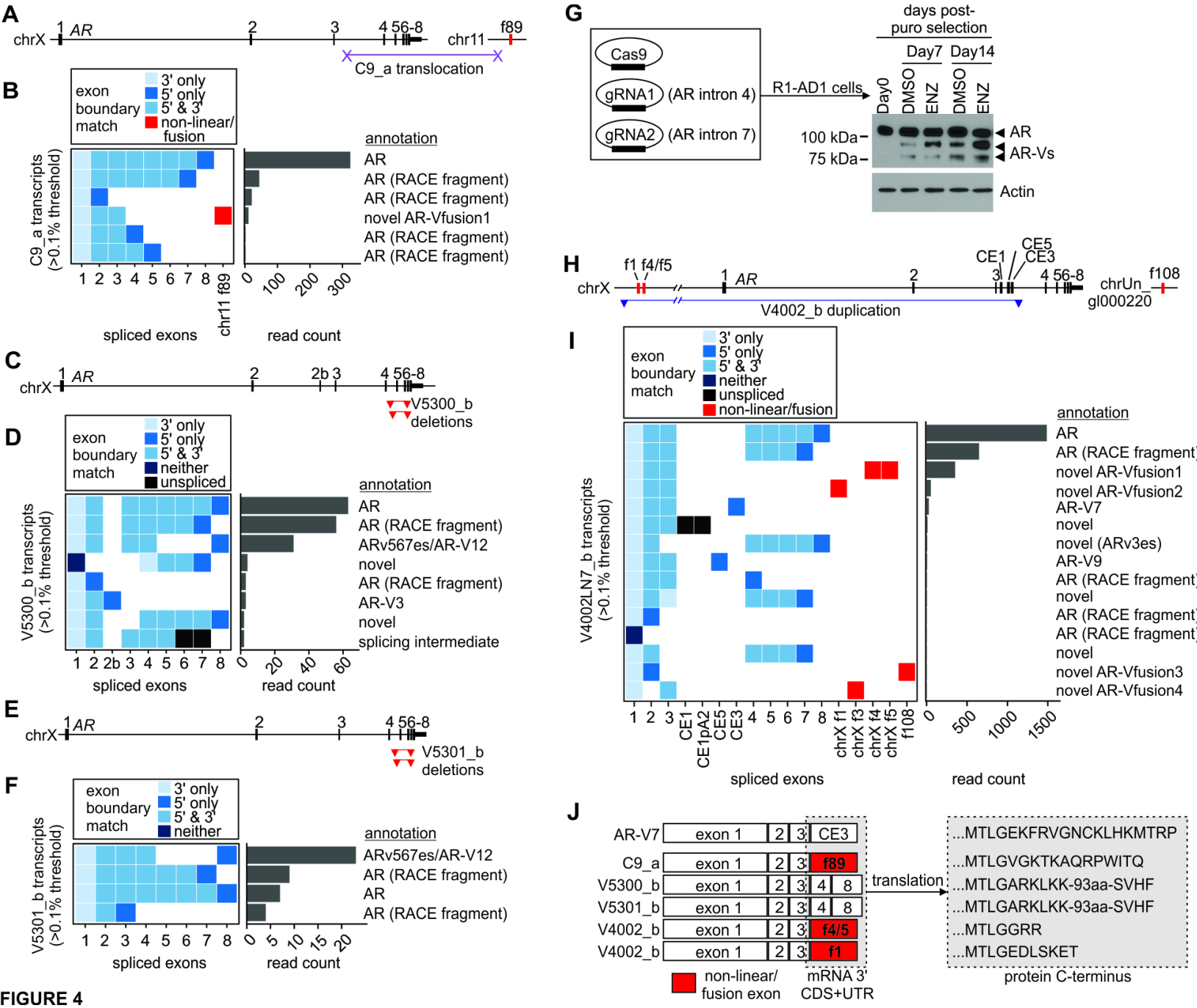


FIGURE 4

Author Manuscript Published OnlineFirst on January 13, 2020; DOI: 10.1158/1078-0432.CCR-19-3023
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

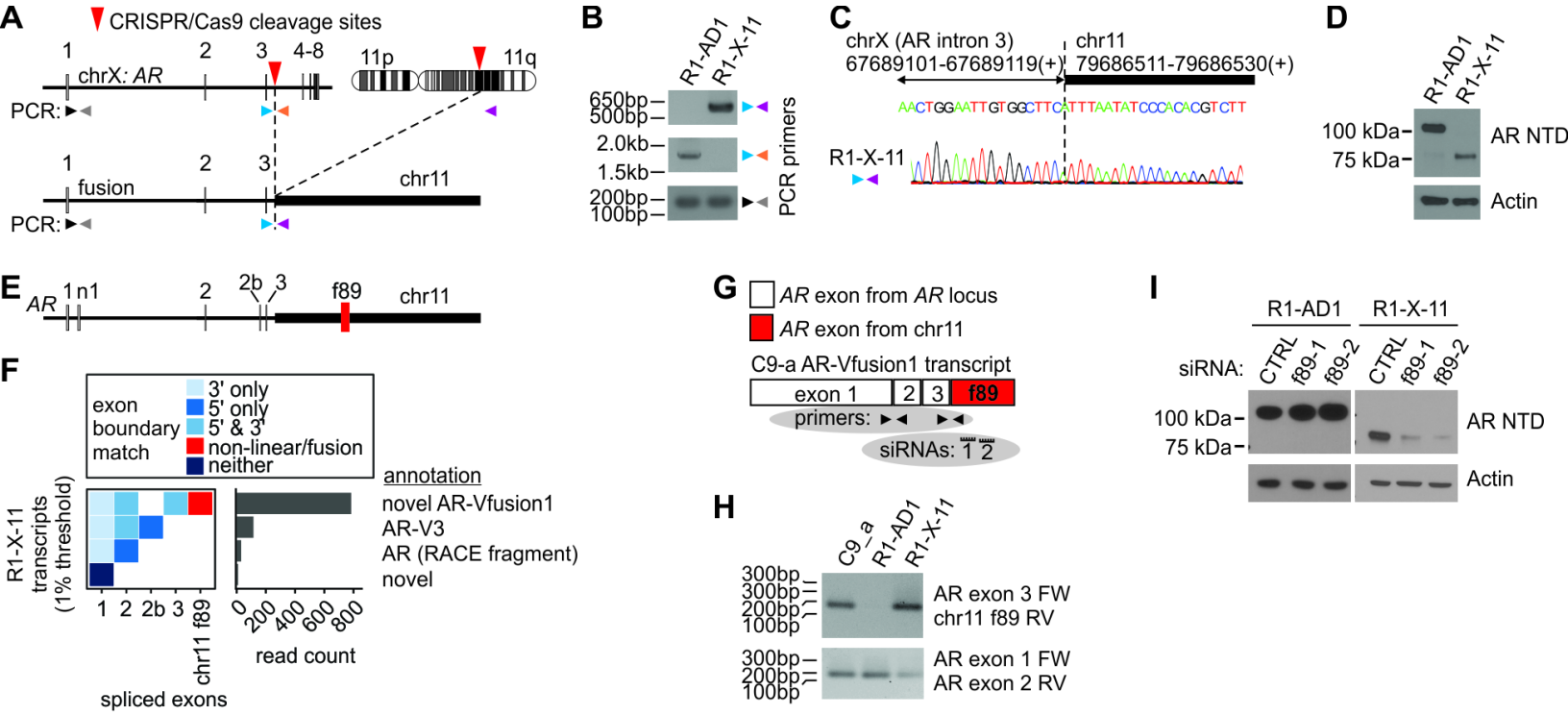


FIGURE 5

Author Manuscript Published OnlineFirst on January 13, 2020; DOI: 10.1158/1078-0432.CCR-19-3023
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

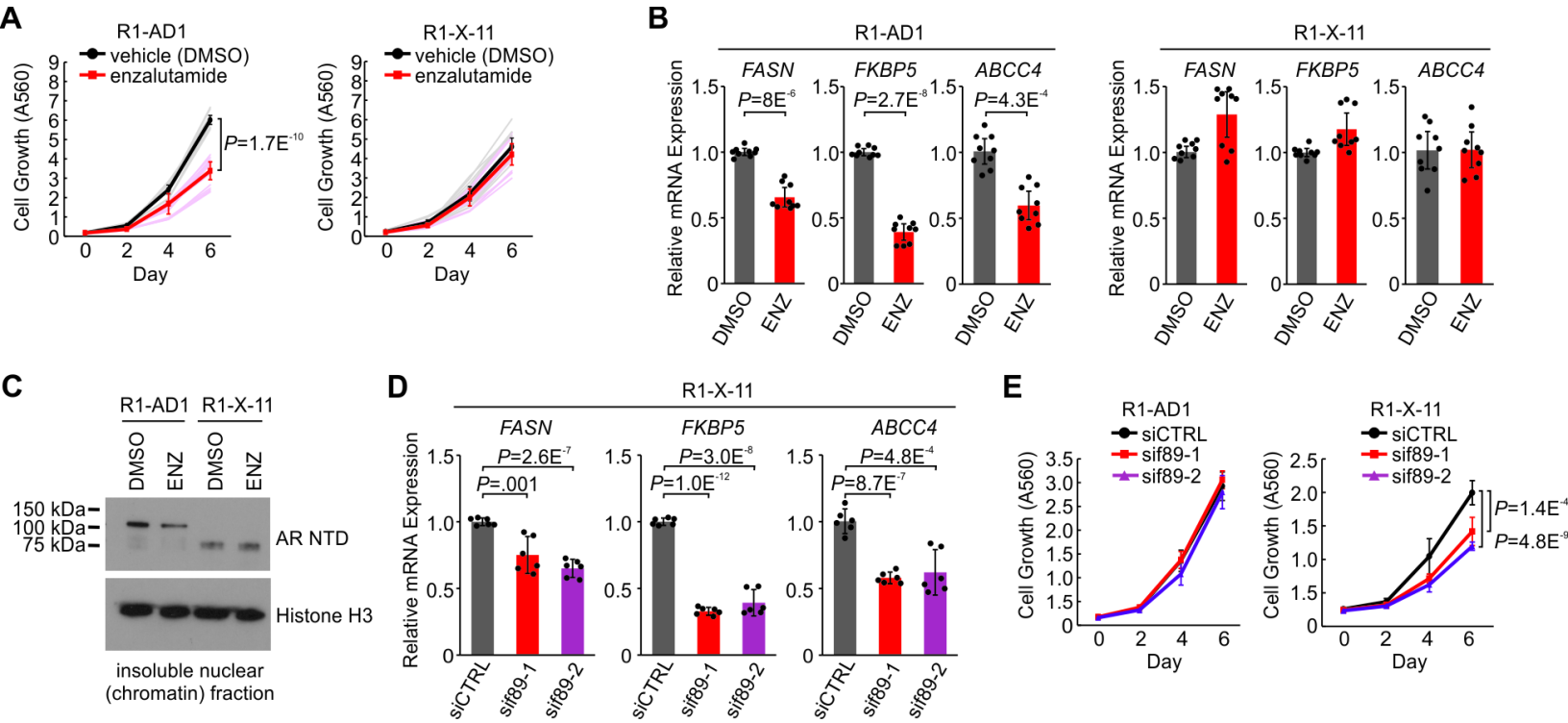


FIGURE 6

AACR American Association for Cancer Research

Clinical Cancer Research

Diverse AR Gene Rearrangements Mediate Resistance to Androgen Receptor Inhibitors in Metastatic Prostate Cancer

Yingming Li, Rendong Yang, Christine M Henzler, et al.

Clin Cancer Res Published OnlineFirst January 13, 2020.