9780–9796 *Nucleic Acids Research, 2022, Vol. 50, No. 17* *Published online 31 August 2022*
https://doi.org/10.1093/nar/gkac737

The androgen receptor couples promoter recruitment of RNA processing factors to regulation of alternative polyadenylation at the 3' end of transcripts

Cinzia Caggiano[1,2], Marco Pieraccioli[1,2], Consuelo Pitolli[1,2], Gabriele Babini[2], Dinghai Zheng[3], Bin Tian[4], Pamela Bielli[5,6,*] and Claudio Sette[1,2,*]

[1]Department of Neuroscience, Section of Human Anatomy, Catholic University of the Sacred Heart, Rome 00168, Italy, [2]IRCCS Fondazione Policlinico A. Gemelli, Rome 00168, Italy, [3]Department of Microbiology, Biochemistry and Molecular Genetics, Rutgers New Jersey Medical School, Newark, NJ 07103, USA, [4]Gene Expression and Regulation Program, The Wistar Institute, Philadelphia, PA 19104, USA, [5]Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome 00133, Italy and [6]IRCCS Fondazione Santa Lucia, Rome 00143, Italy

Received December 05, 2021; Revised July 20, 2022; Editorial Decision August 11, 2022; Accepted August 29, 2022

ABSTRACT

Prostate cancer (PC) relies on androgen receptor (AR) signaling. While hormonal therapy (HT) is efficacious, most patients evolve to an incurable castration-resistant stage (CRPC). To date, most proposed mechanisms of acquired resistance to HT have focused on AR transcriptional activity. Herein, we uncover a new role for the AR in alternative cleavage and polyadenylation (APA). Inhibition of the AR by Enzalutamide globally regulates APA in PC cells, with specific enrichment in genes related to transcription and DNA topology, suggesting their involvement in transcriptome reprogramming. AR inhibition selects promoter-distal polyadenylation sites (pAs) enriched in *cis*-elements recognized by the cleavage and polyadenylation specificity factor (CPSF) complex. Conversely, promoter-proximal intronic pAs relying on the cleavage stimulation factor (CSTF) complex are repressed. Mechanistically, Enzalutamide induces rearrangement of APA subcomplexes and impairs the interaction between CPSF and CSTF. AR inhibition also induces co-transcriptional CPSF recruitment to gene promoters, predisposing the selection of pAs depending on this complex. Importantly, the scaffold CPSF160 protein is up-regulated in CRPC cells and its depletion represses HT-induced APA patterns. These findings uncover an unexpected role for the AR in APA regulation and suggest that APA-mediated transcriptome reprogramming represents an adaptive response of PC cells to HT.

INTRODUCTION

Prostate cancer (PC) is the second most frequent malignancy and the fifth most common cause of cancer-related death in men (1). It originates as prostatic intraepithelial neoplasia (PIN) and progresses to localized, invasive advanced and, eventually, metastatic disease (1,2). Throughout its evolution, PC remains highly dependent on the presence of androgens, which support growth and oncogenic features of tumor cells. Chemical castration by androgen-deprivation therapy (ADT) represents the standard treatment for PC patients and initially causes disease regression (1). However, the tumor generally develops mechanisms of resistance and evolves to castration-resistant prostate cancer (CRPC) and metastatic PC, which represents the leading cause of patients' death (2–4). Importantly, most CRPC cells still require androgen receptor (AR) signaling, but they acquire the ability to withstand ADT conditions (1–4). New-generation anti-hormonal drugs, such as the AR inhibitor Enzalutamide (5), were shown to significantly improve survival of CRPC patients (2–4). Nevertheless, these drugs do not represent a cure in the medium to long term, and the prognosis of CRPC patients remains poor (4). Thus, new insights into the routes involved in acquisition of resistance to hormonal therapy (HT) are urgently needed, as they may lead to development of more efficacious treatments.

The multifactorial mechanisms involved in acquired resistance to HT are partially known. For instance, autocrine production of androgens, *AR* gene amplification, androgen-

*To whom correspondence should be addressed. Email: claudio.sette@unicatt.it
Correspondence may also be addressed to Pamela Bielli. Email: pamela.bielli@uniroma2.it

independent activation of the AR and oncogenic gene translocations, such as the *TMPRSS2–ERG* fusion, conform to the CRPC phenotype (1). More recently, aberrant regulation of alternative splicing and alternative polyadenylation (APA) of pre-mRNAs have emerged as key factors in PC evolution and in acquisition of HT resistance (6–8). For instance, changes in APA can yield pro-oncogenic variants of the *CCND1* (cyclin D1) and *AR* genes (8). *CCND1* codes for cyclin D1a, which includes all five exons, and the cyclin D1b variant generated by usage of a cryptic polyadenylation site (pA) in intron 4 of the gene (9). Up-regulation of oncogenic splicing factors, such as SRSF1 and Sam68 (10,11), enhances recognition of the intronic pA and promotes expression of cyclin D1b, which causes aberrant regulation of AR activity (9). Even more importantly, activation of a cryptic pA in the *AR* intron 3 leads to expression of constitutively active AR variants (AR-Vs) (12), which lack the ligand-binding domain (8). Expression of AR-Vs is favored by *AR* genomic rearrangements, contributes to HT resistance and is associated with worse clinical outcome in patients (13–15). Direct inhibition of the selection of this intronic pA was sufficient to block androgen-independent growth of CRPC cells (12), indicating that modulation of APA can be exploited therapeutically. Interestingly, a similar inhibition of AR-Vs expression and of the CRPC phenotype was observed in cells depleted of CPSF160 (12), a component of the cleavage and polyadenylation specificity factor (CPSF) complex that enhances recruitment of the cleavage and polyadenylation (C/P) complex to the pA and promotes 3'-end cleavage (16). These selected examples highlight the oncogenic role played by dysregulation of APA in PC and the impact exerted by this process on the acquisition of a CRPC phenotype. Nevertheless, it is currently unknown to what extent APA is modulated in response to HT and how this process is regulated. Elucidating these mechanisms may pave the way for the development of new therapies targeting APA dysregulation in PC.

In this study, we have uncovered a genome-wide remodulation of APA during HT, which prevalently leads to transcript lengthening and suppression of premature intronic pA usage. APA-regulated genes are direct targets of the AR, and selection of alternative pAs occurs co-transcriptionally. HT causes rearrangements of the complexes involved in APA regulation, with increased recruitment of CPSF160 and FIP1L1 in the CPSF complex and reduced interaction of the latter with the cleavage stimulation factor (CSTF) complex. Moreover, Enzalutamide promotes the recruitment of CPSF to the promoter region of APA-regulated genes and to the selected pAs in the pre-mRNA. Collectively, our findings indicate that AR links promoter-localized events with 3' end processing events and suggest that APA-mediated transcriptome reprogramming represents an adaptive response of PC cells to HT.

MATERIALS AND METHODS

Cell culture and treatments

LNCaP and 22Rv1 cells were cultured in RPMI 1640 (Euroclone) with 4500 mg/l glucose (Sigma-Aldrich), 10 mM HEPES and 1 mM sodium pyruvate (both obtained from Euroclone). VCaP cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Euroclone) using dishes coated with Geltrex matrix (GIBCO) according to the manufacturer's instructions. Human embryo kidney (HEK) 293T cells were cultured in DMEM. LNCaP-ER cells were obtained by selection in the continuous presence of 1 μM Enzalutamide for 4 months and maintained in the presence of the drug. All cell lines were maintained in growth media supplemented with 10% fetal bovine serum (FBS; GIBCO), 100 U/ml penicillin and 100 μg/ml streptomycin (Euroclone) at 37°C in a humidified atmosphere with 5% CO_2. The cell lines were purchased from ATCC and tested periodically for mycoplasma contamination. In ADT conditions, cells were cultured in growth medium completed with either charcoal-stripped serum (CSS), instead of FBS, or Enzalutamide (MDV3100, Sigma-Aldrich) and Bicalutamide (MedChemExpress) at the indicated concentration.

Lentiviral silencing and transfection experiments

LNCaP cells were infected with pLVTHM lentiviral vectors (12247, Addgene) encoding control short hairpin (sh-scrambled; TGGTTTACATGTCGACTAATC) or shAR.1 (CAAGGGAGGTTACACCAAATT) or shAR.2 (GAAATGATTGCACTATTGATT) using MluI and ClaI as restriction sites. Lentiviral packaging was performed into HEK293T cells. The small interfering RNA (siRNA) transfection experiments were performed using Lipofectamine RNAiMax and Opti-MEM medium (Invitrogen) according to the manufacturer's instruction. CPSF160 siRNAs were purchased from Dharmacon (D-020395–01; D-020395–02) and used at 80 nM final concentration in a single transfection reaction. The siRNAs for CSTF64 and CASK were purchased from Sigma-Aldrich and their sequences are listed in Supplementary Table S1.

RNA isolation, 3'-RACE and real-time PCR

RNA was extracted using the Trizol reagent (Invitrogen, Life Technologies) and treated with RNase-free DNase (Ambion). For APA pattern and gene expression (GE), 1 μg of RNA was retrotranscribed with M-MLV reverse transcriptase (Promega) and oligo(dT) primers (Roche). Quantitative real-time polymerase chain reaction (qRT-PCR) analyses was carried out using the StepOnePlus Real-Time PCR system (Applied Biosystems) and an isoform-specific primer pair designed within 200 bp upstream of the pA, 10 ng of cDNA template and PowerUp SYBR Green Master Mix (Applied Biosystems). The APA pattern induced by Enzalutamide was calculated as the expression of distal pA (d-pA) versus proximal pA (p-pA) by the ΔΔCT method, normalized to the untreated sample. For 3'-random amplification of cDNA ends (RACE) experiments, 2 μg of RNA was retrotranscribed with 0.5 μg of an oligo(dT_{20}) tail followed by an anchor sequence (CTGATCTAGAGGTACCGGATCC). To evaluate the differential expression of 3'-end isoforms, we performed semi-quantitative PCR with 40 ng of anchor-tagged cDNA using a gene-specific forward primer and an anchor reverse primer. PCR products were resolved in an agarose gel, purified and sequenced. Oligonucleotides used in this study are listed in Supplementary Table S1

3' READS and data analysis

Total RNA was processed by 3' region extraction and deep sequencing (3' READS) (17). Briefly, total RNA was subjected to poly(A) selection using the Poly(A)Purist™ MAG kit (Ambion). After fragmentation using the RNA fragmentation kit (Ambion), poly(A)-containing fragments were isolated using MyOne streptavidin C1 beads (Invitrogen) coated with a 5'-biotinylated chimeric U5T45 oligo (Sigma). RNA bound to the beads was digested with RNase H (5 U in 50 μl reaction volume) at 37°C for 1 h, which removed the poly(A) sequence, and eluted. Purified RNA fragments were phosphorylated of the 5' end with T4 kinase (NEB), purified by the RNeasy kit (Qiagen) and ligated to a 5'-adenylated 3' adapter (5'-rApp/NNNNGATCGTCG GACTGTAGAACTCTGAAC/3ddC) with the truncated T4 RNA ligase II (Bioo Scientific) and to a 5' adapter (5'-GUUCAGAGUUCUACAGUCCGACGAUC) by T4 RNA ligase I (NEB). The RNA was reverse transcribed by Superscript III (Invitrogen), followed by 12 cycles of PCR amplification with Phusion high fidelity polymerase (NEB). cDNA libraries were sequenced on an Illumina Hiseq 2000 sequencer. For data analysis, 5'-adapter sequences were removed from reads, and reads <15 nucleotides (nt) after this step were discarded. Reads were mapped in the hg19 genome sequence. We used only reads with mapping quality score (MAPQ) ≥10 and required mismatches to be ≤5% of the read. Reads with two or more unaligned Ts at the 5' end were called poly(A) site-supporting (PASS) reads and were used to identify pAs. Multiple pAs within 24 nt were clustered, but each cluster did not span >48 nt. pA sites mapped to the genome were further assigned to genes by using RefSeq, Ensembl and University of California at Santa Cruz Known Gene databases. To reduce false pA sites, we further required that the number of PASS reads for a pA site was ≥5% of all PASS reads for the gene and detected in at least two samples.

3' READS data were used to identify pA sites and to measure transcript expression. For GE analysis, all PASS reads for a gene were summed to calculate the expression level of the gene (18). *P*-value <0.05 and fold change >20% were used to select significantly regulated genes. The abundance of a pA isoform was defined as the fraction of PASS reads corresponding to the pA site over all PASS reads regarding the gene. For APA analysis, we compared pA site isoform abundance (ΔAbn) between the Enzalutamide sample and the vehicle sample. *P*-value <0.05 was applied as the cut-off to select significantly regulated APA events.

Bioinformatic analysis

The KLK3 pA sites were analyzed utilizing published RNA-seq data (GSE40050) and the PolyA_DB database for annotation of polyadenylation signals (PASs) (19). AR-binding sites were searched using deposited chromatin immunoprecipitation (ChIP) data obtained in LNCaP (GSE55064) and VCaP cells (GSE40050). All data were displayed in the Integrative Genomics Viewer (IGV) tool. Gene Ontology (GO) term enrichment was performed in the R environment (http://www.r-projects.org) using Fisher's exact test with the 'elim' algorithm of the topGO package (R package version 2.36.0). The 16 most significant terms were graphed using ggplot2 package (R package version 3.2.1).

Cis-elements potentially differing between up-regulated and down-regulated pAs were identified by k-mer (5 nt) enrichment analysis using R custom script. Briefly, we counted the number of occurrences of each 5-mer in a pA that was up-regulated or down-regulated and used the Fisher's exact test to examine the significant association of each 5-mer with each group. The enriched 5-mers were ordered for abundance and significance. Next, we aligned the enriched motifs to generate sequence logos using WebLogo 3 (http://weblogo.threeplusone.com/). Lastly, we calculated the frequency of motifs (AAUAAA, UUUU, UGUG and UGUA) in the ± 100 nt transcript region surrounding each group of pAs using custom R script. The frequencies were graphed withthe ggplot2 package using a span value of 15/201 and a 0.95 level of confidence interval.

PC dataset analysis was carried out using the TCGA Pan-Cancer Atlas (20) or other published PC datasets (21,22). *Z*-score values and progression-free survival data relative to patients were downloaded from cBioPortal (https://www.cbioportal.org/) and the webtool developed by Cancer Research UK Cambridge Institute (https://bioinformatics.cruk.cam.ac.uk/apps/camcAPP/). For GE analysis, patients were divided into two groups (castration recurrent and not) and normalized mRNA levels were plotted and tested with Student's *t*-test. Log-rank (Mantel–Cox) test was used to evaluate the survival contribution of C/P factors splitting the high-expression and low-expression cohorts. The median expression value of each gene of interest was calculated in the whole cohort, patients were then segregated in a low-expression group, comprising those with expression levels below the median value, and a high-expression group, comprising patients with expression levels above the median value (GraphPad Prism version 7.0.0 for Mac OS X, GraphPad Software, San Diego, CA, USA).

Isolation of BrU-labeled RNA

LNCaP cells were treated with 1 μM Enzalutamide or vehicle, and bromodeoxyuridine (BrU) was added in the medium at a final concentration of 2 mM in the last hour. After phosphate-buffered saline (PBS) washes, RNA was isolated by standard procedures, DNase treated and quantified. A 30 μg aliquot of RNA was denatured for 5 min at 70°C, diluted at 1 ml in incubation buffer [PBS, 0.05% bovine serum albumin (BSA) and 0.1% Tween] supplemented with RNasin ribonuclease inhibitors (N2518, Promega), and incubated with 1.25 μg of anti-BrdU antibody (sc-32323, Santa Cruz Biotechnology) in the presence of Dynabeads protein G (Invitrogen) at 4°C for 2 h. The bead–antibody–RNA complexes were washed three times with washing buffer (PBS, 0.1% Tween and 1:1000 RNasin ribonuclease inhibitors) and RNA was isolated using the Trizol reagent (Invitrogen), retrotranscribed with random hexamers (Roche) and used for RT-PCR as described for total RNA.

Co-immunoprecipitation assay and western blotting

Nuclear extracts (1 mg) were prepared as previously described (23) and incubated with 3 μg of antibody specific for the protein of interest or immunoglobulin G (IgG) (negative control) in the presence of Dynabeads protein G (Invitrogen) in a total volume of 1 ml, for 3 h at 4°C under rotation. After rinsing, the co-immunoprecipitates were resuspended in Laemmli sample buffer and processed for western blot analysis. Immunoprecipitated proteins and whole-cell extracts prepared as previously described (23) were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), transferred to polyvinylidene fluoride (PVDF) membranes and incubated for western blot analyses, with the antibodies listed in Supplementary Table S2. Densitometric analysis of western blot was performed using ImageJ 1.51g software.

Chromatin immunoprecipitation assay

ChIP experiments were carried out as previously described (24) using LNCaP cells cross-linked with 1% formaldehyde for 10 min at room temperature and quenched with 125 mM glycine solution.

Briefly, nuclei (24) were sonicated in a Bioruptor sonication waterbath (Diagenode) and sonicated chromatin was quantified to allow the dilution of samples at the same concentration. A 100 μg aliquot of chromatin was incubated with the specific antibodies or IgGs (negative control) under rotation at 4°C overnight. In the last 2 h, Dynabeads protein G (Invitrogen) were added to the mixture. The bead–antibody–protein–DNA complexes were washed twice with low salt buffer (0.1% SDS, 2 mM EDTA, 1% Triton, 20 mM Tris pH 8, and 150 mM NaCl), twice with high salt buffer (0.1% SDS, 2 mM EDTA, 1% Triton, 20 mM Tris pH 8, and 500 mM NaCl), twice with TE (1 mM EDTA, 10 mM Tris pH 8) and eluted in 94 μl of TE. RNA was degraded by addition of 4 μl of 5 M NaCl and 2 μl of RNase cocktail (AM2286, Invitrogen) for 1 h at 37°C. Cross-links were reversed overnight at 65°C and proteins were degraded by addition of Proteinase K (150 μg; Invitrogen) for 2 h at 55°C. Immunoprecipitated DNA was isolated using phenol:chloroform:isoamyl alcohol (25:24:1, Sigma) and resuspended in 60 μl of H_2O. Part of the diluted chromatin in each sample was treated to degrade RNA and proteins and used as input. ChIP-PCR was performed with 2 μl of DNA as described (24) and results were shown as a percentage of input, calculated by the ΔCt method.

UV cross-linking RNA/protein immunoprecipitation (CLIP) assay

The CLIP experiments were carried out as previously described (25) by using UV irradiation (400 mJ/cm^2) to cross-link the cells. A total of 1 mg of nuclear extract was incubated with the specific antibodies or IgGs (negative control) in the presence of Dynabeads protein G (Invitrogen, Life Technologies) and 10 μl of diluted RNase I (1:1000, Ambion) for 2 h at 4°C under rotation. Immunoprecipitates were washed as described (25) before treatment with Proteinase K (50 μg) for 1 h at 55°C. RNA was isolated and retrotranscribed by standard procedures. Each sample was normalized with respect to its input. Target RNA associated with proteins in the Enzalutamide-treated condition was represented as fold enrichment relative to control samples (vehicle-treated).

CASK minigene assay

The calcium/calmodulin-dependent serine protein kinase (CASK) internal pA (IPA) and terminal pA (TPA)genomic regions were amplified and cloned into the pRiG vector digested with XhoI and EcoRI restriction enzymes (26). The mutant minigenes for CSTF64- and FIP1L1-binding sites were generated by the megaprimer strategy. All primers are listed in Supplementary Table S1. Transfection was carried out using jetPRIME (Polyplus-transfection) using 100 ng for each well of a 12-well plate when the confluency of cells was 70%.

Fluorescence-activated cell sorting (FACS) analysis

For cell cycle analyses, LNCaP cells were treated for 24, 48 and 72 h with 1 μM Enzalutamide or dimethylsulfoxide (DMSO) as vehicle control and pulsed with 30 μM BrdU (Sigma-Aldrich) for the last 30 min. Collected cells were then processed for analysis of BrdU-positive cells by flow cytometry (FACSCalibur, BD Biosciences) as previously described (24). For reporter plasmids, HEK293T cells were transfected with pRiG minigenes and collected after 24 h to measure the green [enhanced green fluorecent protein (EGFP)] and red [red fluorescent protein (RFP)] fluorescent signals using a CytoFlex Flow Cytometer (Beckam Coulter). Signals were analyzed using CytExpert ver 2.2 software (Beckam Coulter) and FlowJo X ver. 0.7 software, and cells negative for both signals were filtered out. Log2(Red/Green) were calculated for each cell.

RESULTS

AR inhibition modulates a widespread APA program in prostate cancer cells

The androgen-responsive gene *KLK3*, which codes for the prostate-specific antigen (PSA), comprises two actionable pAs that are annotated in the polyA_database (19) (Figure 1A). Inspection of RNA sequencing data (GSE40050) indicated that the distal pA (d-pA) at the 3' end of the untranslated region (UTR) is preferentially used under stimulation of androgen-sensitive LNCaP cells with dihydrotestosterone (DHT), whereas the proximal pA (p-pA) was activated upon androgen deprivation in cells grown with CSS (Figure 1A). To directly test whether *KLK3* APA is modulated by androgens, LNCaP cells were grown under HT conditions by treatment with increasing doses of Enzalutamide (0.3–3 μM). 3'-RACE analysis of *KLK3* transcripts isolated after 24 h of treatment confirmed a dose-dependent increase of p-pA selection, with concomitant reduction in usage of the canonical d-pA (Figure 1B). Similar effects were also observed upon androgen deprivation with CSS-supplemented medium for 48 h (Supplementary Figure S1A). These results document that androgen availability modulates APA regulation of a prototypical androgen-responsive gene.

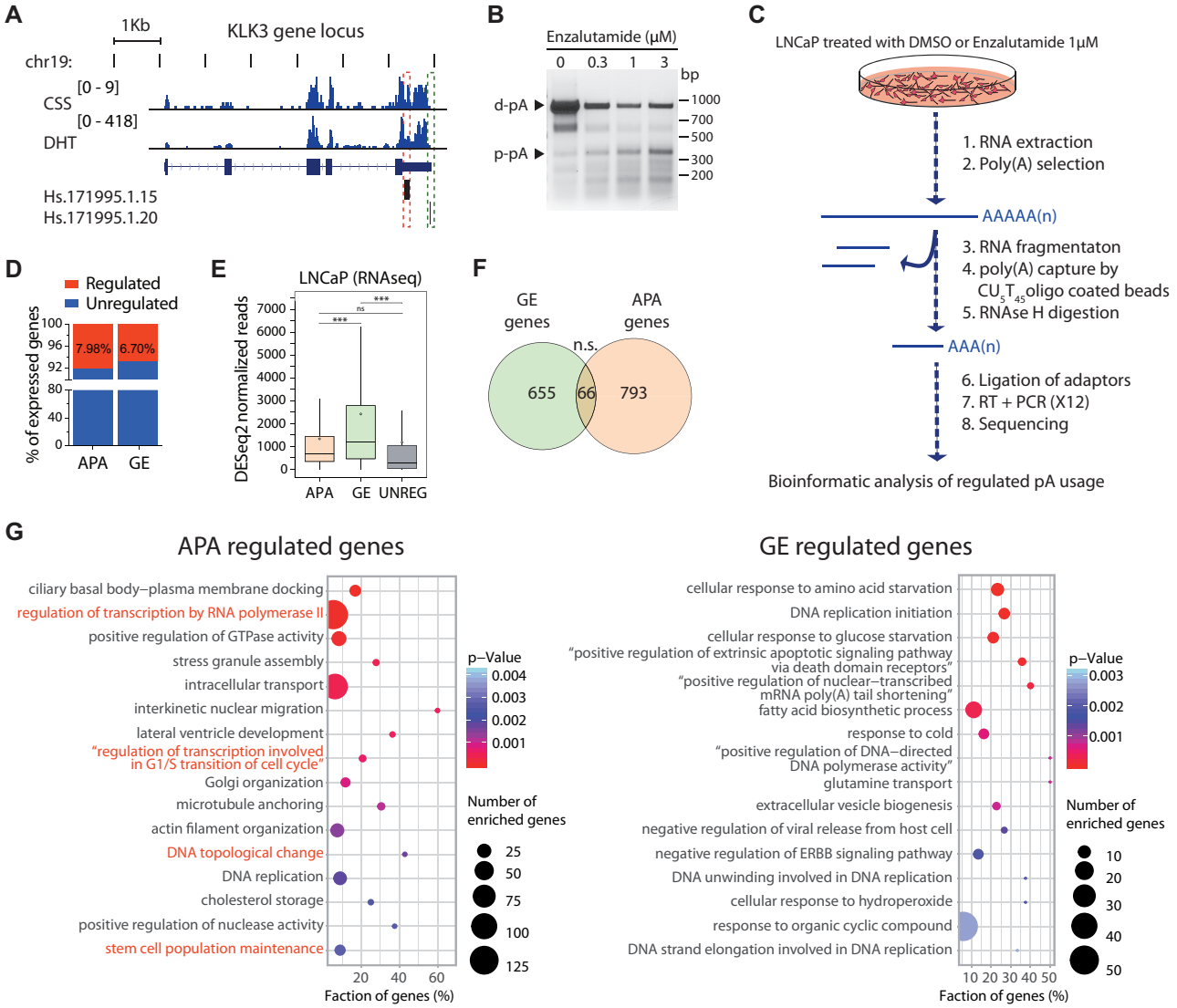

Figure 1. (**A**) Human *KLK3* polyadenylation sites were analyzed by using RNA-seq tracks (GSE40050) of CSS-starved LNCaP cells treated or not with DHT. Two pAs were identified by the polyA_DB tracks in the 3'UTR of *KLK3*, which were either up- (red box) or down-regulated (green box) by hormone treatment. (**B**) LNCaP cells were treated with increasing doses of Enzalutamide for 24 h and 3'-RACE analysis of the *KLK3* 3'UTR was performed. Bands corresponding to p-pA and d-pA are indicated by arrowheads. (**C**) Schematic overview of the 3' READS method. (**D**) Bar graphs representing the percentage of LNCaP expressed genes that are regulated at either the APA or GE level upon treatment with Enzalutamide (highlighted in red). (**E**) Boxplot graph representing the normalized DESeq2 reads relative to APA-regulated, GE-regulated and unregulated genes in LNCaP cells. (**F**) Overlap of APA- and GE-regulated genes. Statistical significance was evaluated by the hypergeometric test. (**G**) GO term enrichment of APA- and GE-regulated genes (modified Fisher's exact test).

Next, we asked whether HT causes a general regulation of APA in androgen-sensitive PC cells. To rule out secondary effects due to the cytostatic action of Enzalutamide, we first evaluated the impact of the drug on cell cycle progression. Flow cytometry analysis of BrdU incorporation in the S phase of the cycle showed that treatment with 1 μM Enzalutamide for 24 and 48 h did not significantly affect cell cycle progression, whereas a reduction of cells in S phase was observed at 72 h of treatment (Supplementary Figure S1B). On this basis, we selected 24 h of treatment with 1 μM Enzalutamide for the analysis of APA regulation by 3' READS (Figure 1C; GSE190153), which allows detection of pA usage together with overall expression of transcripts (17). Under these conditions, Enzalutamide caused a switch in pA selection in 7.98% ($n = 859$) of the genes expressed in LNCaP cells, whereas 6.7% of them ($n = 721$) were affected at the overall gene expression (GE) level (Figure 1D; Supplementary Figure S1C; Supplementary Tables S3 and S4). Analysis of a deposited LNCaP transcriptomic dataset (GSE195916) indicated that GE-regulated genes are expressed at significantly higher level than unregulated genes in LNCaP cells, while APA-regulated genes showed no significant difference, albeit their median expression level was slightly higher than that of unregulated genes (Figure 1E). There was a very limited overlap between APA and GE regulation, with only 66 genes being modulated at both levels (Figure 1F; Supplementary Figure S1D; Supplementary Table S5). Moreover, genes regulated at APA

and GE levels were enriched in functional categories related to different biological processes. APA-regulated genes are involved in transcriptional regulation, DNA topological changes and stem cell maintenance (Figure 1G), which suggests an important impact of APA regulation on transcriptome reprogramming of PC cells in response to HT. On the other hand, functional categories related to starvation and biosynthetic pathways were up-regulated among GE-regulated genes, whereas genes involved in regulation of DNA replication were down-regulated (Figure 1G; Supplementary Figure S1E). Collectively, these findings uncover a widespread APA program that is set in motion by HT in androgen-responsive PC cells.

HT induces preferential lengthening of APA-regulated transcripts

APA can lead to changes in the protein coding sequence of the transcript (CDS-APA) or in the length of the regulatory 3'UTR (UTR-APA) (16). In CDS-APA, the alternative pAs are localized within the transcription unit (IPA), mainly in alternative last exons or in introns (Figure 2A). In UTR-APA, instead, all regulated pAs are localized in the distal-most last exon (TPA) and their utilization does not affect the protein output (Figure 2A). However, the different UTR length downstream of the stop codon can confer specific regulatory features to UTR-APA variants (16). HT affected both CDS-APA and UTR-APA events (Supplementary Figure S2A). However, while CDS-APA events were significantly enriched with respect to their representation in LNCaP cells ($P = 3.69^{e-4}$, Fisher's exact test), UTR-APA events were under-represented ($P = 4.80^{e-2}$, Fisher's exact test; Figure 2B). Among the CDS-APA, IPA events were significantly repressed by Enzalutamide ($P < 0.039$, Fisher's exact test, Figure 2C), whereas unique pAs at the end of the transcription unit (S) were up-regulated ($P < 0.0004$, Fisher's exact test, Figure 2C). Furthermore, there was a significant increase in usage of the distal-most pAs (L) in the UTR ($P = 0.027$, Fisher's exact test; Figure 2C).

These observations suggested that HT prevalently promotes lengthening of transcripts in LNCaP cells. To test this possibility, we analyzed the regulation of pA selection at the transcript level. Each Enzalutamide-regulated pA was classified as proximal or distal according to its position with respect to the other regulated pAs or to the position of the most used pAs in the same gene. First, by quantifying the extent of changes in APA isoform abundance (ΔAbn), we observed a preferential repression of p-pAs (mean = –4.89, $n = 569$) and up-regulation of d-pAs (mean = +0.88, $n = 488$; $P < 0.0001$, t-test; Figure 2D). This trend was even more pronounced when only transcripts regulated by CDS-APA were analyzed (means –4.9 versus +5.4, $P < 0.0001$, t-test; Supplementary Figure S2B). Accordingly, a higher number of genes underwent transcript lengthening with respect to shortening (57.78%; Supplementary Figure S2C). Lengthened transcripts tended to have a significantly larger distance between p-pA and d-pA than transcripts undergoing shortening ($P = 0.0179$, Mann–Whitney test; Supplementary Figure S2D). In line with this result, transcript lengthening was more represented in CDS-APA- (63.98%) than in UTR-APA-regulated genes (54.31%, $P = 6.47^{e-3}$, Fisher's exact test; Figure 2E; Supplementary Table S6). Lastly, analysis of the relative usage of pAs in 182 genes that displayed significant modulation of at least two alternative pAs (Supplementary Table S7) highlighted a highly significant negative correlation between the abundance of the two APA isoforms for most genes (92.85%; $r = –0.8936$, $P < 0.0001$, Pearson correlation; Figure 2F), thus confirming the effect on transcript lengthening induced by HT. Moreover, CDS-APA-regulated transcripts were mostly lengthened (74%; Figure 2F). These results confirm that Enzalutamide widely affects transcript length through APA regulation in LNCaP, leading to preferential expression of longer transcripts.

HT causes direct APA regulation of AR target genes

qRT-PCR analysis of a subset of APA events ($n = 13$) confirmed the results of the 3' READS analysis (Figure 3A, B; Supplementary Figure S3A). Most of these HT-modulated APA events (80%) were also observed in Enzalutamide-treated VCaP cells (Figure 3C; Supplementary Figure S3B), a PC cell line that is also responsive to androgenic stimulation (27). Moreover, AR inhibition with Bicalutamide, another AR antagonist (3), modulated these APA events in the same direction as Enzalutamide (Supplementary Figure S4A, B). Lastly, knockdown of AR expression by transient transduction of lentiviral vectors encoding two different shRNAs (Supplementary Figure S4C) faithfully recapitulated the APA changes induced by Enzalutamide in LNCaP cells (Figure 3D; Supplementary Figure S4D). These results indicate that repression of AR activity is directly involved in APA regulation under HT conditions.

AR signaling was recently shown to impact on pre-mRNA processing by regulating the expression of selected splicing factors, such as ESRP1 and ESRP2 (28,29). However, we did not detect any change in expression or APA for the main components of the CPSF or of other complexes that regulate cleavage and polyadenylation [i.e. the CSTF and the cleavage factor I (CFI) and II (CFII) complexes; Supplementary Tables S3 and S4]. Changes in abundance of APA isoforms could also be related to their differential stability in the presence of Enzalutamide. To test this possibility, transcription was blocked by treatment with actinomycin D and transcript decay was determined in the following 12 h. Nevertheless, this analysis showed that the stability of each APA isoform was not significantly affected by AR inhibition (Supplementary Figure S5A).

Since APA regulation generally occurs co-transcriptionally (16,30–32), we next analyzed production of HT-induced APA isoforms in nascent transcripts labeled with BrU (Figure 3E). Importantly, analysis of transcripts synthesized during the last hour of treatment displayed the same switch in APA that was identified at steady state (Figure 3F; Supplementary Figure S5B, C), suggesting that Enzalutamide co-transcriptionally induces changes in APA. Moreover, by querying ChIP-sequencing (ChIP-seq) datasets (33) from the ENCODE project (ChEA 2016), we found a significant enrichment in APA-regulated genes displaying AR binding in their promoter/enhancer regions in PC samples ($n = 133$, $P = 4.41^{e-07}$) and LNCaP cells ($n = 29$, $P = 0.001$; Supplementary Figure S6A). The other

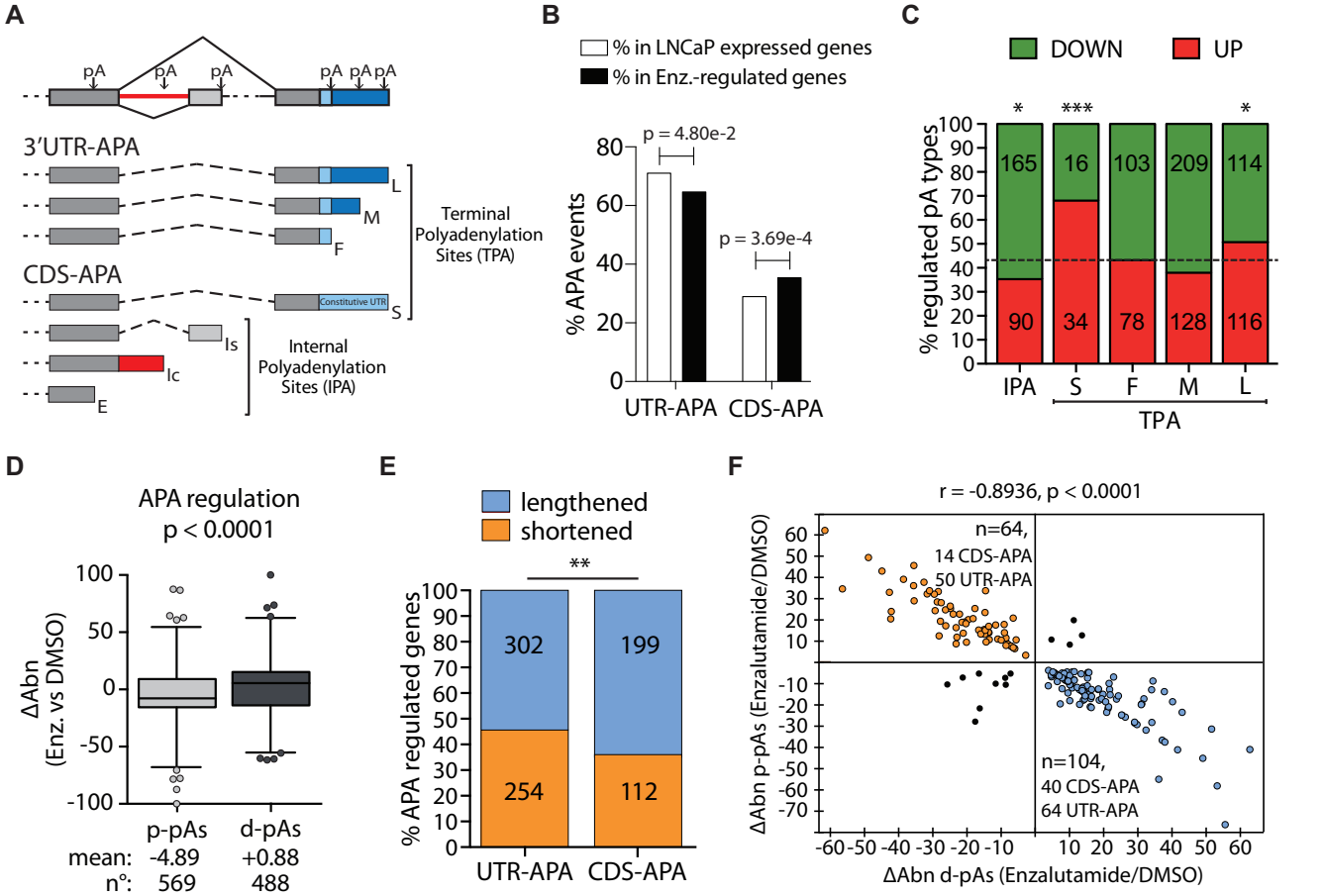

Figure 2. (**A**) Schematic representation of the APA patterns. In 3'UTR-APA (top), the pAs were classified based on its relative location in the gene as first (F), middle (M) and last (L). The region upstream of the first pA is the constitutive 3'UTR (light blue), the downstream region is alternative (dark blue). CDS-APA (bottom) can have a single pA (S) or different pAs in the last exon (TPAs) that compete with the choice of the IPA, which are named composite terminal exon (Ic), skipped terminal exon (Is) and internal exonic pA (E). (**B**) Bar graph showing the percentage of UTR-APA and CDS-APA events annotated (white columns) and regulated by Enzalutamide (black columns) in LNCaP cells. Statistical analysis was performed by Fisher's exact test. (**C**) Bar graphs showing the percentage and number of up-regulated (red) and down-regulated (green) pAs by Enzalutamide, classified into five types as described in (A). The dotted line indicates the percentage of up-regulated pAs in all Enzalutamide-regulated events used as reference for Fisher's exact test (*P <0.05, ***P <0.001). (**D**) Box plot representing changes (Enzalutamide/DMSO) in APA isoform abundance (ΔAbn) expressed in normalized read density for p-pA and d-pA isoforms. Mean values and number of pA isoforms (n°) are reported. Statistical analysis was performed by Student's t-test. (**E**) The number of genes that undergo shortening and lengthening of the transcript are shown in the bar graph with the relative percentage. $P = 0.00647$, Fisher's exact test. (**F**) Scatter plot showing changes of isoform abundance upon Enzalutamide treatment referred to 182 genes that displayed both a statistically significant (P <0.05) regulated p-pA (y-axis) and d-pA (x-axis). Pearson's correlation coefficient (r) and P-value are reported.

most enriched transcription factors were classified as AR-interactors in the STRING database (https://string-db.org) (Supplementary Figure S6B). On this basis, we tested whether AR is recruited to the promoter/enhancer regions of APA-regulated genes by performing ChIP assays in LNCaP cells treated with DHT for 24 h in the presence or absence of Enzalutamide. To identify the putative AR-binding sites, we searched for ChIP signals in deposited PC cell datasets (GSE55064; GSE40050) and designed PCR primers flanking the most relevant AR peaks (Supplementary Figure S6C). As control for AR binding, we also tested two GE-regulated genes: *RBM47*, a new AR target gene identified by our analysis; and *KLK3* (Supplementary Figure S6D). ChIP assays indicated that AR is specifically recruited to the regulatory regions of both GE- and APA-regulated genes. Moreover, binding of AR to the regulatory regions of APA-regulated genes was significantly reduced by treatment with Enzalutamide (Figure 3G), similarly to what was observed for GE-regulated genes (Supplementary Figure S6E). These results indicate that APA-regulated genes are direct targets of AR, whose regulation is sensitive to androgenic stimulation.

AR inhibition induces rearrangements of the cleavage and polyadenylation complexes

Processing of pre-mRNAs involves the recognition of the cleavage site (CS) by the C/P machinery, transcript cleavage and subsequent recruitment of the poly(A) polymerase (PAP) that catalyzes polyadenylation of the 3' end (16,34). Recognition of *cis*-acting elements located near the CS by C/P subcomplexes occurs co-transcriptionally (35), with some factors being recruited at the promoter regions of transcribed genes (31,32). This process is mediated by di-

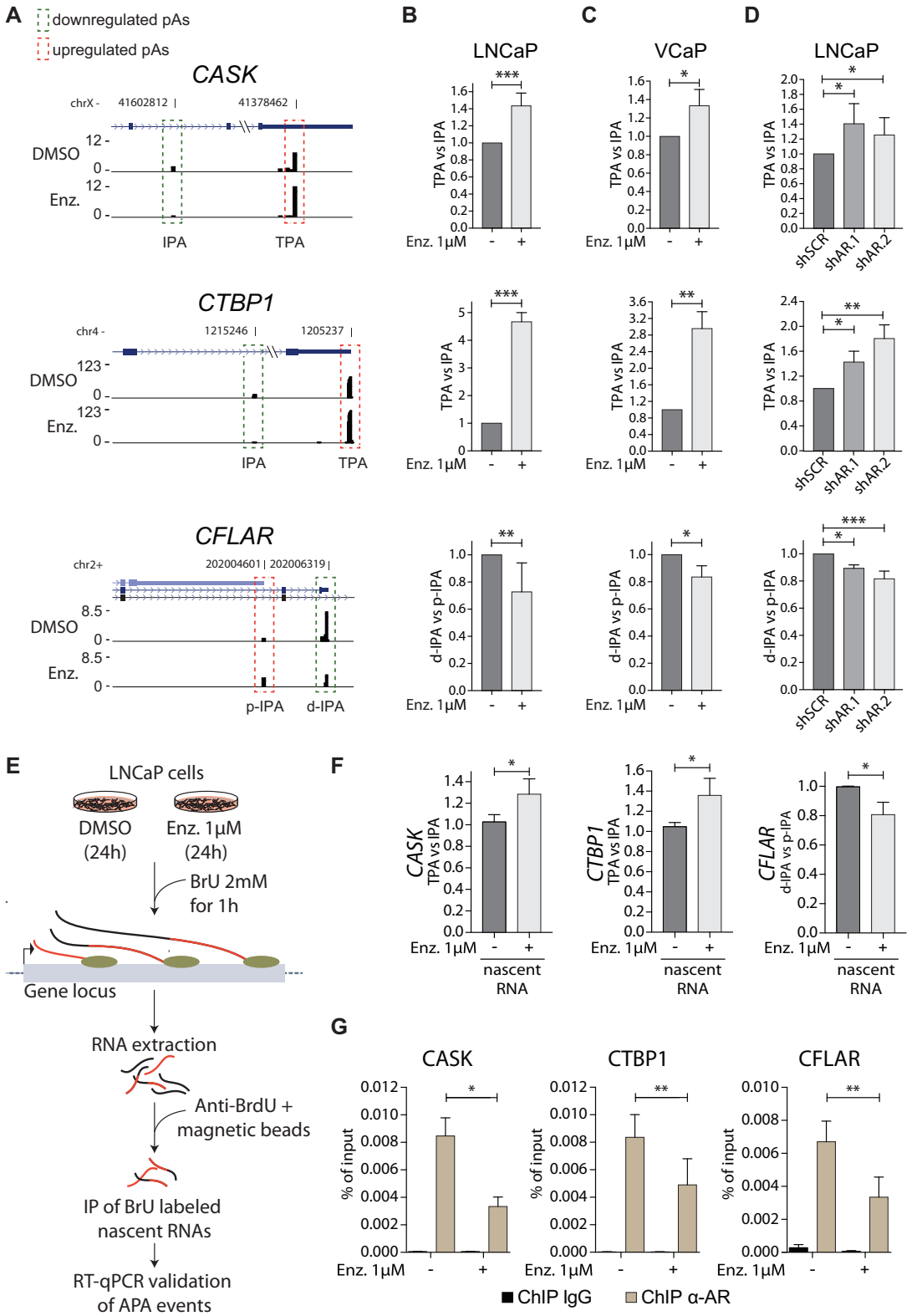

Figure 3. (**A**) 3' READS data (modified from UCSC genome browser tracks) for three representative APA-regulated genes in LNCaP cells treated or not with Enzalutamide. Red and green boxes indicate up- and down-regulated pA events, respectively. (**B–D**) qRT-PCR analysis of the APA events represented in (A) performed in LNCaP (B) and VCaP (C) cells treated or not with Enzalutamide, or LNCaP cells silenced for AR (D). Bar graphs represent the expression of TPA or d-pA versus IPA or p-pA, expressed as fold change with respect to untreated or sh-control samples. (**E**) Schematic representation of the isolation and analysis of nascent RNAs. (**F**) qRT-PCR analysis of the APA events in nascent RNA isolated as described in (E). (**G**) qRT-PCR analyses of ChIP experiments performed in LNCaP cells using AR antibody or rabbit IgG, as negative control. The associated DNA amount was expressed as a percentage (%) of input. All qRT-PCR analyses represent the mean of at least three independent experiments with relative standard deviation (SD). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ (Student's *t*-test).

rect binding of C/P components either to general transcription factors (36,37) or to the C-terminal domain (CTD) of RNA polymerase II (RNAPII) (30). Serine 2 phosphorylation of the RNAPII CTD, which peaks in proximity to the CS and affects APA regulation (38), is under control of AR interactors (39). However, treatment with Enzalutamide did not significantly affect this post-translational modification in LNCaP cells (Supplementary Figure S7A). AR inhibition could also modulate the assembly of C/P subcomplexes on nascent transcripts, as factors involved in pre-mRNA processing were reported to interact with the AR in a ligand-dependent manner (40). Among others, mass spectrometry analyses suggested an interaction between the AR and CSTF50 after androgen stimulation (41). In line with this previous observation, CSTF50 co-immunoprecipitated with the AR in LNCaP cells grown under basal conditions (Figure 4A). However, treatment with Enzalutamide, which reduced the pool of AR in the nuclear- and chromatin-associated fractions (Supplementary Figure S7B), disrupted this interaction (Figure 4A). To test whether AR inhibition more generally caused the rearrangement of the C/P machinery, we performed co-immunoprecipitation experiments of the three major subcomplexes involved in 3' end processing (16). The CPSF complex recognizes the PAS motif and catalyzes pre-mRNA cleavage. The CFI and CSTF complexes bind, respectively, to auxiliary upstream (USE) and downstream sequence elements (DSE) flanking the PAS (Figure 4B). To enrich for functionally active complexes, we immunoprecipitated the subunit that directly binds the RNA in each complex: CPSF30 (PAS), CSTF64 (GU-rich DSE) and CFIm25 (UGUA USE) (16). CPSF30 efficiently co-immunoprecipitated with the other CPSF components in LNCaP cells, and its interaction with CPSF160 and FIP1L1 was significantly increased in the presence of Enzalutamide (Figure 4C, D). Increased interaction of FIP1L1 with CPSF30 was also observed upon immunoprecipitation of the former (Supplementary Figure S7C). In contrast, no differences were observed in the composition of the CSTF and CFI complexes (Supplementary Figure S7D, E). Enzalutamide treatment also impaired the interaction of CSTF64 with CPSF proteins, and particularly with CPSF160 (Figure 4E, F). The reduced interaction between CSTF and CPSF was confirmed by immunoprecipitating CSTF50 (Supplementary Figure S7F). These results suggest that AR activity modulates the assembly of the C/P subcomplexes and its inhibition impairs the interaction between CPSF and CSTF while increasing the recruitment of CPSF160 and FIP1L1 within the CPSF.

AR inhibition promotes selection of pAs flanked by upstream U-rich elements

Next, we asked whether APA-regulated genes were characterized by specific *cis*-elements that confer sensitivity to HT. First, we searched for pentameric motifs enriched in the sequences in four subregions encompassing ± 100 bp from the CS (Figure 5A), a region where most of the C/P subcomplexes are recruited (16,34). CPSF30 directly binds to the PAS sequence located at 10–40 nt upstream of the CS (42,43), while upstream U-rich elements and the UGUA motif recruit FIP1L1, a component of the CPSF complex that interacts with CPSF160, and the CFI complex, respectively (16,34,44,45). On the other hand, the CSTF complex binds to UG/U-rich elements located within 40 nt downstream of the CS (46). In addition, strong pAs generally display G-rich elements beyond this region (+40/+100 nt) (16). We found a significant enrichment of U-rich pentamers upstream of the up-regulated pAs, whereas UGUA elements were enriched in the same regions of down-regulated pAs (Figure 5B). Moreover, UG/U-rich elements were enriched in the downstream region of both up- and down-regulated pAs (Figure 5B; Supplementary Figure S8A).

To analyze in more detail the distribution of *cis*-elements in these regions, we calculated the positional frequencies of UUUU (FIP1L1), UGUG (CSTF64) and UGUA (CFI25) motifs (Supplementary Figure S8B, C). Enrichment of U-rich elements was observed in the whole upstream region of up-regulated pAs (Figure 5C). Notably, recruitment of FIP1L1 to U-rich sequences in this region is promoted by its interaction with CPSF160, a scaffold protein that stabilizes the CPSF complex (43) and whose recruitment to nascent transcripts is also increased in U-rich regions upstream of the PAS (35,44). Conversely, the UGUG motif was significantly enriched only within the first 25 nt downstream of the down-regulated pAs, in line with the known region of CSTF64 binding (35,46,47) (Figure 5C; Supplementary Figure S8B). Moreover, the UGUG motif was significantly depleted in the up-regulated pAs with respect to both down-regulated and unregulated pAs (Figure 5C). The UGUA motif was not significantly enriched in the –40/–100 nt region, with the exception of an enrichment in a small region near –75 nt (Supplementary Figure S8C), which does not correspond to the CFI positional recruitment peaking at –40/–50 nt from the CS (16,35). These results suggest that HT-induced rearrangements within the CSPF complex and its decreased interaction with the CSTF64 may contribute to differential selection of pAs that depend on these complexes.

AR inhibition promotes the selection of distal pAs that are not dependent on CSTF

Promoter-proximal pAs are frequently weaker than downstream sites, due to lower frequency of the canonical PAS sequence and auxiliary *cis*-elements (48). Accordingly, bioinformatics analyses of sequence elements in all genes expressed in LNCaP cells highlighted reduced frequency of the AAUAAA motif as well as of motifs recognized by the FIP1L1/CPSF160 (UUUU), CFI (UGUA) and CSTF (UGUG) near p-pAs with respect to d-pAs (Supplementary Figure S9A, C–F). Thus, although kinetically favored by their position, p-pAs need to compete with stronger d-pAs that are endowed with all canonical *cis*-elements. On the other hand, acquisition of CSTF motifs in p-pAs was shown to increase their selection with respect to d-pAs (46,48,49). Since we observed an overall lengthening of transcripts in Enzalutamide-treated cells, we asked whether the pAs in these transcripts also followed this general rule. Unexpectedly, the CSTF motif (UGUG) was significantly depleted in regulated d-pAs with respect to d-pAs of other expressed genes and was not enriched with respect to regulated p-

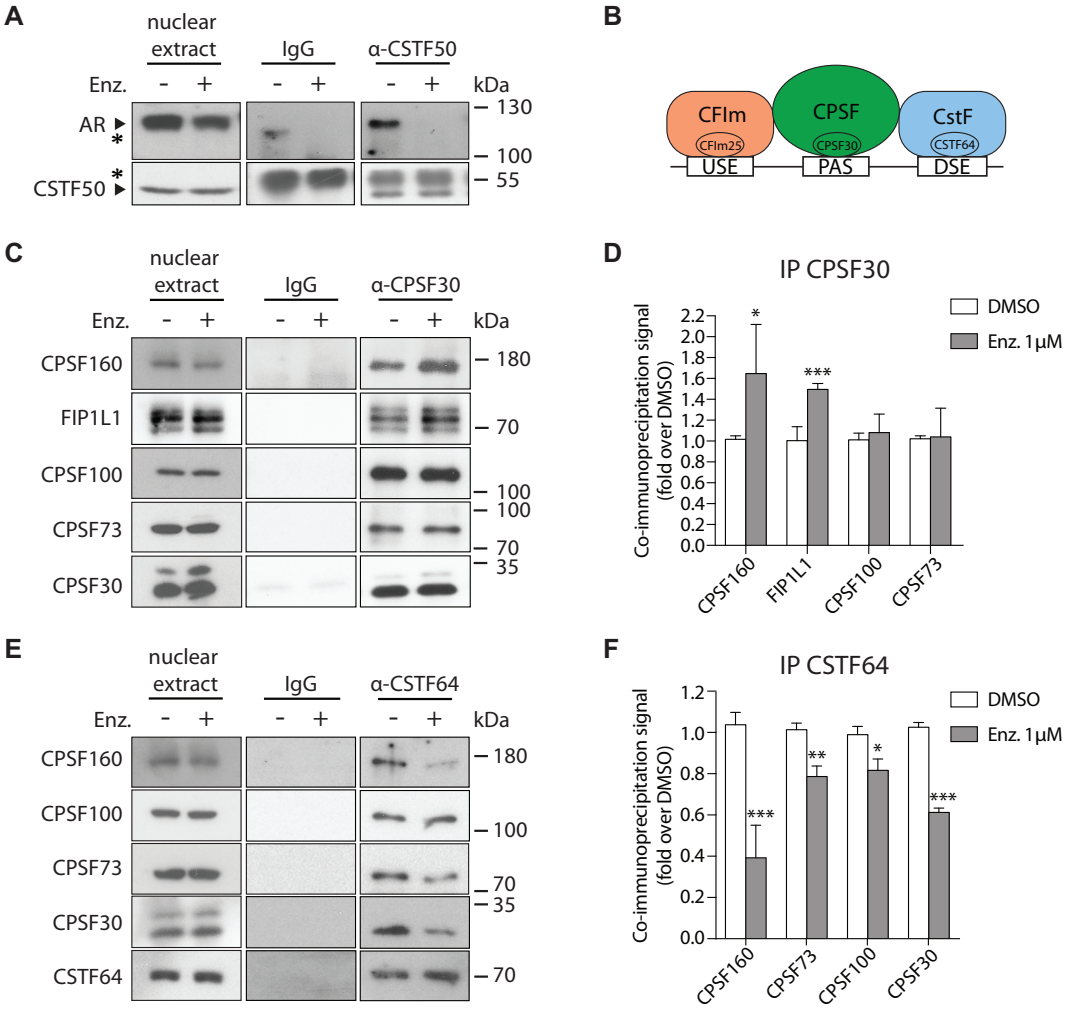

Figure 4. (**A**) Representative western blot analysis of the co-immunoprecipitation between the AR and CSTF50 using nuclear extracts from LNCaP cells treated or not with Enzalutamide. Rabbit IgGs were used as negative control. Asterisks indicate the non-specific signals and/or eluted antibody; arrows indicate the specific signals related to the AR and CSTF50 proteins. (**B**) Schematic diagram of the binding position of the main C/P complexes with their cognate *cis*-elements. Upstream (USE) and downstream sequence elements (DSE) flanking the PAS are indicated. (**C–F**) Western blot and densitometric analyses of the proteins that co-immunoprecipitate with CPSF30 (C, D) or CSTF64 (E, F) in LNCaP cells treated or not with Enzalutamide. Rabbit IgGs were used as negative control. Nuclear extracts represent 10–15% of input used in co-immunoprecipitation assays. In the bar graphs, the signals of proteins co-immunoprecipitated were represented as fold change in Enzalutamide- versus vehicle (DMSO)-treated cells. Mean of at least three independent experiments with relative SD are shown. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$ (Student's *t*-test).

pAs (Supplementary Figure S9E). The canonical PAS sequence and other *cis*-elements had a comparable frequency distribution between pAs in APA-regulated and unregulated genes (Supplementary Figure S9C–F). Notably, genes undergoing transcript lengthening during HT displayed a significant enrichment of FIP1L1/CPSF160- (UUUU) and depletion of CSTF64- (UGUG) binding sites in the d-pAs with respect to p-pAs, whereas shortened transcripts did not show significant differences in frequency distribution of these motifs (Figure 5D).

To test whether Enzalutamide promoted the selection of pAs without the auxiliary role of the CSTF, we performed CLIP assays for CPSF30, FIP1L1 and CSTF64 in our gene models. In line with the depletion of UGUG motifs near the d-pAs of lengthened transcripts (Figure 5D), we observed that recruitment of CSTF64 was stronger in p-pAs for both *CASK* and *CTBP1* transcripts (Supplementary Figure S10A, B). However, no significant changes in CSTF64 recruitment were observed under AR inhibition (Figure 5E). In contrast, CPSF30 and FIP1L1 binding to the d-pAs was induced by Enzalutamide treatment (Figure 5E). To better understand the contribution of FIP1L1 and CSTF64 to the regulation of these APA events, the alternative pAs of CASK were cloned in the pRiG plasmid, which contains an RFP and a GFP separated by an IRES sequence that allows concomitant translation of the two reporter genes (26). The CASK p-pA (IPA) and d-pA (TPA) (± 100 bp from cleavage sites) were cloned downstream of the RFP and upstream of the IRES sequence that drives translation of GFP in the bicistronic transcript (Figure 5F, G). Selection of the inserted pA would terminate the transcript before the IRES and increase the RFP/GFP ratio, which can be detected by flow cytometric (FACS) analysis (Supplementary Figure S10C–F). By using these mini-

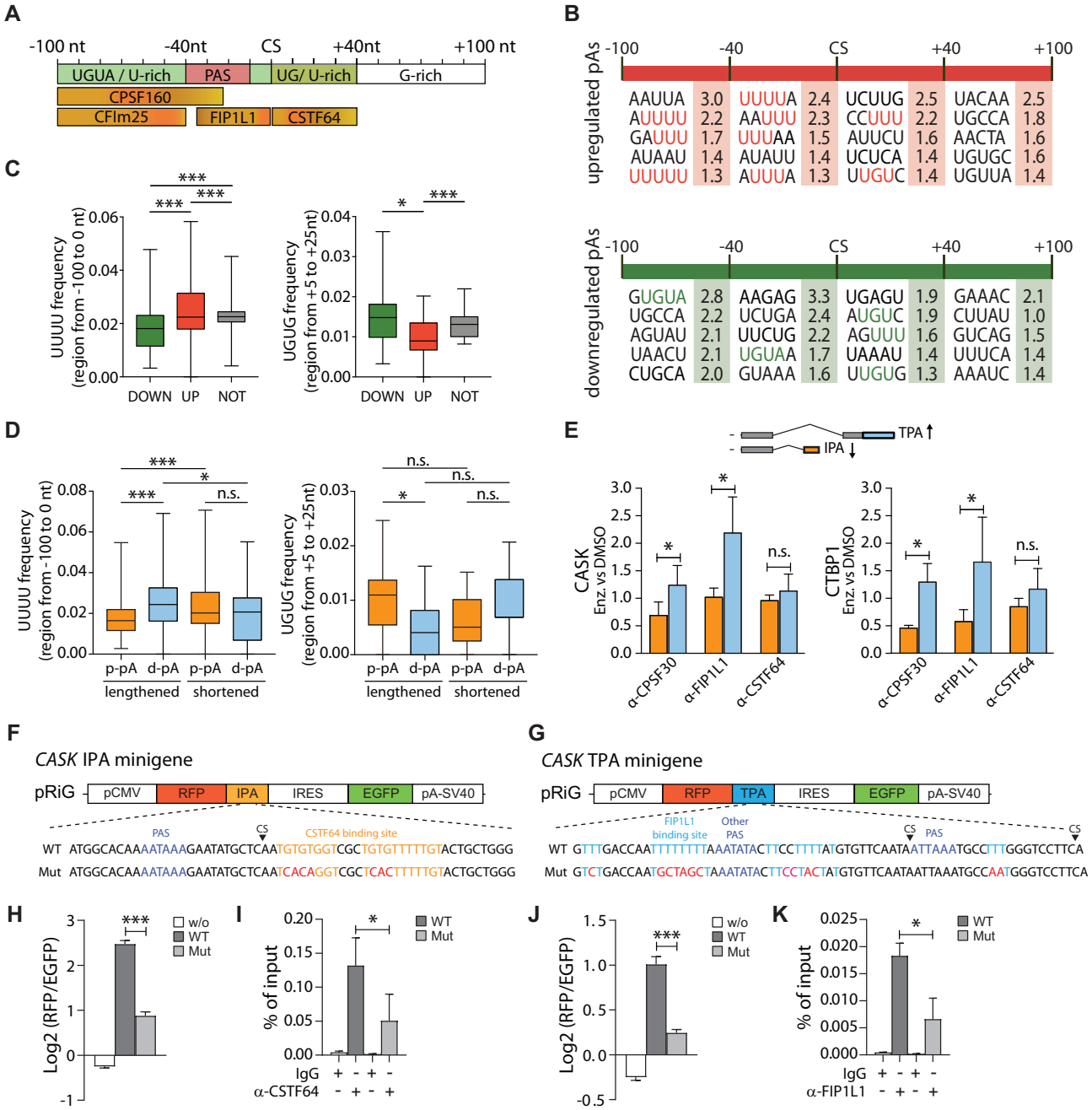

Figure 5. (**A**) Schematic diagram showing the position of *cis*-acting RNA elements between –100 nt and +100 nt relative to the CS and the complexes that bind to them (16,34). (**B**) Pentameric *cis*-elements that are significantly enriched in the indicated four regions surrounding the up- and down-regulated pAs. Enrichment analysis was performed using Fisher's exact test and expressed as –log10 of the *P*-value (highlighted numbers). Sequences that resemble the consensus binding motifs described in (A) are highlighted in red (up-regulated pAs) and green (down-regulated pAs). (**C**) Box plots represent the distribution of the indicated tetrameric motif in up- (red), down- (green) and unregulated (gray) pAs. UUUU frequency was evaluated in a region up to –100 nt from the CS (left graph), whereas UGUG frequency was evaluated in a region from +5 to +25 nt relative to the CS (right graph). ***P <0.001; *P <0.05; n.s. = not significant (Student's t-test). (**D**) Motif distribution at the p-pAs (orange) and d-pAs (light blue) of lengthened and shortened transcripts analyzed in the regions described in (C). (**E**) CLIP assays of CPSF30, FIP1L1 and CSTF64 in the region near the regulated pAs. IgGs were used as negative control (Supplementary Figure S10A, B). CLIP signals for p-pAs (orange) and d-pAs (light blue) of CASK and CTBP1 were quantified by qRT-PCR and represented as the ratio of the signals in Enzalutamide- versus vehicle-treated cells (mean of at least three independent experiments with relative SD; *P <0.05; n.s. = not significant, Student's t-test). (**F, G**) Scheme of the pRiG minigenes harboring CASK IPA and TPA sequences (± 100 bp from cleavage sites). PAS-, FIP1L1- and CSTF64-binding sites are highlighted in dark blue, light blue and orange, respectively. The position of the CS and mutated bases (in red) is also shown. (**H, J**) Comparison of log2(RFP/EGFP) obtained with minigenes harboring the wild type or mutated sequences (as in F, G), or empty pRiG vector (w/o). (**I, K**) CLIP assays to detect FIP1L1 and CSTF64 recruitment to CASK TPA and IPA in the wild type or mutated sequences. IgGs were used as negative control. Data represent the mean of at least three independent experiments with relative SD. *P <0.05; **P <0.01; ***P <0.001 (Student's t-test).

genes, we confirmed the functionality of both the IPA and TPA sequences of CASK (Figure 5H, J). Moreover, usage of the IPA and the TPA was significantly reduced when the CSTF64 (Figure 5F, H) or FIP1L1 (Figure 5G, J) motifs were mutated, respectively. Accordingly, CLIP assays indicated that FIP1L1 and CSTF64 efficiently bound the transcripts harboring the wild-type TPA and IPA sequences and that these interactions were reduced upon mutation of the U-rich and GU-rich elements, respectively (Figure 5I, K). Together with the reduced interaction between CPSF and CSTF (Figure 4E, F), these results suggest that AR inhibition induces the selection of a subset of U-rich d-pAs that are not dependent on CSTF. In support of this notion, depletion of CSTF64 in LNCaP cells did not impair the Enzalutamide-induced up-regulation of the *CTBP1* and *CASK* d-pA isoforms (Supplementary Figure 10G).

Transcriptional coupling of APA by the AR involves differential recruitment of CPSF30 to the promoter region

The CPSF and CSTF complexes can associate with RNAPII in gene promoter regions and are then deposited at the 3' end of the nascent transcripts, where they mediate pre-mRNA processing at the pA (31,36–38,50). To explore the recruitment of CPSF30 and CSTF64 at the promoter of APA-regulated genes, we performed ChIP assays. In line with the effects on their transcriptional regulation (Supplementary Figure S5D; Supplementary Table S4), Enzalutamide treatment significantly reduced RNAPII recruitment on the *KLK3* promoter but not on the *CASK* and *CTBP1* promoters (Figure 6A). Interestingly, CPSF30 recruitment was increased at the promoter region of these APA-regulated genes, regardless of the transcriptional effect (Figure 6B), whereas it did not bind to the promoter of a non-regulated gene (*KHDRBS1*; Supplementary Figure S10H). Recruitment of CSTF64 was not significantly affected by hormone treatment, albeit a slight reduction was observed in all genes tested (Figure 6C). These results suggest that AR activity primes APA regulation by affecting the assembly of C/P factors at the gene promoter regions. Moreover, our findings indicate that hormone treatments predispose the selection of U-rich pAs that are less dependent on CSTF function by favoring the recruitment of CPSF30 at their promoter region (Figure 6D).

Increased CPSF160 expression in CRPC cells promotes the HT-induced APA signature and correlates with poor prognosis

We asked whether CRPC cells that have acquired resistance to HT are characterized by changes in expression of the C/P factors that contribute to Enzalutamide-induced APA events. First, by querying expression data from PC patients in The Tumor Genome Atlas (TCGA; PanCancer Atlas) (20), we found that CPSF160 is the most up-regulated C/P component in PC samples from patients who underwent disease recurrence, which is associated with resistance to HT, with respect to non-recurrent patients (Figure 7A). CPSF73 and CFIm68 were also up-regulated in recurrent patients, albeit to a lesser extent (Supplementary Figure S11A), whereas CSTF50 and CFIm25 were significantly down-regulated (Figure 7A). We also observed a significant reduction in KLK3 expression in recurrent patients (Figure 7A), as well as in 22Rv1 cells with respect to the androgen-sensitive LNCaP cells (Supplementary Figure S11B, C). Querying of two additional datasets (21,22) confirmed the significant up-regulation of CPSF160 in CRPC patients with respect to primary tumors and primary/HT-sensitive PCs (Supplementary Figure S11D). Moreover, high expression of CPSF160 was significantly associated with worse prognosis in PC patients (Figure 7B, C), a feature related to acquisition of HT resistance and metastatic progression of the disease (1,2). Interestingly, combined high CPSF160 expression and low CSTF expression (CSTF50 or CSTF64) was associated with an even more significant negative prognosis (Figure 7C; Supplementary Figure S11E). Conversely, no improvement of prognostic value was observed in patients displaying high CPSF160 and low CFIm25 expression (Supplementary Figure S11E).

To test whether CPSF160 plays a role in CRPC cells, we evaluated its expression in 22Rv1 cells and in LNCaP cells that were selected upon chronic exposure to Enzalutamide (LNCaP-ER), which maintained the resistance to the drug and expressed lower *KLK3* levels like 22Rv1 cells (Supplementary Figure S11B, C). CPSF160 was up-regulated in both these CRPC cell lines with respect to LNCaP cells (Figure 7D). Further analysis in 22Rv1 cells confirmed that CPSF160 was the most up-regulated C/P factor (Supplementary Figure S11F, G). Moreover, depletion of CPSF160 in 22Rv1 cells reduced their colony-forming ability (Figure 7E; Supplementary Figure S11H), a metastatic feature of cancer cells. The increase in CPSF160 expression in 22Rv1 cells correlated with the same APA pattern of *CASK* (Figure 7F), *CTBP1* and *KLK3* (Supplementary Figure S11I, J) that was observed in Enzalutamide-treated LNCaP cells, whereas CPSF160 knockdown switched pA selection toward the pattern observed in untreated LNCaP cells (Figure 7G; Supplementary Figure 11K). These observations suggest that increased activity of CPSF160 promotes the HT-induced APA signature in CRPC cells.

To investigate whether the APA pattern induced by Enzalutamide contributes to the oncogenic features of PC cells, we focused on *CASK*, which encodes a kinase recently shown to promote PC progression (51). Selection of IPA generates a transcript variant terminating in the fourth intron of *CASK*, which comprises 27 exons. Enzalutamide treatment promotes the full-length APA isoform and is expected to increase CASK protein levels. Accordingly, the APA switch triggered by Enzalutamide correlates with an increase in CASK protein in LNCaP cells (Figure 7H). Increased CASK protein levels were also observed in 22Rv1 and LNCaP-ER cells (Figure 7H), which also show preferential use of the TPA (Figure 7F). It is noteworthy that knockdown of CASK in 22Rv1 cells limited their malignant phenotype, as indicated by the reduced colony-forming (Figure 7I; Supplementary Figure 11L) and invasive (Figure 7J) potentials. Collectively, these findings suggest that uncoupling CPSF and CSTF activity in PC cells exposed to HT reprograms the transcriptome through APA (Figure 7K), a process that might be associated with progression toward a CRPC phenotype.

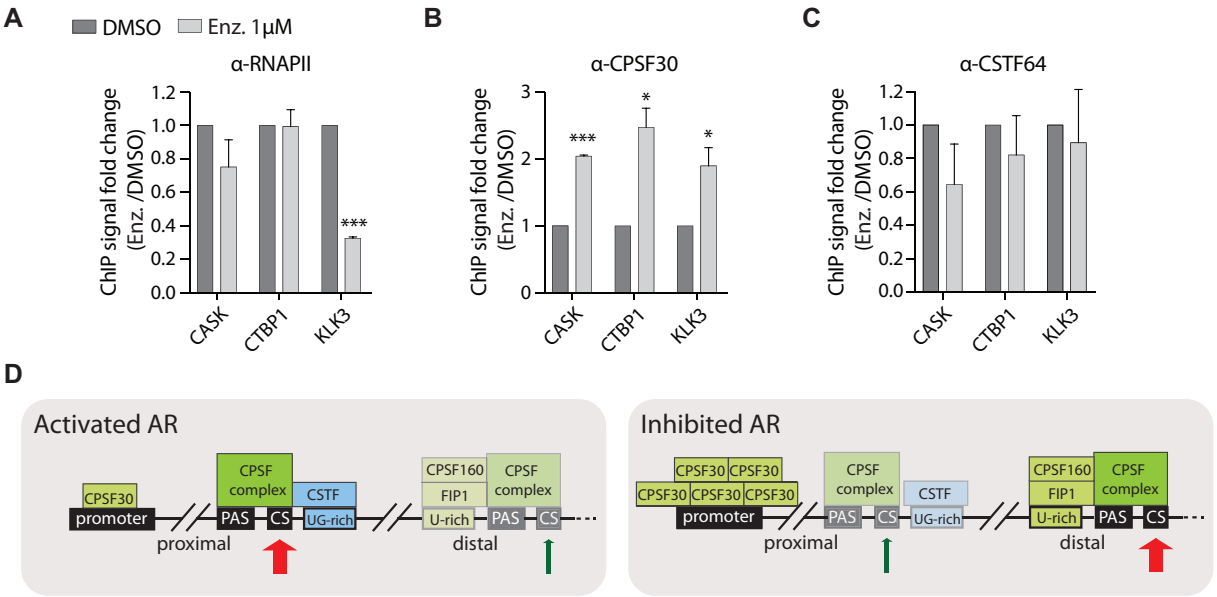

Figure 6. (**A–C**) qRT-PCR analyses of ChIP experiments performed in LNCaP cells using RNAPII, CPSF30 and CSTF64 antibodies, as indicated. Bar graphs show the amount of DNA immunoprecipitated in Enzalutamide-treated cells expressed as fold change with respect to untreated cells (mean of three independent experiments with relative SD. *P <0.05; ***P <0.001, Student's t-test). (**D**) Schematic model of the effects elicited by AR inhibition on APA regulation. Left panel: when AR is activated, p-pAs characterized by UG-rich elements recruit CSTF, which promote CPSF binding to the PAS. Right panel: AR inhibition allows increased recruitment of CPSF30 to the gene promoter, favors the disassembly of the CSTF–CPSF interaction (light colors) and promotes the usage of d-pAs that are independent of CSTF, thus causing transcript lengthening.

DISCUSSION

The AR guides a transcriptional program required for prostate gland development and function (1). On the other hand, uncontrolled AR activity promotes neoplastic transformation of prostate cells and is also involved in the progression of PC toward more advanced, therapy-resistant phenotypes (1–4). Thus, elucidating the mechanisms by which the AR modulates the transcriptome of prostate cells may highlight new molecular targets of its function, paving the way for the development of more efficacious therapies. In this work, we have discovered that hormone treatment triggers a widespread remodeling of the PC cell transcriptome through APA regulation. Inhibition of AR activity by Enzalutamide caused rearrangements in the CPSF complex and reduced its interaction with the CSTF complex, thus favoring pAs flanked by U-rich motifs. Notably, the APA pattern induced by short-term HT treatment appears to be maintained, at least in part, in PC cells that have acquired resistance to Enzalutamide, such as 22Rv1 cells. Moreover, this long-term response is associated with increased expression of CPSF160 in CRPC cells, a feature that correlates with disease recurrence and poor prognosis in PC patients. Thus, our study suggests that APA regulation contributes to the acquisition of HT resistance in advanced PC and that targeting CPSF160 expression or activity may represent a valuable approach to revert such resistance.

The AR is a ligand-activated transcription factor, and most efforts have focused on the transcriptional programs orchestrated by the AR in both PC and CRPC cells (51–53). Nevertheless, more recent evidence has indicated that AR activation also modulates gene expression at the post-transcriptional level, by causing global changes in splicing in the PC transcriptome (30,31). This effect was largely indirect and mediated by transcriptional control of the expression of the splicing factors ESRP1 and ESRP2. The ESRP1/2 proteins are expressed by epithelial cells and promote splicing of mRNA variants that are typical of the epithelial status (30,31). Accordingly, inhibition of ESRP1/2 expression by HT switched splicing regulation of many cancer-relevant genes toward a mesenchymal pattern (30,31), which is generally associated with metastatic behavior of cancer cells (54). These results highlighted an AR-dependent splicing program whose dysregulation may have important implications for disease progression. Our findings now reveal that AR activity also globally impacts on 3' end processing of pre-mRNAs (Figure 7K). However, in contrast to splicing regulation, several observations suggest a direct effect of the AR on APA. First, although most APA-regulated genes are not regulated at the transcriptional level, AR is recruited to their promoters and this association is sensitive to short-term HT treatment. Second, AR interacts with CSTF50, and Enzalutamide impaired this interaction and the association between the CSTF and CPSF complexes. HT also caused rearrangements within the CPSF complex, by enhancing the recruitment of the scaffold protein CPSF160 and the U-rich binding protein FIP1L1. Lastly, we observed that Enzalutamide treatment affected the recruitment of CPSF30 to the promoter region of APA-regulated genes, suggesting that removal of AR from these promoters favors the interaction between the transcriptional machinery and CPSF, predisposing the selection of pAs that mainly depend on this complex. To our knowledge, these findings document for the first time the regulation of APA by a transcription factor through a

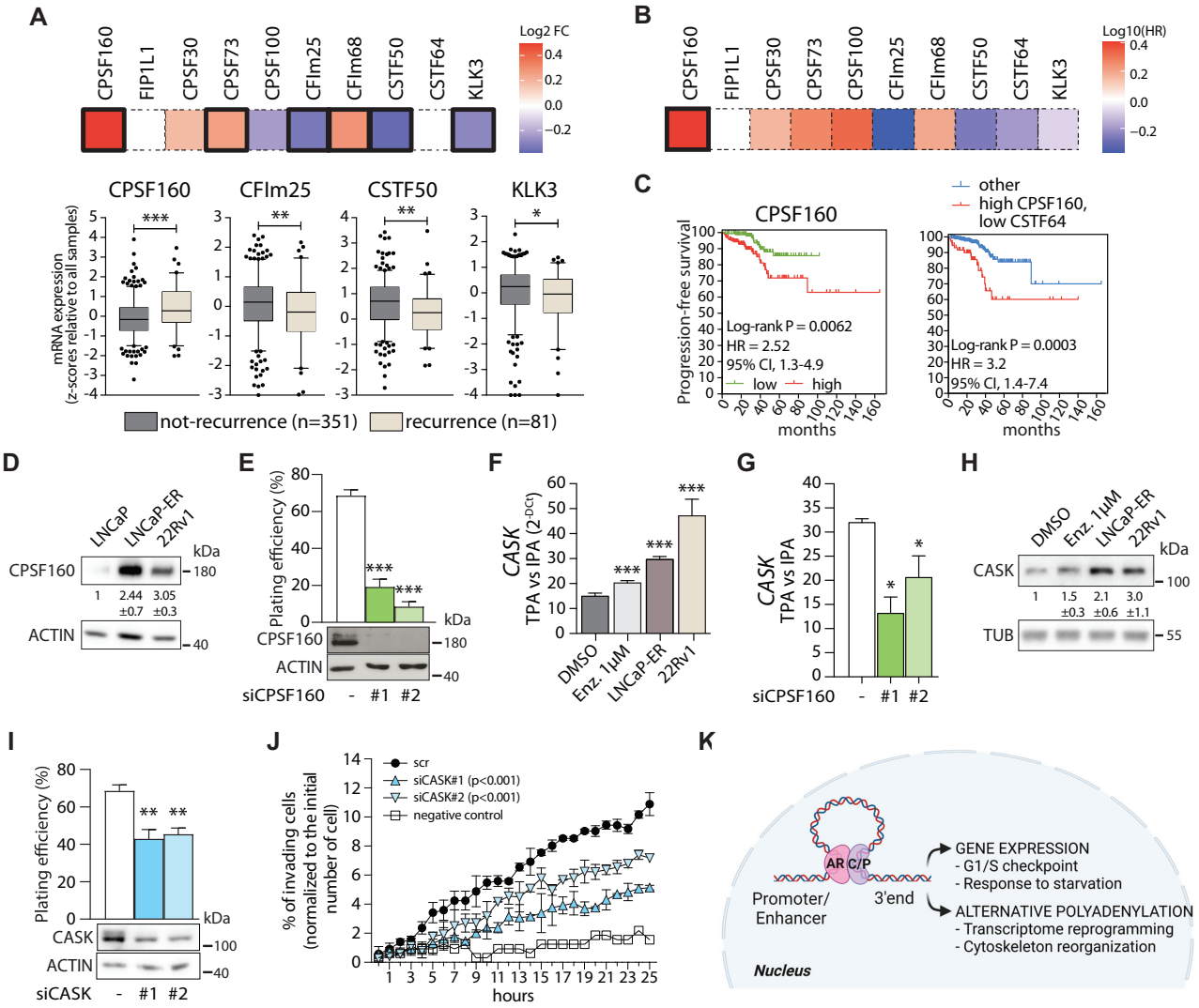

Figure 7. (**A**) Heatmap of log2-fold changes in mRNA expression of the indicated C/P factors and KLK3 in 351 PC samples with clinical response versus 81 PC samples with resistance to HT (TCGA, PanCancer Atlas) (upper panel). The thick line highlights significantly modulated factors in the two groups analyzed (P <0.05; Student's t-test). Box plots showing the normalized expression of CPSF160, CFIm25, CSTF50 and KLK3 with relative median and 5th/95th percentiles (whiskers) are shown in the lower panels (*P <0.05; **P <0.01; ***P <0.001; Student's t-test). (**B**) Correlation of expression of C/P factors with survival in PC patients. Survival heatmap shows the log10 of the hazard ratio (HR), calculated using log-rank (Mantel–Cox) test. For each gene, the HR was estimated using the progression-free survival data and the expression levels using the median threshold for splitting the high-expression and low-expression cohorts. The thick frame indicates a significant correlation (P <0.01). (**C**) Kaplan–Meier plot comparing the progression-free survival of patients with high (red) and low (green) expression of CPSF160 transcript (left) or with concomitant high CPSF160 and low CSTF64 expression (red line) with respect to all other patients (blue line; right). (**D**) Representative western blot analysis of CPSF160 protein levels in the indicated cell lines. Ratio of CPSF160 expression relative to LNCaP cells was calculated by densitometric analysis and is shown below (mean ± SD, $n \geq 3$). (**E**) Colony formation efficiency of 22Rv1 cells silenced (#1, #2) or not (–; control siRNA) for CPSF160. (**F**) qRT-PCR analysis of the CASK TPA versus IPA ratio in LNCaP cells treated or not with Enzalutamide, LNCaP-ER and 22Rv1 cells. (**G**) qRT-PCR analysis of the CASK TPA versus IPA ratio in 22Rv1 silenced (#1, #2) or not (–; control siRNA) for CPSF160. (**H**) Representative western blot analysis of CASK protein levels in the indicated cell lines. The ratio of CASK expression relative to LNCaP cells was calculated by densitometric analysis and shown below (mean ± SD, $n \geq 3$). (**I**) Colony formation efficiency of 22Rv1 cells silenced (#1, #2) or not (–; control siRNA) for CASK. (**J**) Line graph showing the time course of 22Rv1 cells, silenced or not for CASK, that invade Matrigel-coated transwells of the IncuCyte Clearview 96-well insert system and expressed as a percentage of total seeded cells. Negative control (white squares) was performed with 22Rv1 cells that were not starved to prevent their migration. The data represent the mean ± SD of two independent experiments performed in triplicate (two-way ANOVA for each siCASK relative to scr). (**K**) Schematic model of the AR-regulated transcriptional and APA programs.

mechanism that involves rearrangements of C/P subcomplexes on the gene promoter region.

Our analysis revealed that short-term HT treatment preferentially causes the selection of d-pAs with respect to IPAs, thus leading to transcript lengthening. Since selection of IPAs changes the structure of the proteins and often results in unproductive transcripts, HT may lead to increased production of proteins whose transcripts undergo lengthening. Importantly, among the genes subjected to CDS-APA regulation, some play an oncogenic role in PC. For instance, the kinase CASK was recently shown to promote PC progression (51). We now show that *CASK* transcript lengthening correlates with an increase in CASK protein levels and that its depletion in 22Rv1 cells impairs their clonogenic and invasive potential. Moreover, we observed that genes undergoing transcript lengthening are enriched in functional categories related to transcription regulation, suggesting that APA may also contribute to reprogram the PC cell transcriptome in response to HT. In this regard, it is interesting to mention that CASK was also shown to positively regulate the NOTCH pathway and other stemness pathways in other cancers (55,56). Thus, the increased CASK levels may also contribute to switch PC cells toward a stem-like phenotype in response to HT.

APA changes in cancer have been generally associated with transcript shortening, which is a feature of actively proliferating cells with respect to differentiated cells (34). Thus, transcript lengthening induced by HT could be part of the mechanisms that lead to reduced proliferation and neuroendocrine differentiation of PC cells under androgen deprivation (57). Moreover, a recent study suggested that transcript lengthening through APA characterizes the quiescent status of hematopoietic stem cells, which then switch to express shorter APA isoforms upon transition from quiescence to activation (58). Likewise, fibroblasts induced into quiescence by contact inhibition were shown to preferentially express longer transcripts through selection of d-pAs, which was associated with transcript stabilization (59). Thus, the HT-induced APA reprogramming may represent a functional mechanism set in motion by PC cells to withstand the hostile environment caused by androgen deprivation.

Computational analysis of sequence elements that characterize APA-regulated events highlighted an enrichment of U-rich motifs upstream of the up-regulated pAs and of UG-rich motifs downstream of the repressed pAs. Upstream U-rich elements are directly recognized by FIP1L1, a CPSF component that acts in pA processing together with CPSF160 (44). In addition, CPSF160 also acts as a scaffold that pre-organizes WDR33 and CPSF30 for PAS recognition (43). Thus, the increased interaction of CPSF160 and FIP1L1 with CPSF30 induced by HT may favor the selection of pAs comprising U-rich motifs (Figure 6D). Since U-rich motifs are preferentially enriched in d-pAs, their selection preferentially leads to transcript lengthening. In support of this hypothesis, Enzalutamide induced the binding of CPSF30 and FIP1L1 near to the selected pAs in LNCaP cells. It is noteworthy that upon inhibition and reduced binding of the AR, CPSF30 was also recruited more efficiently at the promoter region of these APA-regulated genes. This observation is in line with the known co-recruitment of components of the C/P machinery at promoters by RNAPII (30) and suggests a new role for the AR in coupling promoter-associated events with RNA processing regulation. In our experiments, we immunoprecipitated CPSF30 because it binds more strongly to RNAPII than other CPSF components (50), it directly contacts the PAS sequence in the pre-mRNA, and CPSF30 antibodies were very efficient under CLIP and ChIP conditions. Conversely, CPSF160 does not contact the RNA directly and cannot be efficiently cross-linked to the pre-mRNA (42). Nevertheless, it is likely that the whole CPSF complex is concomitantly recruited by RNAPII at promoters of these APA-regulated genes.

The other sequence element significantly enriched in our analysis was the UG-rich motif. UG-rich elements downstream of the CS are bound by CSTF64 and this factor globally promotes the selection of p-pAs (60). In particular, CSTF64 binds to thousands of IPAs, thus globally affecting CDS-APA regulation (46). Interestingly, CSTF64 is down-regulated in quiescent fibroblasts, concomitantly with widespread transcript lengthening, whereas its knockdown in proliferating fibroblasts was sufficient to recapitulate the selection of APA isoforms associated with quiescence (59). While we did not observe a reduction in CSTF64 expression in HT-treated cells, its interaction with the CPSF complex was significantly reduced. Thus, it is likely that dissociation of the CSTF from the CPSF complex upon AR inhibition impairs the selection of pAs that contain UG-rich motifs and that rely more strictly on CSTF function. As already observed in fibroblasts, reduced CSTF efficiency leads to preferential down-regulation of p-pAs also in PC cells.

Analysis of the expression levels of C/P components in PC datasets revealed that CPSF160 (*CPSF1* gene) is the only factor that positively correlates with both PC recurrence and progression in patients. Moreover, we found that CPSF160 was up-regulated in 22Rv1 cells and its depletion was sufficient to repress the HT-induced APA pattern in these CRPC cells. Therefore, we suggest that up-regulation of CPSF160 in CRPC cells stabilizes CPSF recruitment to U-rich pAs that depend less strictly on the CSTF function. Interestingly, CPSF160 was also shown to promote expression of AR-Vs in 22Rv1 cells through APA regulation (12), and these truncated receptor variants are strongly associated with tumor progression and resistance to therapies (6,14,15). These observations suggest that HT induces a short-term response through CPSF160-mediated APA regulation in PC cells to withstand androgen deprivation. This pathway might then be stabilized by the increased expression of CPSF160, an event that probably contributes to tumor progression and acquisition of the CRPC phenotype in PC cells and patients. CPSF160 expression was also positively correlated with tumorigenesis in other human cancers, including triple-negative breast cancer (61) and hepatocellular carcinoma (62). Our findings now suggest that global APA reprogramming is part of the oncogenic activity of CPSF160 in PC cells. Thus, CPSF160 may represent an actionable therapeutic target in CRPC to revert resistance to HT and improve clinical outcome of patients.

DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article (and its supplementary in-

formation files 1 and 2). The RNA sequencing data have been deposited in GEO (accession GSE190153).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This manuscript is dedicated to the memory of Dr Gabriele Babini. We wish to thank all members of the laboratory for fruitful discussion throughout the course of this study.
Authors' contributions: C.S., C.C. and P.B. conceived the study. C.C., M.P. and C.P. performed the experimental work. D.Z., G.B and C.C. performed the computational analysis. B.T. provided support to the experimental work. C.S and C.C. interpreted the data and wrote the manuscript. All authors approved the final version of the manuscript.

FUNDING

The research was supported by Ministero della Salute 'Ricerca Finalizzata 2011' [GR-2011–02348423 to P.B.], 'Ricerca Finalizzata 2016' [RF-2016–02363460 to C.S.] and 'Ricerca Corrente 2022' [to IRCCS Fondazione Policlinico A. Gemelli], by the Associazione Italiana Ricerca sul Cancro [AIRC IG23416 to C.S.], and by the National Institutes of Health [GM084089 and GM129069 to B.T.]. Università Cattolica del Sacro Cuore contributed to the funding of this research project and its publication. Funding for open access charge: Associazione Italiana Ricerca sul Cancro Catholic University of the Sacred Heart; Ministero della Salute Ricerca Corrente 2022.
Conflict of interest statement. None declared.

REFERENCES

1. Wang,G., Zhao,D., Spring,D.J. and Depinho,R.A. (2018) Genetics and biology of prostate cancer. *Genes Dev.*, **32**, 1105–1140.
2. Sartor,O. and de Bono,J.S. (2018) Metastatic prostate cancer. *N. Engl. J. Med.*, **378**, 645–657
3. Watson,P.A., Arora,V.K. and Sawyers,C.L. (2015) Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. *Nat. Rev. Cancer*, **15**, 701–711.
4. Schmidt,K.T., Huitema,A.D.R., Chau,C.H. and Figg,W.D. (2021) Resistance to second-generation androgen receptor antagonists in prostate cancer. *Nat. Rev. Urol.*, **18**, 209–226.
5. Scher,H.I., Fizazi,K., Saad,F., Taplin,M.-E., Sternberg,C.N., Miller,K., de Wit,R., Mulders,P., Chi,K.N., Shore,N.D. *et al.* (2012) Increased survival with enzalutamide in prostate cancer after chemotherapy. *N. Engl. J. Med.*, **367**, 1187–1197.
6. Paschalis,A., Sharp,A., Welti,J.C., Neeb,A., Raj,G.V., Luo,J., Plymate,S.R. and de Bono,J.S. (2018) Alternative splicing in prostate cancer. *Nat. Rev. Clin. Oncol.*, **15**, 663–675.
7. Dong,X. and Chen,R. (2020) Understanding aberrant RNA splicing to facilitate cancer diagnosis and therapy. *Oncogene*, **39**, 2231–2242.
8. Sette,C. (2013) Alternative splicing programs in prostate cancer. *Int. J. Cell Biol.*, **2013**, 458727–458737.
9. Burd,C.J., Petre,C.E., Morey,L.M., Wang,Y., Revelo,M.P., Haiman,C.A., Lu,S., Fenoglio-Preiser,C.M., Li,J., Knudsen,E.S. *et al.* (2006) Cyclin D1b variant influences prostate cancer growth through aberrant androgen receptor regulation. *Proc. Natl Acad. Sci. USA*, **103**, 2190–2195.
10. Olshavsky,N.A., Comstock,C.E.S., Schiewer,M.J., Augello,M.A., Hyslop,T., Sette,C., Zhang,J., Parysek,L.M. and Knudsen,K.E. (2010) Identification of ASF/SF2 as a critical, allele-specific effector of the cyclin D1b oncogene. *Cancer Res.*, **70**, 3975–3984.
11. Paronetto,M.P., Cappellari,M., Busà,R., Pedrotti,S., Vitali,R., Comstock,C., Hyslop,T., Knudsen,K.E. and Sette,C. (2010) Alternative splicing of the cyclin D1 proto-oncogene is regulated by the RNA-binding protein sam68. *Cancer Res.*, **70**, 229–239.
12. Van Etten,J.L., Nyquist,M., Li,Y., Yang,R., Ho,Y., Johnson,R., Ondigi,O., Voytas,D.F., Henzler,C. and Dehm,S.M. (2017) Targeting a single alternative polyadenylation site coordinately blocks expression of androgen receptor mRNA splice variants in prostate cancer. *Cancer Res.*, **77**, 5228–5235.
13. Henzler,C., Li,Y., Yang,R., McBride,T., Ho,Y., Sprenger,C., Liu,G., Coleman,I., Lakely,B., Li,R. *et al.* (2016) Truncation and constitutive activation of the androgen receptor by diverse genomic rearrangements in prostate cancer. *Nat. Commun.*, **7**, 13668–13680.
14. Antonarakis,E.S., Lu,C., Wang,H., Luber,B., Nakazawa,M., Roeser,J.C., Chen,Y., Mohammad,T.A., Chen,Y., Fedor,H.L. *et al.* (2014) AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N. Engl. J. Med.*, **371**, 1028–1038.
15. Zhu,Y., Dalrymple,S.L., Coleman,I., Zheng,S.L., Xu,J., Hooper,J.E., Antonarakis,E.S., De Marzo,A.M., Meeker,A.K., Nelson,P.S. *et al.* (2020) Role of androgen receptor splice variant-7 (AR-V7) in prostate cancer resistance to 2nd-generation androgen receptor signaling inhibitors. *Oncogene*, **39**, 6935–6949.
16. Tian,B. and Manley,J.L. (2016) Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.*, **18**, 18–30.
17. Hoque,M., Ji,Z., Zheng,D., Luo,W., Li,W., You,B., Park,J.Y., Yehia,G. and Tian,B. (2013) Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. *Nat. Methods*, **10**, 133–139.
18. Zheng,D., Liu,X. and Tian,B. (2016) 3′READS+, a sensitive and accurate method for 3′ end sequencing of polyadenylated RNA. *RNA*, **22**, 1631–1639.
19. Wang,R., Nambiar,R., Zheng,D. and Tian,B. (2018) PolyA-DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res.*, **46**, D315–D319.
20. Liu,J., Lichtenberg,T., Hoadley,K.A., Poisson,L.M., Lazar,A.J., Cherniack,A.D., Kovatich,A.J., Benz,C.C., Levine,D.A., Lee,A.V. *et al.* (2018) An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*, **173**, 400–416.
21. Grasso,C.S., Wu,Y.M., Robinson,D.R., Cao,X., Dhanasekaran,S.M., Khan,A.P., Quist,M.J., Jing,X., Lonigro,R.J., Brenner,J.C. *et al.* (2012) The mutational landscape of lethal castration-resistant prostate cancer. *Nature*, **487**, 239–243.
22. Ross-Adams,H., Lamb,A., Dunning,M., Halim,S., Lindberg,J., Massie,C., Egevad,L., Russell,R., Ramos-Montoya,A., Vowler,S. *et al.* (2015) Integration of copy number and transcriptomics provides risk stratification in prostate cancer: a discovery and validation cohort study. *EBioMedicine*, **2**, 1133–1144.
23. Paronetto,M.P., Achsel,T., Massiello,A., Chalfant,C.E and Sette,C. (2007) The RNA-binding protein sam68 modulates the alternative splicing of Bcl-x. *J. Cell Biol.*, **176**, 929–939.
24. Caggiano,C., Pieraccioli,M., Panzeri,V., Sette,C. and Bielli,P. (2019) C-MYC empowers transcription and productive splicing of the oncogenic splicing factor sam68 in cancer. *Nucleic Acids Res.*, **47**, 6160–6171.
25. Bielli,P. and Sette,C. (2017) Analysis of in vivo interaction between RNA binding proteins and their RNA targets by UV cross-linking and immunoprecipitation (CLIP) method. *Bio-Protocol*, **7**, e2274.
26. Shin,J., Ding,Q., Wang,L., Cui,Y., Baljinnyam,E., Guvenek,A. and Tian,B. (2022) CRISPRpas: programmable regulation of alternative polyadenylation by dCas9. *Nucleic Acids Res.*, **50**, e25.
27. Korenchuk,S., Lehr,J.E., McLean,L., Lee,Y.G., Whitney,S., Vessella,R., Lin,D.L. and Pienta,K.J. (2001) VCaP, a cell-based model system of human prostate cancer. *In Vivo (Brooklyn)*, **15**, 163–168.
28. Munkley,J., Li,L., Krishnan,S.R.G., Hysenaj,G., Scott,E., Dalgliesh,C., Oo,H.Z., Maia,T.M., Cheung,K., Ehrmann,I. *et al.* (2019) Androgen-regulated transcription of esrp2 drives alternative splicing patterns in prostate cancer. *Elife*, **8**, e47678.
29. Shah,K., Gagliano,T., Garland,L., O'Hanlon,T., Bortolotti,D., Gentili,V., Rizzo,R., Giamas,G. and Dean,M. (2020) Androgen receptor signaling regulates the transcriptome of prostate cancer cells by modulating global alternative splicing. *Oncogene*, **39**, 6172–6189.
30. Hsin,J.P. and Manley,J.L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.*, **26**, 2119–2137.

31. Glover-Cutter,K., Kim,S., Espinosa,J. and Bentley,D.L. (2008) RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat. Struct. Mol. Biol.*, **15**, 71–78.
32. Nojima,T., Gomes,T., Grosso,A.R.F., Kimura,H., Dye,M.J., Dhir,S., Carmo-Fonseca,M. and Proudfoot,N.J. (2015) Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell*, **161**, 526–540.
33. Kuleshov,M.V., Jones,M.R., Rouillard,A.D., Fernandez,N.F., Duan,Q., Wang,Z., Koplev,S., Jenkins,S.L., Jagodnik,K.M., Lachmann,A. *et al.* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.*, **44**, 90–97.
34. Gruber,A.J. and Zavolan,M. (2019) Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.*, **20**, 599–614.
35. Martin,G., Gruber,A.R., Keller,W. and Zavolan,M. (2012) Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep.*, **1**, 753–763.
36. Dantonel,J.C., Murthy,K.G.K., Manjey,J.L. and Tora,L. (1997) Transcription factor TFIID recruits factor CPSF for formation of 3 end of mRNA. *Nature*, **389**, 399–402.
37. Wang,Y., Fairley,J.A. and Roberts,S.G.E. (2010) Phosphorylation of TFIIB links transcription initiation and termination. *Curr. Biol.*, **20**, 548–553.
38. Davidson,L., Muniz,L. and West,S. (2014) 3' end formation of pre-mRNA and phosphorylation of ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev.*, **28**, 342–356.
39. Zhao,Y., Wang,L., Ren,S., Wang,L., Blackburn,P.R., McNulty,M.S., Gao,X., Qiao,M., Vessella,R.L., Kohli,M. *et al.* (2016) Activation of P-TEFb by androgen receptor-regulated enhancer RNAs in castration-resistant prostate cancer. *Cell Rep.*, **15**, 599–610.
40. Heemers,H.V. and Tindall,D.J. (2007) Androgen receptor (AR) coregulators: a diversity of functions converging on and regulating the AR transcriptional complex. *Endocr. Rev.*, **28**, 778–808.
41. Stelloo,S., Nevedomskaya,E., Kim,Y., Hoekman,L., Bleijerveld,O.B., Mirza,T., Wessels,L.F.A., Van Weerden,W.M., Altelaar,A.F.M., Bergman,A.M. *et al.* (2018) Endogenous androgen receptor proteomic profiling reveals genomic subcomplex involved in prostate tumorigenesis. *Oncogene*, **37**, 313–322.
42. Chan,S.L., Huppertz,I., Yao,C., Weng,L., Moresco,J.J., Yates,J.R., Ule,J., Manley,J.L. and Shi,Y. (2014) CPSF30 and wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. *Genes Dev.*, **28**, 2370–2380.
43. Clerici,M., Faini,M., Aebersold,R. and Jinek,M. (2017) Structural insights into the assembly and polyA signal recognition mechanism of the human CPSF complex. *Elife*, **6**, e33111.
44. Kaufmann,I., Martin,G., Friedlein,A., Langen,H. and Keller,W. (2004) Human fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.*, **23**, 616–626.
45. Lackford,B., Yao,C., Charles,G.M., Weng,L., Zheng,X., Choi,E.A., Xie,X., Wan,J., Xing,Y., Freudenberg,J.M. *et al.* (2014) Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J.*, **33**, 878–889.
46. Yao,C., Biesinger,J., Wan,J., Weng,L., Xing,Y., Xie,X. and Shi,Y. (2012) Transcriptome-wide analyses of CstF64–RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc. Natl Acad. Sci. USA*, **109**, 18773–18778.
47. Yang,W., Hsu,P.L., Yang,F., Song,J.E. and Varani,G. (2018) Reconstitution of the CstF complex unveils a regulatory rolefor cstf-50 in recognition of 3-end processing signals. *Nucleic Acids Res.*, **46**, 493–503.
48. Di Giammartino,D.C., Nishida,K. and Manley,J.L. (2011) Mechanisms and consequences of alternative polyadenylation. *Mol. Cell*, **43**, 853–866.
49. Nunes,N.M., Li,W., Tian,B. and Furger,A. (2010) A functional human poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J.*, **29**, 1523–1536.
50. Nag,A., Narsinh,K. and Martinson,H.G. (2007) The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. *Nat. Struct. Mol. Biol.*, **14**, 662–669.
51. Zhang,W., Liu,Q., Zhao,J., Wang,T. and Wang,J. (2021) Long noncoding RNA AATBC promotes the proliferation and migration of prostate cancer cell through miR-1245b-5p/CASK axis. *Cancer Manag. Res.*, **13**, 5091–5100.
52. Mills,I.G. (2014) Maintaining and reprogramming genomic androgen receptor activity in prostate cancer. *Nat. Rev. Cancer*, **14**, 187–198.
53. Sharma,N.L., Massie,C.E., Ramos-Montoya,A., Zecchini,V., Scott,H.E., Lamb,A.D., MacArthur,S., Stark,R., Warren,A.Y., Mills,I.G. *et al.* (2013) The androgen receptor induces a distinct transcriptional program in castration-resistant prostate cancer in man. *Cancer Cell*, **23**, 35–47.
54. Pradella,D., Naro,C., Sette,C. and Ghigna,C. (2017) EMT and stemness: flexible processes tuned by alternative splicing in development and cancer progression. *Mol. Cancer*, **16**, 8.
55. Qu,J., Zhou,Y., Li,Y., Yu,J. and Wang,W. (2021) CASK regulates notch pathway and functions as a tumor promoter in pancreatic cancer. *Arch. Biochem. Biophys.*, **701**, 108789.
56. Farahani,D.B., Akrami,H., Moradi,B., Mehdizadeh,K. and Fattahi,M.R. (2021) The effect of hsa-miR-451b knockdown on biological functions of gastric cancer stem-like cells. *Biochem. Genet.*, **59**, 1203–1224.
57. Yuan,T.C., Veeramani,S. and Lin,M.F. (2007) Neuroendocrine-like prostate cancer cells: neuroendocrine transdifferentiation of prostate adenocarcinoma cells. *Endocr. Relat. Cancer*, **14**, 531–547.
58. Sommerkamp,P., Altamura,S., Renders,S., Narr,A., Ladel,L., Zeisberger,P., Eiben,P.L., Fawaz,M., Rieger,M.A., Cabezas-Wallscheid,N. *et al.* (2020) Differential alternative polyadenylation landscapes mediate hematopoietic stem cell activation and regulate glutamine metabolism. *Cell Stem Cell*, **26**, 722–738.
59. Mitra,M., Johnson,E.L., Swamy,V.S., Nersesian,L.E., Corney,D.C., Robinson,D.G., Taylor,D.G., Ambrus,A.M., Jelinek,D., Wang,W. *et al.* (2018) Alternative polyadenylation factors link cell cycle to migration. *Genome Biol.*, **19**, 176.
60. Hwang,H.W., Park,C.Y., Goodarzi,H., Fak,J.J., Mele,A., Moore,M.J., Saito,Y. and Darnell,R.B. (2016) PAPERCLIP identifies microRNA targets and a role of cstf64/64tau in promoting non-canonical poly(A) site usage. *Cell Rep.*, **15**, 423–435.
61. Wang,L., Lang,G.T., Xue,M.Z., Yang,L., Chen,L., Yao,L., Li,X.G., Wang,P., Hu,X., Shao,Z.M. *et al.* (2020) Dissecting the heterogeneity of the alternative polyadenylation profiles in triple-negative breast cancers. *Theranostics*, **10**, 10531–10547.
62. Chen,S.L., Zhu,Z.X., Yang,X., Liu,L.L., He,Y.F., Yang,M.M., Guan,X.Y., Wang,X. and Yun,J.P. (2021) Cleavage and polyadenylation specific factor 1 promotes tumor progression via alternative polyadenylation and splicing in hepatocellular carcinoma. *Front. Cell Dev. Biol.*, **9**, 616835.