SI appendix for

Dysregulation of spliceosome gene expression in advanced prostate cancer by RNA-binding protein PSF

Ken-ichi Takayama, Takashi Suzuki, Tetsuya Fujimura, Yuta Yamada, Satoru Takahashi, Yukio Homma, Yutaka Suzuki, Satoshi Inoue.

SI text 1

PSF regulates the cell cycle inhibitors and tumor suppressive genes transcriptionally

To confirm the function of PSF as a transcriptional regulator in CRPC, we performed PSF chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) in LTAD cells (SI Appendix, Fig. S3A). By using ChIP-seq in LNCaP cells, we previously showed that PSF transcriptionally regulated cell cycle inhibitors such as p53 and that its association with DNA was dependent on *CTBP1-AS* (1). Numbers of global PSF-binding sites (PSFBSs) at DNA level were decreased by *CTBP1-AS* depletion, although PSF recruitments were enhanced in some genomic regions, suggesting both inhibitory and promoting functions of *CTBP1-AS* in PSF bindings (SI Appendix, Fig. S3B, C). In these PSF-binding genes at DNA level, we found tumor suppressor genes, such as CDKN1A/p21, CDKN1B/p27, CDKN2B/p15 and p53, were transcriptional targets of PSF/histone deacetylase 1 (HDAC1) complex in CRPC cells (SI Appendix, Fig. S3D, E). We hypothesized that PSF represses tumor suppressor genes by HDAC-mediated transcriptional regulation. This is supported by our pathway analysis using DAVID (2) which revealed that PSF represses cancer-associated genes such as p53 signal and cell cycle-associated genes (SI Appendix, Fig. S4A-C) transcriptionally through binding to histone modified regions or HDAC1 interaction (SI Appendix, Fig. S4D, E). In addition, PSFBSs partially overlapped with AR binding sites (ARBSs) (1.7-8%) (SI Appendix, Fig. S5), suggesting a possible role of direct PSF action on AR activity in a small subset of

ARBSs. Taken together, we showed that PSF mainly regulates enhancer or promoters epigenetically to repress nearby tumor suppressor gens including cell cycle regulators in prostate cancer as a transcriptional repressor.

SI text 2

The different role of PSF as a transcriptional regulator from as an RNA-binding protein

We further compared and examined the overlap of PSF targets at RNA and DNA levels identified in the present global study. In LNCaP cells, we found 261 genes out of 1139 (22.9%) PSF target genes at RNA level were overlapped with those at DNA level (SI Appendix, Fig. S9A). Our analysis using RNA-seq data in LNCaP cells showed that PSF target genes at DNA level were regulated negatively by PSF recruitment to their promoters generally (SI Appendix, Fig. S9B), consistent with the notion that PSF mediates transcriptional repression by interaction with HDAC complex. Among PSF-binding genes overlapped with androgen-mediated deacetylated regions, significant enrichment of PSF-mediated gene repression with androgen treatment (SI Appendix, Fig. S9C left). However, induction of gene expression by PSF recruitment was also observed, suggesting a possible role of PSF for transcriptional activation (SI Appendix, Fig. S9B left and S9C right). Marked overlap of PSFBSs with activated histone maeks (such as K4me3 and AcH3) suggests that these roles of PSF in transcriptional regulation by repressing or activating these epigenetic status (SI Appendix, Fig. S4D, E). On the other hand, the role of PSF on the expressions of overlapped genes between

PSF-binding genes at RNA and DNA level is mainly positive, suggesting the importance of the PSF function as an RNA-binding protein for the expression of this gene set (SI Appendix, Fig. S9B, C).

SI Materials and Methods

ChIP-seq

ChIP was performed as described previously (1, 3), using PSF (Sigma), HDAC1 (Santa cruz) and AcH3 (Millipore) antibodies. Briefly, cells were crosslinked for 10 minutes with 1% formaldehyde and the crosslinking was inactivated with 0.125 M glycine for 5 minutes at room temperature. Cells were lysed and sonicated to an average of 500 bp. Chromatin were incubated with antibody overnight at 4℃. Beads were added and incubated for 2 h, washed and eluted in elution buffer. Reverse crosslinking was performed by incubating the eluted product with 0.3 M NaCl at 65℃ overnight. ChIP products were precipitated with ethanol. ChIP-seq analyses were performed using an Illumina Genome Analyzer or HiSeq platform (Illumina, San Diego, CA). Libraries were prepared according to the manufacturer's instructions. Unfiltered 36 bp sequence readings were aligned to the human reference genome (hg19) using CASAVA v1.7 (Illumina). Signal scores for significant bindings were calculated using model-based analysis for ChIP-Seq (MACS) (4) and the threshold for AR binding sites was P-value < 10^{-5}.

RIP and RIP-seq

RIP was performed using EZ-magna RIP RNA binding protein immunoprecipitation kit (Millipore

#17-701) according to the manufacturer's protocol. Briefly, nuclei of prostate cancer cells were lysed in lysis buffer. The lysate was diluted and incubated using PSF antibodies or IgG magnetic beads at 4℃ overnight. The RNA/antibody complex was washed four times. The RNA was extracted using acid-phenol:chloroform and subjected to qRT-PCR or sequencing. Library construction was performed using sureselect Strand Specific RNA system (Agilent Technologies, Santa Clara, CA). Sequencing was performed by illumine HiSeq 2500. To remove rRNA sequence, we used bowtie 2 version 2.2.6. Mapping to human genome (hg19) was performed using tophat 2.1.0.with bowtie. Alignments were generated in the SAM format from given single-end reads. The reads were mapped to the human RefSeq mRNA database (to exons, introns and antisense regions) or GENCODE/NONCODE database. The expression levels of mapped transcripts were normalized into reads per kilobase of exon per million mapped reads (RPKM) to facilitate comparison among different samples. Fisher exact test was performed to determine the difference of PSF-IP and input samples statistically. Fold enrichment was calculated and identified enriched transcripts (Fold > 2.0, *P*-value < 10^{-5}). For RIP assay, we used ABI step one Real-Time PCR system and used GAPDH for normalization and Myoglobin (MB) for negative control (1).

RNA sequencing

RNA sequencing (RNA-seq) library construction and sequencing were performed according to the manufacturer's protocol for the Applied Biosystems SOLiD 4 System (Applied Biosystems). The

library was sequenced using SOLiD Opti Fragment Library Sequencing kit Master Mix 50 chemistry (Applied Biosystems), which gives 50 bp readings. The RNA-seq readings were analyzed using whole transcriptome software tools from Applied Biosystems. Matching locations were subsequently used to generate counts for annotated features, exons, transcripts or genes of RefSeq. Finally, gene expression was determined as RPKM.

qRT-PCR

Total RNA was isolated using the ISOGEN reagent (Nippon gene, Tokyo, Japan). First-strand cDNA was generated using a PrimeScript RT reagent kit (Takara, Kyoto, Japan). Resulting cDNA was then analyzed by quantitative PCR (qPCR) using KAPA SYBR green mix (KAPA) on Step one Real-Time PCR system. The primer sequences are listed in SI Appendix, Table S1 or previously described (1, 5).

miRNA quantitative PCR

Total RNA was extracted using a mirVana miRNA Isolation Kit (Ambion). RNA (10 ng) was reverse transcribed to cDNA using a TaqMan micro-RNA reverse transcription kit (Applied Biosystems). Quantitative PCR (qPCR) for cDNA was performed using TaqMan microRNA assay primers and a Step One Real-Time PCR System (Applied Biosystems). The relative level of miRNA was calculated by normalizing the target gene signal by the signal for U6B small nuclear RNA.

siRNA

siRNAs targeting *PSF (PSF #1) and NONO (NONO #1)* was as described (1). We designed and

purchased additional siPSF#2 (5'-GGCACGUUUGAGUACGAAUAU-3') from Sigma Genosys Japan. Silencer select siRNA targeting NONO (NONO #2:s9613), splicing factors (U2AF2:s22364, SF1: s194850, DDX23: s18026, CHERP: s20627, SF3B3: s23848, SF3B2: s21643, ACIN1: s22777, U2SURP: s23622, PRPF3: s17435, HNRNPU: s6744) were also purchased from Thermo fisher. All siRNA experiments were performed at the concentration of 10 nM. Cells were transfected with siRNAs using Lipofectamine RNAiMAX (Thermofisher) 48 - 72 h before each experiment.

Immunoprecipitation and Western blot analysis

Whole cell lysates were prepared using lysis buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 1% NP-40, protease inhibitor cocktail). Protein concentration was determined by the Bradford assay (Bio-Rad Laboratories, Tokyo, Japan). Fifty micrograms of each protein lysate was loaded onto SDS-polyacrylamide gels, separated by electrophoresis, and electrotransferred onto Immobilon-P Membranes (Millipore, Billerica, MA). Membranes were incubated with primary antibodies overnight and then incubated with secondary antibodies. Antibody-antigen complexes were detected using ECL Prime Western Blotting Detection Reagents (GE Healthcare Life Sciences, Tokyo, Japan).

Immunohistochemistry

Formalin-fixed tissues were embedded in paraffin and sectioned. A Histofine kit (Nichirei, Tokyo, Japan), which employs the streptavidin-biotin amplification method, was used for immunohistochemical analysis of PSF (61045, NOVUS), NONO (611278, BD bioscience) and

AR-V7 (198394, Abcam). Antigen retrieval was performed by heating the slides in an autoclave at 120°C for 5 min in citric acid buffer (2 mM citric acid and 9 mM trisodium citrate dehydrate, pH 6.0) for PSF and NONO and in microwave for 20 min in H buffer, pH7.0 (LSI Medience) for AR-V7. Dilution of primary antibody used was 1/8,000 (PSF), 1.2500 (NONO), and 1/50 (AR-V7) in this study. The antigen-antibody complex was visualized with a 3,3'-diaminobenzidine solution (1 mM 3,3'-diaminobenzidine, 50 mM Tris-HCl buffer, pH 7.6, and 0.006% H_2O_2). During immunohistochemical analysis, immunoreactivity was evaluated in more than 1,000 carcinoma cells in each case and the percentage of immunoreactivity (labeling index (LI)) was determined by specialized pathologists. It was difficult for us to perform immunohitochemistry for AR-V7 in this study, which was partly due to the fact that we could not find an appropriate positive control tissue. To overcome this problem, we used a cell block of 22Rv1 cells, which highly expressed AR-V7 as shown in Figure 3C, from formalin-fixed and paraffin-embedded specimens, and determined the optimal condition of the AR-V7 immunostaining.

Cell culture and reagents

VCaP and 293T cells were grown in DMEM medium supplemented with 10% fetal bovine serum (FBS), 50 U/mL penicillin, and 50 μg/mL streptomycin. 22Rv1, DU145 and LNCaP cells were grown in RPMI medium supplemented with 10% FBS, 50 U/mL penicillin, and 50 μg/mL streptomycin. PrEC cells were obtained from Lonze as described. LTAD cells were grown in phenol red free RPMI

medium supplemented with 10% charcoal-dextran stripped FBS, 50 U/mL penicillin, and 50 μg/mL streptomycin. All cell lines were validated for identity by short tandem repeat (STR) analysis and routinely checked for Mycoplasma contamination. The antibodies used in this study were AR-V7 (198394), DDX23 (70461) from abcam, AcH3(06599) from Millipore, NONO (611278) from BD bioscience, AR-full (441), U2AF2 (MC3), HNRNPU (3G6), p53 (Do1), p27(C-19 sc528), p21 (C-19 sc-397), HDAC1 (H51) from Santa cruz, rabbit polyclonal PSF (61045) from NOVUS, mouse monoclonal PSF (WH0006421) and β-actin from Sigma. Other antibodies and reagents used have been previously described (1).

CLIP-seq

LNCaP cells were crosslinked using UV crosslinker (150 mJ/cm2). Beads were prepared by mixing protein G beads with PSF specific antibodies. Harvested cells were lysed in lysis buffer (50mM Tris-HCl, PH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, proteinase inhibitors) and added RNase dilution and DNase to digest RNA and DNA. Cell lysates were immunoprecipitated by prepared protein G beads overnight. Beads were washed and eluted by incubation at 70℃ for 10 min in SDS-PAGE sample buffer. Samples were loaded onto a SDS-PAGE electrophoresis. Then the protein-RNA complexes were transferred to a nitrocellulose membrane. The PSF protein band was confirmed by western blotting and protein-RNA complexes at higher molecular weight than PSF were cut and isolated using PK buffer (100 mM Tris-HCl PH 7.4, 50 mM NaCl, 10

mM EDTA) and protein kinase K (10μL). Extracted RNA was collected for library construction. Briefly, ligated RNA was reverse transcribed and followed by PCR. PCR product was analyzed and isolated DNA was extracted for sequencing. We also performed this study by using normal mouse IgG as a negative control. Overlapped peaks between PSF-IP and IgG-IP are excluded from PSF targets. For analysis of CLIP-seq we used MACS for visualizing and confirming the peak positions and PIPE-CLIP (6) to obtained enriched clusters with reliable mutations. This study was performed in biological duplicate samples and identified common target genes as PSF target genes.

Immunofluorescence

Cells grown on 12-mm circular coverslips (Matsunami, Tokyo, Japan) in 24-well plates were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) for 15 min at room temperature and permeabilized with 0.2% Triton X-100/PBS for 10 min. Cells were washed in PBS and blocked in 5% normal goat serum/PBS for 30 min. Cells were first incubated with primary antibodies in 5% normal goat serum/PBS overnight at 4°C. Cells were washed 3 times with PBS and incubated with anti-mouse IgG conjugated to Alexa Fluor 546 and anti-rabbit IgG conjugated to Alexa Fluor 488 (Life Technologies) in goat serum/PBS for 1 h. Nuclei were counterstained with 4', 6-diamidino-2-phenylindole (DAPI). Cells were washed 3 times with PBS, coverslips were mounted in glycerol, and cells were visualized using an Olympus confocal laser scanning microscope FV10i (Olympus, Tokyo, Japan).

Clinical tissue samples

We obtained 102 prostate cancer samples from surgeries performed at the University of Tokyo Hospital (Tokyo, Japan). The University of Tokyo ethics committee approved our study (G10044-2), and informed consent was obtained from each patient before surgery. The ages of the patients ranged from 52 to 78 years (mean, 67 years), and pretreatment serum prostate-specific antigen (PSA) levels ranged from 1.2 to 136 ng/mL (mean, 16.9 ng/mL). The prostate tissue sections submitted for this study contained 95 benign and 102 cancerous foci. Other clinicopathological parameters are shown in SI Appendix, Table S2. We prepared RNA samples by surgeries performed at the University of Tokyo Hospital (Tokyo, Japan). RNA was obtained by laser capture microdisection (LCM) as described (1, 7) from prostate cancer and benign prostate tissues. Tumor samples of CRPC patients (from prostate, bone, liver and lymph node) were obtained by Computed tomography (CT) guided biopsy or rapid pathological anatomy (5 samples from 4 patients). Tumor samples were homogenized in ISOGEN for RNA extraction.

Cell proliferation assay

Cells were plated at 3×10^3 cells per well in 96-well plates. For RNAi experiments, cells were transfected with siRNA 24 h after plating. The MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt] assays were performed using the CellTiter 96® Aqueous MTS reagent (Promega) according to the

manufacturer's instructions. The experiments were performed in quintuplicate.

***In vivo* tumor formation assay**

LNCaP or 22Rv1 (3×10^6) cells were harvested and mixed with an equal volume of Matrigel (BD Biosciences, Bedford, MA). Cells were subcutaneously injected into one side of twenty 5-week-old male BALB/C nude mice. When the tumor volumes reached 100 mm^3, the mice were randomly divided into two groups. For the experiment of 22Rv1 tumor, we performed castration to reduce androgen concentration in mice. Each tumor was injected intratumorally with siPSF or siControl (5 μg) mixed with Lipofectamine RNAi MAX Transfection Reagent (10 μL) (Thermofisher, Carlsbad, CA) in 50 μL OPTI-MEM three times per week. The tumor dimensions were measured before each injection using a caliper. Tumor volume was determined using the following formula: $V = 0.5 \times r1 \times r2 \times r3$ ($r1 < r2 < r3$). The experimenter was blinded to the group allocation during the experiment and when assessing the outcome. This study protocol was approved by the ethic committee of animal experiments in the University of Tokyo. After the monitoring the tumor sizes for three weeks or exceeded the limit of 20 mm in one dimension, mice were killed for obtaining tumor samples. Tumors were homogenized in 1 × Radioimmunoprecipitation (RIPA) buffer with protease inhibitor cocktail (Nakarai, Tokyo, Japan). The protein concentration was measured by Pierce bicinchoninic acid (BCA) assay kit (Thermo Fisher) and used for Western blotting analysis.

ChIP qPCR

Quantitative PCR (qPCR) for ChIPed DNA was performed using KAPA SYBR green mix (KAPA) on Step One Real-Time PCR System (Applied Biosystems). The fold enrichment relative to input samples (% input) was calculated. *MB* locus was used as a negative control. Primer sequences for PSF binding sites were listed in SI Appendix, Table S3.

Microarray

The GeneChip Human Gene 1.0 ST Array (Affymetrix, Santa Clara, CA) was used according to the manufacturer's protocol for gene expression microarrays. Data analysis was performed using the Affymetrix Microarray Suite software. To compare arrays, normalization was performed on data from all probe sets.

Statistical analyses

For the cell proliferation assay, the average of 4 wells was analyzed. For cell line experiments, statistical differences (*P* values) among groups were obtained using the two-sided Student's *t*-test. All experiments were performed at least twice and similar results were obtained. *P* values less than 0.05 were considered statistically significant. Statistical analysis was performed using Graphpad Prism 5 software (GraphPad Software, San Diego, CA) or Microsoft Excel.

The association between immunoreactivity and clinicopathological factors was evaluated using the Student's *t*-test, cross-table using the chi square-test, or the correlation coefficient (*r*) and regression

equation. A survival curve was generated according to the Kaplan-Meier method, and statistical significance was calculated using the log-rank test.

Data accession

Microarray, RIP-seq, CLIP-seq, RNA-seq and ChIP-seq data are deposited in GEO (GSE94577, GSE94028, GSE94243, GSE100239).

Table S1. qPCR primers

	Forward (5'-3')	Reverse (5'-3')
n376985	TGTGAGGTGTGCATTGGATT	CAAGCGTTTGTCCATCATGT
n342314	TGTCTTCATTGCCAGTG	CTTTTGCAGAAACGCACGTA
MALAT1	TGTGTGCCAATGTTTCGTTT	AGGAGAAAGTGCCATGGTTG
miR99AHG	TGGATATGAAAAGTGAAGAAAGTGA	CATTCTGACCTCAGCCTCCT
SchLAP1	GGAGTTCACCAATGATGAGGA	TCACCGGACTCTGGTTCAAT
PPP3CA	AAGGCGCATCTTATGAAGGA	TTTTTCCCCAACAAATGGAA
CAMKK2	TCTCACCACGTCTCCATCAC	TTTGGACAGCACCTTCATTG
NONO	AACCACATTTCCTCGTCCTG	TTCTCACGAGCCTCCTTGAT
U2AF1	TGGAGGAGAAGTATGGGGAAG	ATCGGCTGTCCATTAAACCA
U2AF2	GGACGACGAGGAGTATGAGG	GGTGATAAGAGTCGGGGTCA
SF3B3	ACCTTTGGCAAGAGTGGATG	CCAGGGGAGATGAAATGGTA
SF3B2	GCTGAAGTTGGCTCAGCAG	ATCTCCTGCTGTCGTTCCTG
SNRPA	GGTATCACGGAGCCTGAAGA	GTGCCTTTCATCTTGGCAAT
SNRPF	CCTGCGGTACCTGCTGTAGT	CCTTGTACTCCATTCCCCACT
SRSF4	GTTACGGTTCTGGACGCAGT	ATTTTTGCGTCCCTTGTGAG
SRSF6	TCAAAAATGGGTACGGCTTC	GCTGTATCCACCTCCACCAC
SF1	CAAGCCACCTGCAGATTACA	GCTCATCTTCTCCTGGCAAC
ACIN1	GCCACCACACAGAAGAAACC	ATCGCCATTACGCTCTGTCT
TRA2A	GGACCATTGAGTGGTGTCAA	CCACCCGAATTCTTCTACCA
EFTUD2	ATCCCCTGGACAAGAGCATT	TGGGGTAATTGAGCACAACA
PRPF40B	GAGAAGACAGGCTGGGACAC	AGTCTGAGGCAGACGGTGAG
DDX39B	CGCTCCAACTCAAGATCCTC	TTCCTGGGGAGTACGACTTG
DDX23	ATGTGGCTGGTCGTGGTATT	CAGGAAGACACTGGGCTTTC
HNRNPU	CTACCAGAACATGCGGTCCT	CTGGTGGAAGAGCCTTTTTG
PRPF3	GCATCGGATAAAGTGGGATG	TGCTTTTTGAAATGCTCACG
CHERP	CGCTCCAACTCAAGATCCTC	TTCCTGGGGAGTACGACTTG
U2SURP	GAGGCCACCACCATTAAATC	CAGAAAACCATTGCATCTCC
SchLAP1 intron#1	GGAACACAGGTGGAGGTTTG	AGCCAAGTTCCCTTCCTTTT
#2	GGGAGGAAAGTTGGAACAGG	GCCTCCAGGGAAATCTTCTT
#3	GAGAAAGGTTGGCTGGTTGT	CCCCTTCCTATGTCCTGGAT
AR intron #1	TGAACAAAGGGGTGAACTGA	CCATGCCATGCCAATTTTAT
#2	GTGCCCCCAAAACTAATGAA	ACTTTCAGGCAGCGTGTTTC

Table S2. Relationship of IHC immunoreactivity with clinicopathological parameters in prostate cancer tissues (n = 102)

(a) PSF

Value		PSF		*P* value
		High (n = 51)	Low (n = 51)	
Age* (years)		66.0±0.8	67.6±0.9	0.19
PSA* (ng / ml)		17.6±3.3	15.5±2.1	0.60
Stage (Jewett Staging Sysyem)				
	B	9	23	
	C	34	24	
	D1	8	4	**0.010**
pT				
	pT1-2	14	18	
	pT3-4	37	33	0.39
pN				
	pN0	43	47	
	pN1	8	4	0.22
Gleason score				
	2-6	17	22	
	7	12	16	
	8-10	22	13	0.17

(b)NONO

Value		NONO		*P* value
		High (n = 51)	Low (n = 51)	
Age* (years)		67.0±0.8	66.6±0.9	0.80
PSA* (ng / ml)		20.3±3.2	12.7±2.0	0.49
Stage (Jewett Staging Sysyem)				
	B	12	20	
	C	32	26	
	D1	7	5	0.23
pT				
	pT1-2	10	22	
	pT3-4	41	29	**0.010**
pN				
	pN0	44	46	
	pN1	7	5	0.54
Gleason score				
	2-6	18	21	
	7	7	21	
	8-10	26	9	**0.004**

(C) AR-V7

Value	AR-V7		*P* value
	High (n = 16)	Low (n = 86)	
Age* (years)	65.8±1.5	67.0±0.7	0.44
PSA* (ng / ml)	15.8±4	16.7±2.1	0.86
Stage (Jewett Staging Sysyem)			
B	3	29	
C	10	48	
D1	3	9	0.40
pT			
pT1-2	3	29	
pT3-4	13	57	0.24
pN			
pN0	13	77	
pN1	3	9	0.34
Gleason score (GS3)			
2-6	5	34	
7	5	23	
8-10	6	29	0.82
AR LI (%)	52.1±7.0	37.0±2.9	**0.043**

Table S3. ChIP qPCR primers

	Forward (5'-3')	Reverse (5'-3')
CDKN1A	CTGTGGCTCTGATTGGCTTT	GGGGTCTTTAGAGGTCTCCTG
CDKN1B	CGAAGAGTTAACCCGGGACT	AGTAGAACTCGGGCAAGCTG
CDKN2B	ACCTGGGCTCAGCTTCATTA	CTAGGCGCTTTTTCCCAGA
TP53	GGCAAAAAGAAACCGGAAAT	GCTGTCAGTCGTGGAAGTGA
ILK	GGTCTAGTTGCCTGCTCTCG	GTGTTTGTGTTGGGGGTAGG
DSEL	ACAAGTGCTCCACACATCCA	ATTTGGTTTGGGCACTCAAG

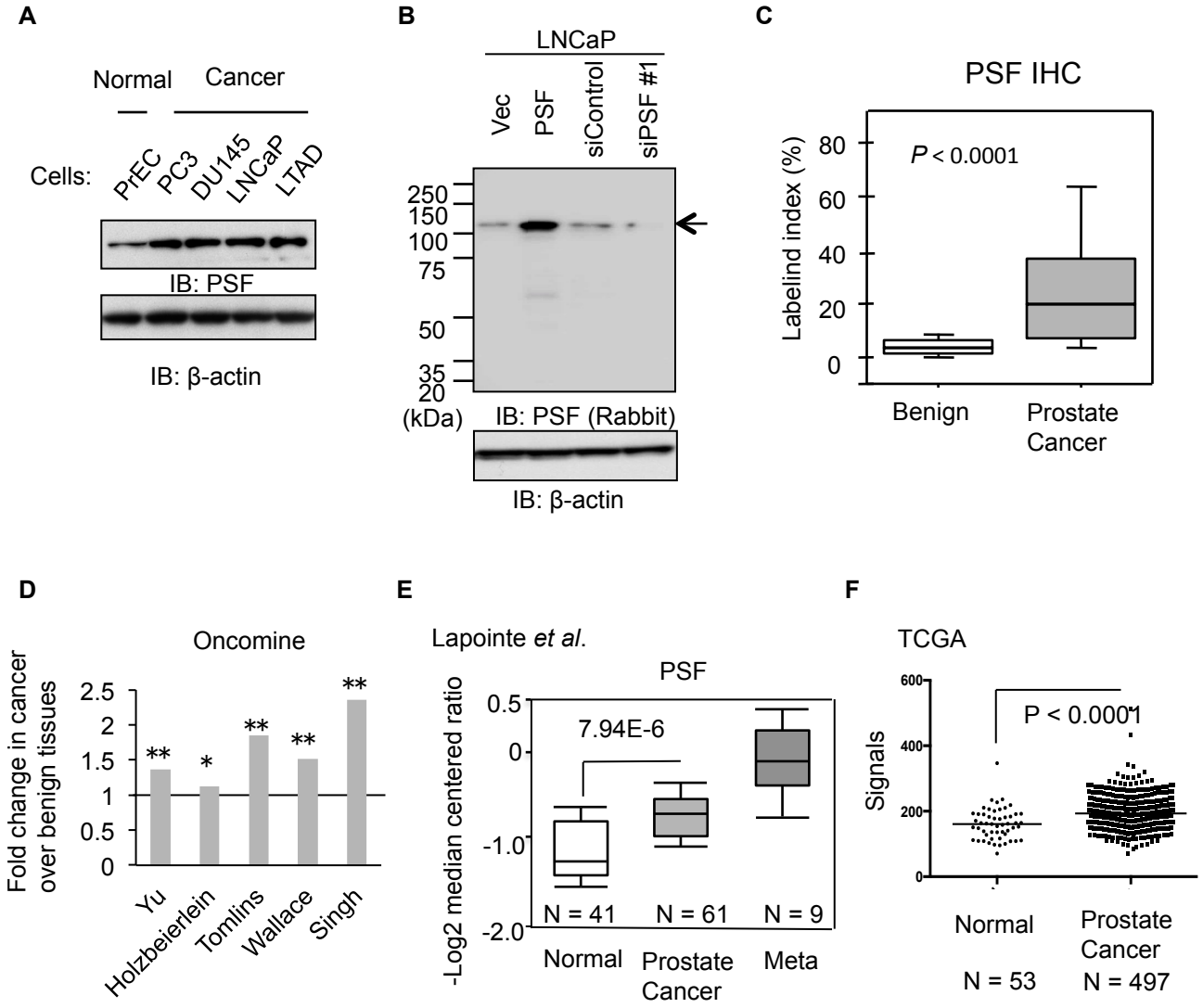

Fig. S1. PSF is upregulated in prostate cancer tissues compared with normal prostate.
(A) Expression level of PSF in normal prostate epithelial cells (PrEC) or prostate cancer cells. Immunoblot of cell lysates from PrEC, PC3, DU145, LNCaP and LTAD cells to detect PSF protein levels. (B) Evaluation of PSF rabbit polyclonal antibody used for immunohistochemistry (IHC). LNCaP cells were transfected with PSF-pcDNA3, empty vector, siControl and siPSF #1. Expression level of PSF protein was examined by western blot analysis. (C) Quantification of PSF protein expression level in prostate cancer tissues. Labeling index (%) of IHC was calculated in cancerous regions and benign regions in prostate cancer tissues (N=102). *P*-value was determined by paired t-test. (D) PSF is upregulated in cancer compared with benign tissues. By databases registered in Oncomine, we analyzed expression level of PSF and then fold changes in cancer compared with benign tissues are shown in five data sets. $*P < 0.05$, $**P < 0.01$. (E) One example of Oncomine data sets (Lapointe *et al.*) is shown. Meta: metastatic regions. (F) PSF is upregulated in cancer compared with benign tissues (Normal) using the data of the Cancer Genome Atlas (TCGA). *P*-value was obtained by Mann-Whitney U-test.

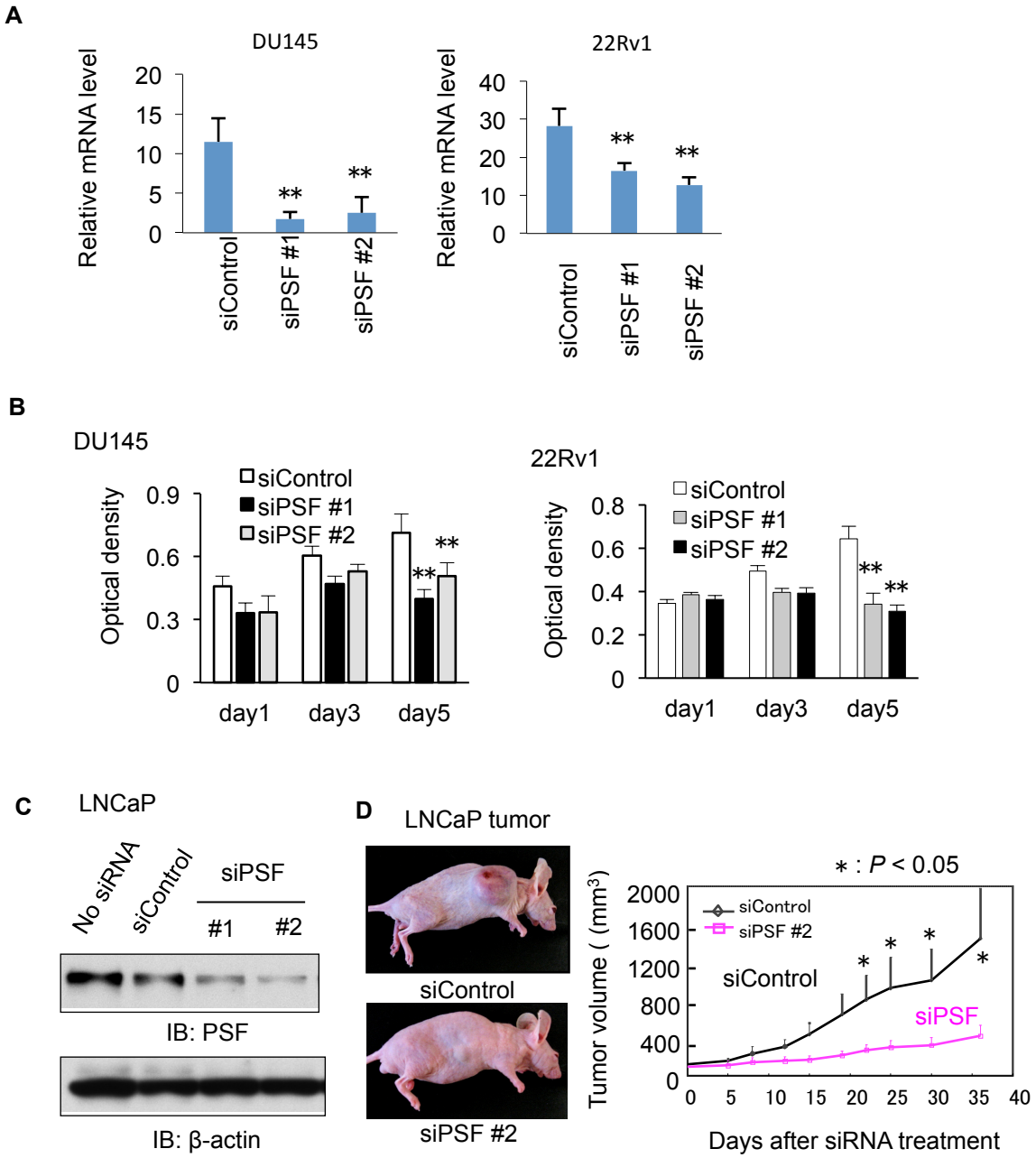

Fig. S2. Knockdown of PSF represses tumor growth derived from prostate cancer cells.
(A) Knockdown efficiency of siPSF in DU145 and 22Rv1 cells. We measured PSF expression levels by qRT−PCR (n = 3). (B) Growth of DU145, 22Rv1 prostate cancer cells after transfection of siControl or siPSF (n = 4). (C) Knockdown efficiency of siPSF #1, #2 in LNCaP cells. (D) PSF siRNA significantly reduced the volume of tumors when compared with the volume of those treated with siControl (N = 8). Representative photographs of mice harboring tumors are shown. Values represent the mean ± S.D. *P < 0.05, **P < 0.01.

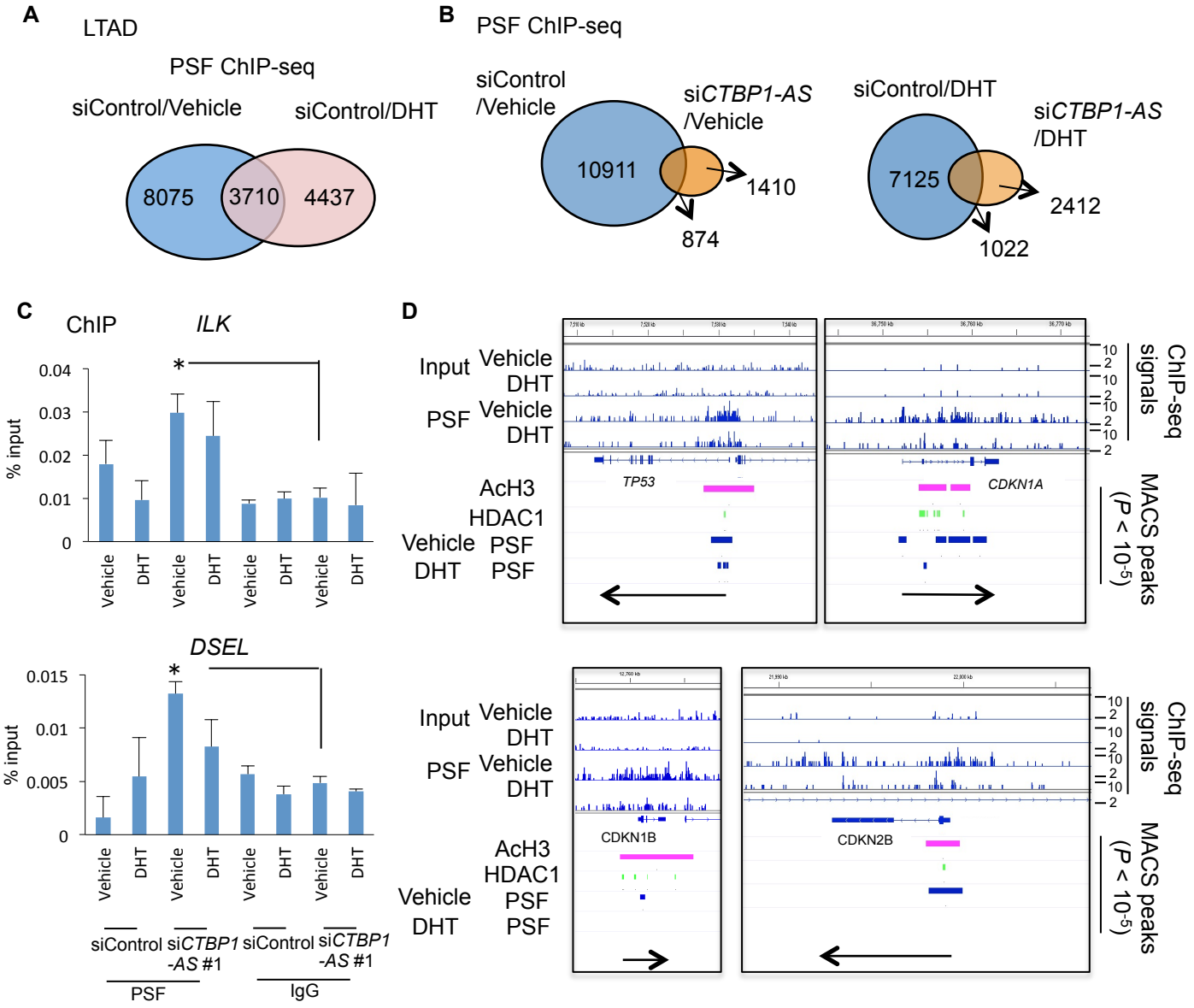
A
LTAD
PSF ChIP-seq
siControl/Vehicle
siControl/DHT
8075
3710
4437
B
PSF ChIP-seq
siControl /Vehicle
siCTBP1-AS /Vehicle
10911
1410
874
siControl/DHT
siCTBP1-AS /DHT
7125
2412
1022
C
ChIP
ILK
% input
0.04
0.03
0.02
0.01
0
Vehicle
DHT
DSEL
0.015
0.01
0.005
siControl
siCTBP1 -AS #1
PSF
IgG
D
Input
PSF
AcH3
HDAC1
ChIP-seq signals
MACS peaks (P < 10-5)
TP53
CDKN1A
CDKN1B
CDKN2B
E
PSF ChIP
siControl
siCTBP1-AS #1
% input
0.012
0.008
0.004
N.S
N.C
CDKN1A
CDKN1B
CDKN2B
TP53
PSF Vehicle
PSF DHT
IgG Vehicle
IgG DHT

Fig. S3. The ChIP-seq analysis of PSF genomic action reveals the associated genomic pathway in the progression into castration-resistant prostate cancer.

(A) Identification of PSF-binding sites at DNA level by chromatin immunoprecipitation-sequencing (ChIP-seq). PSF ChIP-seq analysis was performed using CRPC model cells derived from LNCaP cells, LTAD. (B) Cells were treated with siControl or si*CTBP1-AS* for 48 h and then treated with vehicle or DHT for 24 h. PSF-binding sites ($P < 10^{-4}$) at DNA level were determined by model-based analysis for ChIP-seq (MACS) in the presence or absence of DHT treatment. (C) Validation of PSF recruitments to the binding sites which are obtained by *CTBP1-AS* knockdown. We selected two binding sites, which are obtained by ChIP-seq only in the cells treated with si*CTBP1-AS*/vehicle, for ChIP assay. We treated LTAD cells with siControl or si*CTBP1-AS* #1 for 48 h. ChIP assay was performed using PSF antibody. Enrichments of PSF binding regions relative to input samples were measured by qPCR. (D) Tumor suppressive cell cycle regulators are representative targets of PSF at DNA level in LTAD cells. AcH3, HDAC1- and PSF-binding sites at DNA level ($P < 10^{-4}$) are shown. (E) Validation of PSF recruitments to the promoter regions of cell cycle inhibitors. We treated LTAD cells with siControl or si*CTBP1-AS* #1 for 48 h. ChIP assay was performed using PSF antibody. Enrichments of PSF-binding regions relative to input samples were measured by qPCR. Values represent the mean ± S.D. *P < 0.05, **P < 0.01. N.S: not significant.

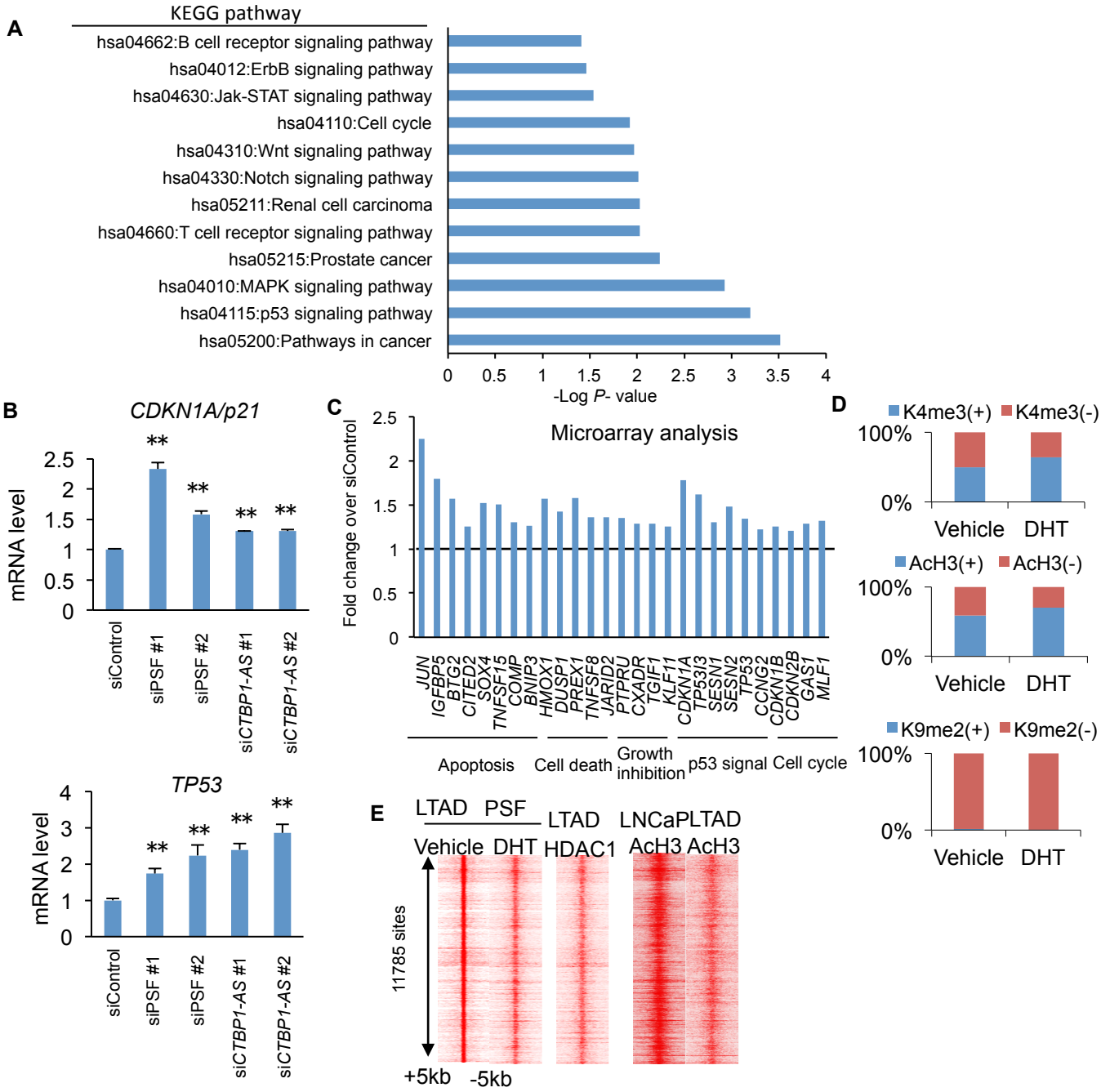

Fig. S4. PSF is associated with HDAC and histone modified regions and represses tumor suppressors transcriptionally.

(A) Pathway analysis of PSF-binding genes at DNA level repressed by PSF. PSF-binding genes at promoter regions in the absence of DHT were found by ChIP-seq. We then obtained 406 genes upregulated by siPSF (> 1.4 fold) using microarray analysis. Pathway analysis was performed in DAVID. (B) Validation of negative regulation of *TP53* and *CDKN1A* in LTAD cells. Cells were treated with siPSF (#1 and #2) or si*CTBP1-AS* (#1 and #2) for 48 h. qRT-PCR analysis was performed to examine the expression levels. **P < 0.01. (C) Tumor suppressive genes with PSF bindings at promoters were negatively regulated. We identified tumor suppressive genes (apoptosis, cell death, cell cycle, negative regulation of cell proliferation and p53 signal) by analyzing GO-terms of PSF target genes. (D) Overlap of PSF-binding sites with histone modified regions. The rate of PSF binding sites overlapped with histone modified regions. We observed marked overlaps of PSF-binding sites with AcH3 and K4me3 regions. (E) Genome-wide distribution of HDAC1 binding peaks and deacetylated regions around PSF binding sites in LTAD cells.

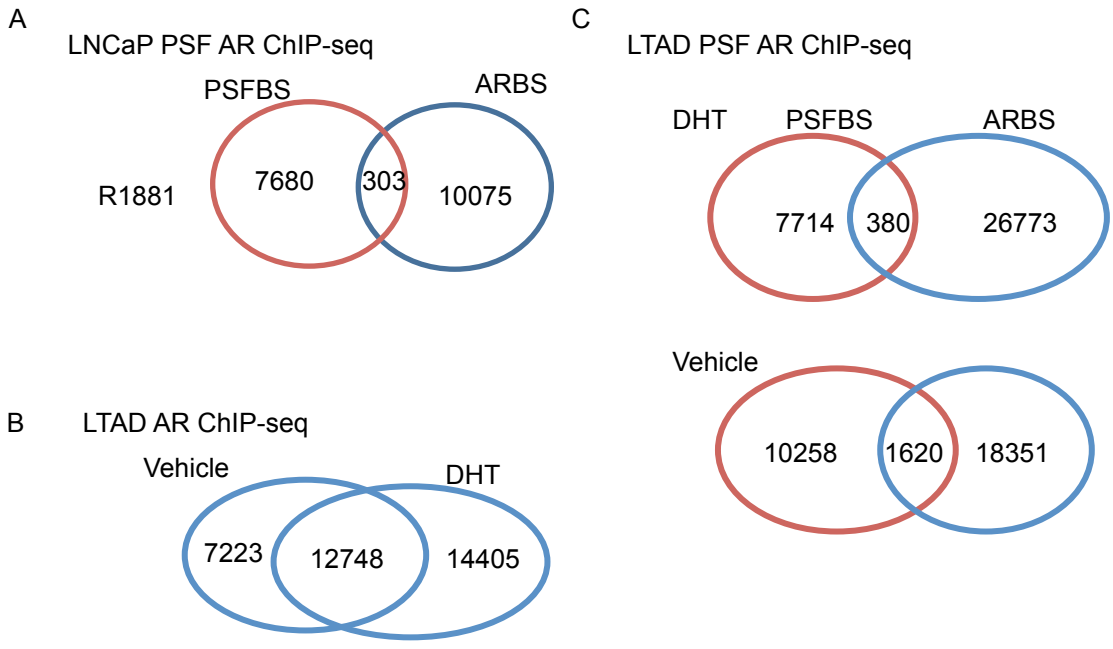

Fig. S5. PSF-binding sites (PSFBSs) overlap with a subset of AR-binding sites (ARBSs).
(A) Analysis of PSFBSs and ARBSs in LNCaP cells. Both binding sites were obtained in the presence of androgen (R1881 10 nM for 24 h) by ChIP-seq. (B) Identification of ARBSs in LTAD cells. We performed ChIP-seq to determine ARBSs in LTAD cells. Cells were treated with vehicle or DHT 10 nM for 24 h. (C) Overlaps of PSFBSs with ARBSs were analyzed in LTAD cells.

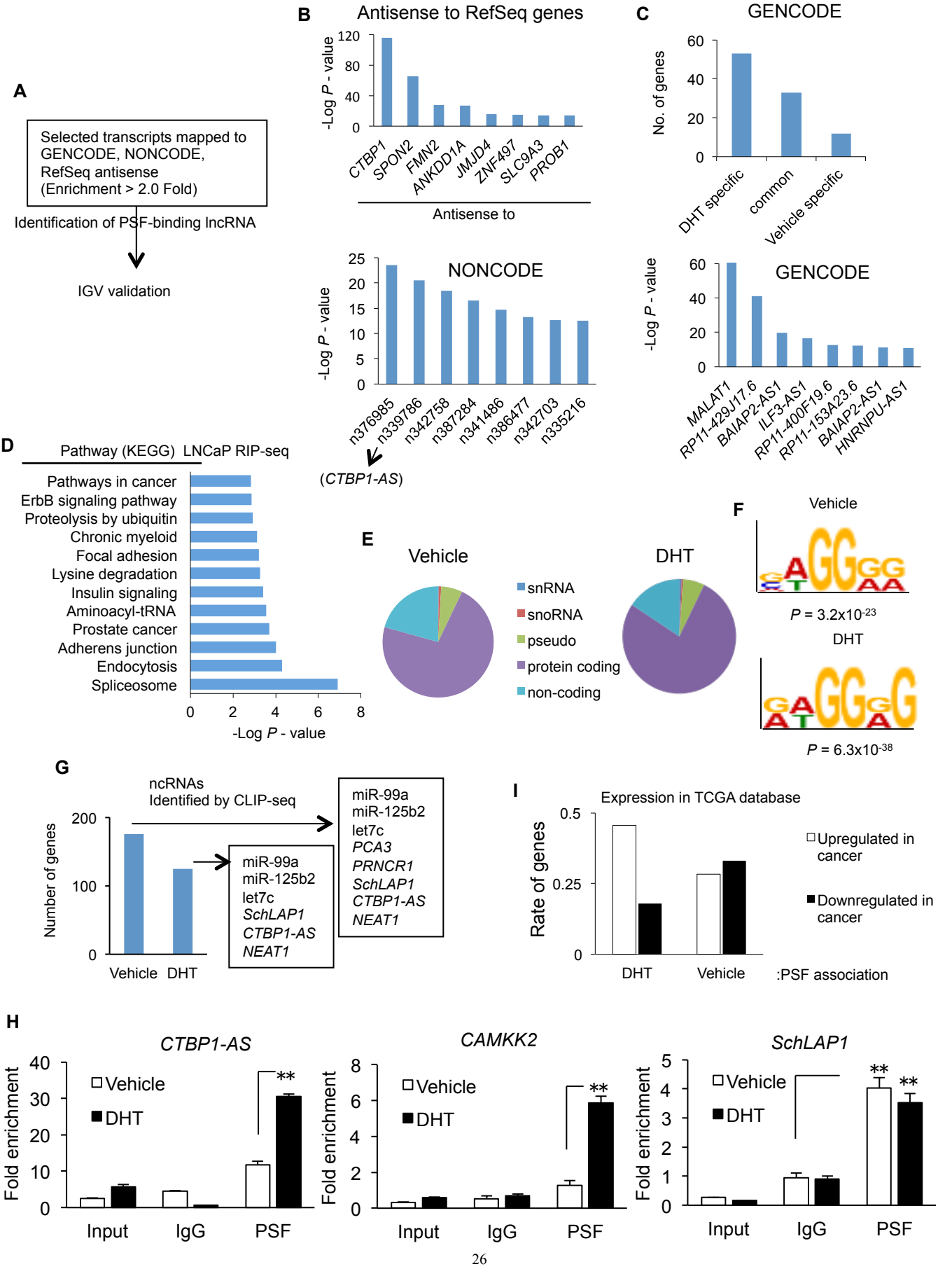
A
Selected transcripts mapped to GENCODE, NONCODE, RefSeq antisense (Enrichment > 2.0 Fold)
Identification of PSF-binding lncRNA
IGV validation
B
Antisense to RefSeq genes
-Log P - value
CTBP1
SPON2
FMN2
ANKDD1A
JMJD4
ZNF497
SLC9A3
PROB1
Antisense to
NONCODE
n376985
n339786
n342758
n387284
n341486
n386477
n342703
n335216
(CTBP1-AS)
C
GENCODE
No. of genes
DHT specific
common
Vehicle specific
GENCODE
MALAT1
RP11-429J17.6
BAIAP2-AS1
ILF3-AS1
RP11-400F19.6
RP11-153A23.6
BAIAP2-AS1
HNRNPU-AS1
D
Pathway (KEGG) LNCaP RIP-seq
Pathways in cancer
ErbB signaling pathway
Proteolysis by ubiquitin
Chronic myeloid
Focal adhesion
Lysine degradation
Insulin signaling
Aminoacyl-tRNA
Prostate cancer
Adherens junction
Endocytosis
Spliceosome
-Log P - value
E
Vehicle
DHT
snRNA
snoRNA
pseudo
protein coding
non-coding
F
Vehicle
P = 3.2x10-23
DHT
P = 6.3x10-38
G
ncRNAs Identified by CLIP-seq
Number of genes
Vehicle
DHT
miR-99a
miR-125b2
let7c
SchLAP1
CTBP1-AS
NEAT1
miR-99a
miR-125b2
let7c
PCA3
PRNCR1
SchLAP1
CTBP1-AS
NEAT1
I
Expression in TCGA database
Rate of genes
Upregulated in cancer
Downregulated in cancer
DHT
Vehicle
:PSF association
H
CTBP1-AS
CAMKK2
SchLAP1
Fold enrichment
Vehicle
DHT
Input
IgG
PSF

Fig. S6. Identification of PSF-binding lncRNAs in prostate cancer cells.
(A) Schematic summary to analyze PSF-binding lncRNAs by RIP-seq is shown. (B) Top ten lncRNAs (Antisense to RefSeq or NONCODE) bound with PSF in prostate cancer cells. These transcripts were validated to be lncRNAs by manually referring to the IGV. n339786 is *CTBP1-AS*. (C) Number of genes identified in the annotated regions of GENCODE. The validated genes by IGV as lncRNAs are listed. (D) Pathway analysis of PSF-binding genes in LNCaP cells is shown. (E) Classification of PSF target genes identified by CLIP-seq. (F) Motif analysis of identified peaks. Flanking regions (- 10 to + 10 nucleotides) around identified reliable mutation sites within the cross-linking regions were used as input for MEME. (G) Identification of lncRNAs and miRNAs by CLIP-seq in LNCaP cells. (H) Validation of PSF bindings with transcripts (*CTBP1-AS*, *CAMKK2* and *SchLAP1*) identified by CLIP/RIP-seq. (I) PSF-binding transcripts in the presence of DHT are upregulated in prostate cancer tissues. TCGA data base was used to analyze the expression of genes in prostate cancer (n = 497) or normal prostate (n = 52). We performed Mann-Whitney U-test to analyze the significant change in tumors. Values represent the mean ± S.D. **P < 0.01.

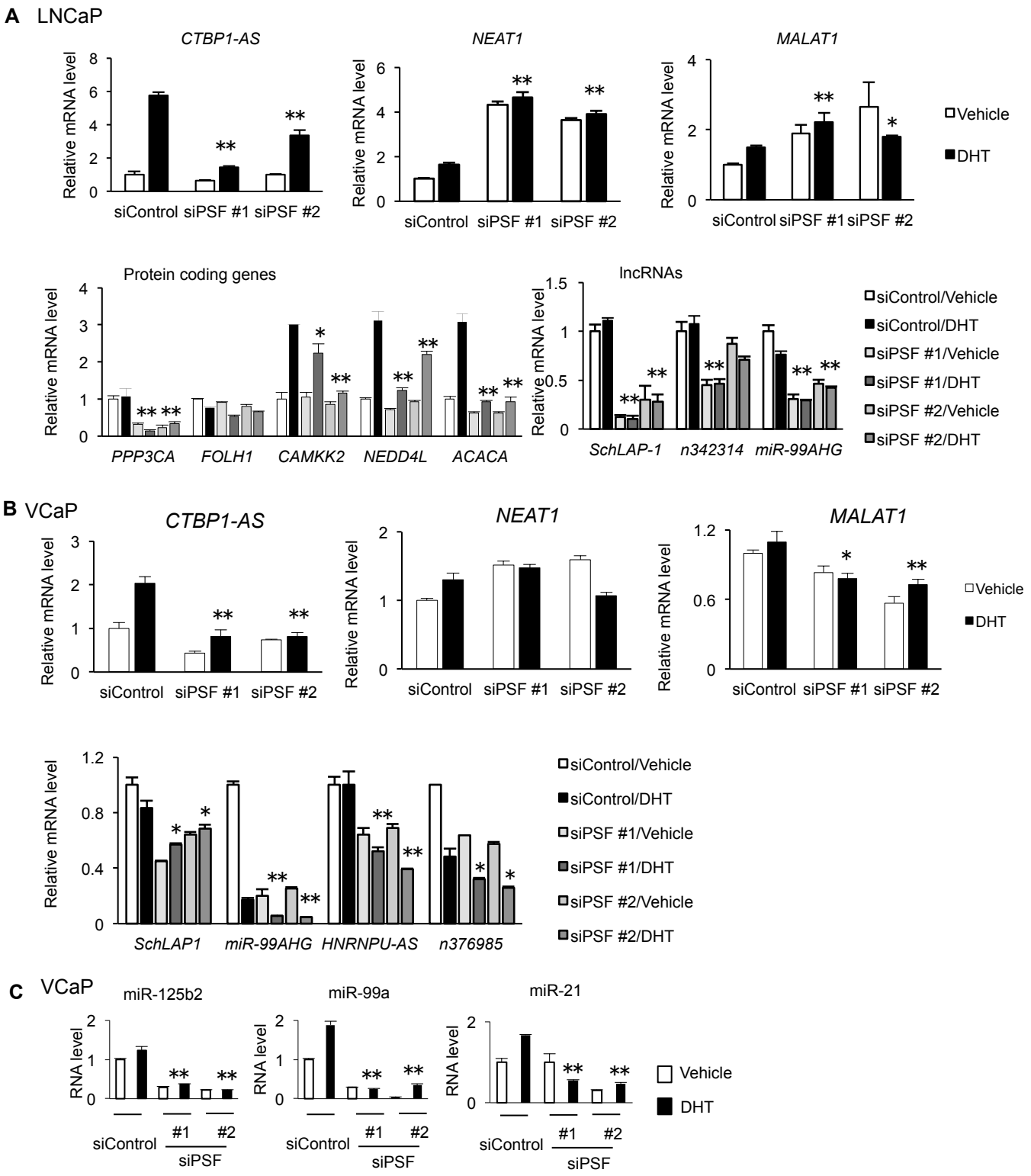

Fig. S7. PSF is associated with lncRNAs and miRNA maturation by binding to the primary transcripts. (A, B) Expressions of PSF-binding genes are repressed by PSF knockdown. LNCaP (A) or VCaP (B) cells were treated with siControl, si*PSF* #1, or si*PSF* #2 for 48 h. Cells were treated with vehicle or 10 nM DHT for 24 h. mRNA expression levels were measured by qRT-PCR (n = 3). (C) Expression of androgen-induced miRNAs were repressed by knockdown of PSF. VCaP cells were treated with siControl or siPSF for 48 h. We then measured miRNA expression levels by qRT-PCR in cells after treatment of vehicle or DHT for 24 h (n =3). Values represent the mean ± S.D. *P < 0.05, **P < 0.01.

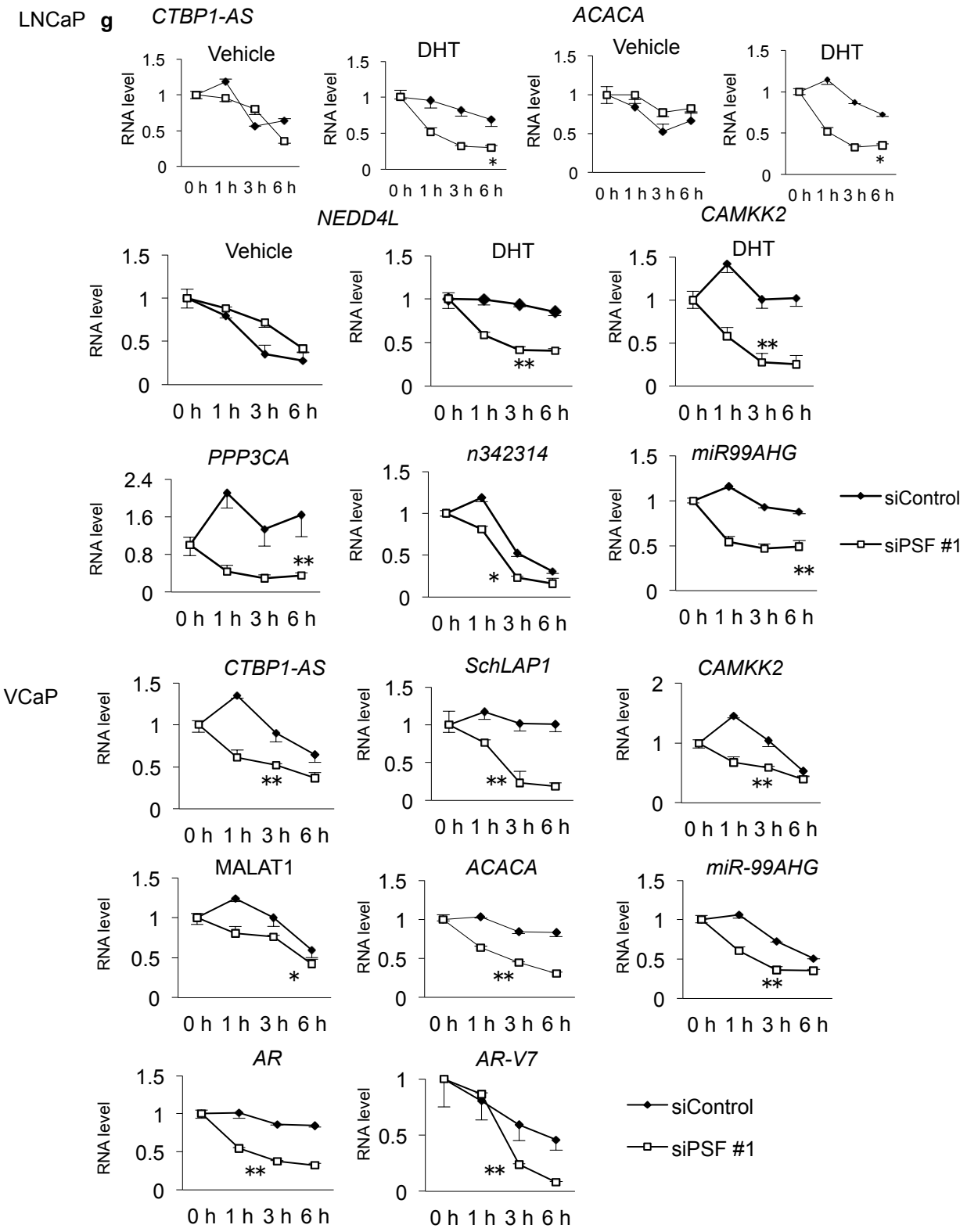

Fig. S8. PSF is responsible for the stability of its target mRNAs.
LNCaP or VCaP cells were transfected with siControl or siPSF#2. After 48 h incubation, cells were treated with vehicle or DHT for 18 h to analyze androgen-regulated genes. After transcriptional inhibition with 1 μg/ml actinomycin-D, expression levels of PSF-binding transcripts relative to GAPDH were measured by qRT-PCR (n = 3). Values represent the mean ± S.D. *P < 0.05, **P < 0.01. Lnc RNAs or androgen-regulated genes such as ACACA (8), CAMKK2 and NEDD4L were examined.

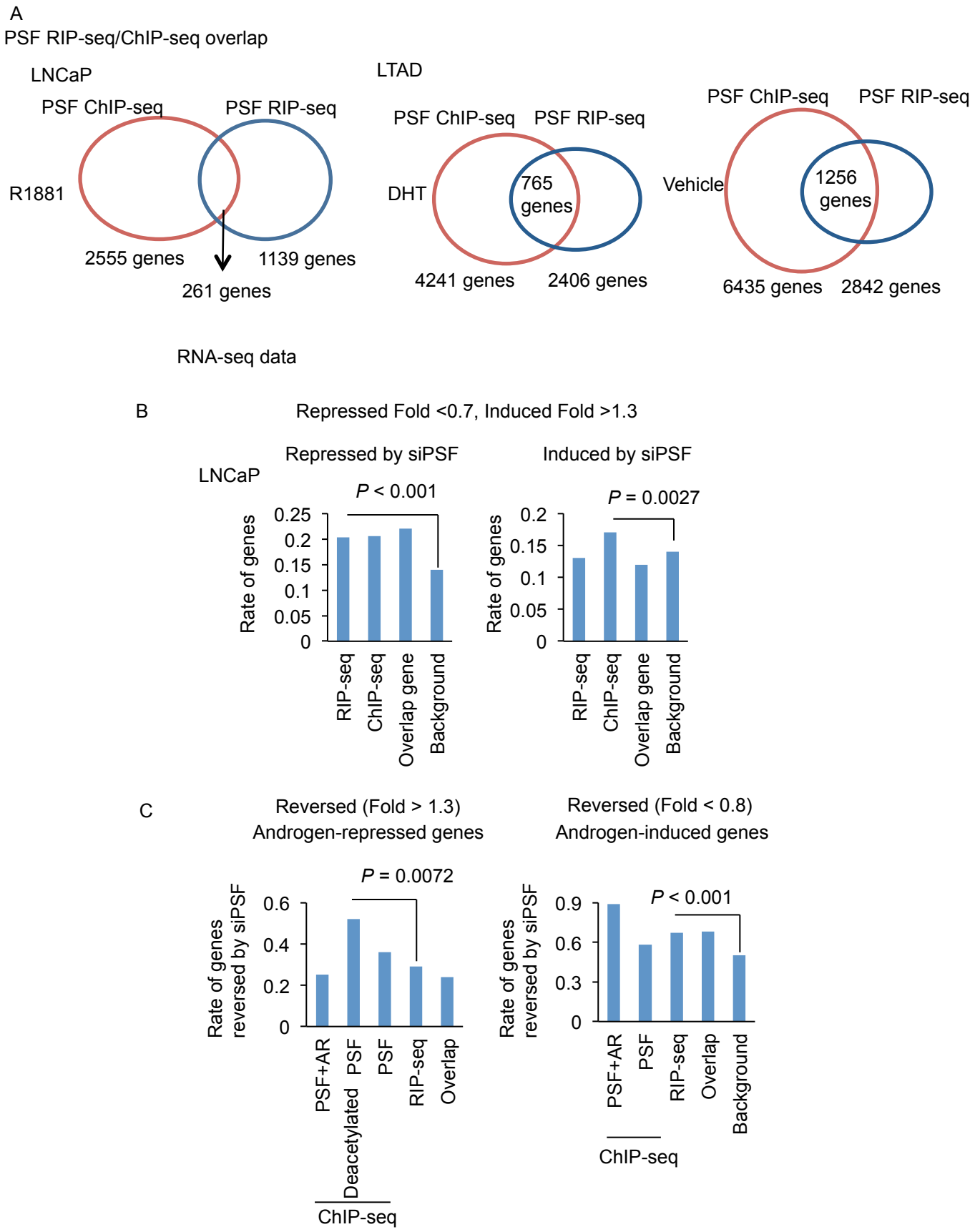
A
PSF RIP-seq/ChIP-seq overlap
LNCaP
PSF ChIP-seq
PSF RIP-seq
R1881
2555 genes
1139 genes
261 genes
LTAD
PSF ChIP-seq
PSF RIP-seq
DHT
765 genes
4241 genes
2406 genes
PSF ChIP-seq
PSF RIP-seq
Vehicle
1256 genes
6435 genes
2842 genes
RNA-seq data
B
Repressed Fold <0.7, Induced Fold >1.3
Repressed by siPSF
Induced by siPSF
LNCaP
P < 0.001
P = 0.0027
Rate of genes
0.25
0.2
0.15
0.1
0.05
0
RIP-seq
ChIP-seq
Overlap gene
Background
C
Reversed (Fold > 1.3)
Androgen-repressed genes
Reversed (Fold < 0.8)
Androgen-induced genes
P = 0.0072
P < 0.001
Rate of genes reversed by siPSF
0.6
0.4
0.2
0
0.9
0.6
0.3
PSF+AR
Deacetylated PSF
PSF
RIP-seq
Overlap
Background
ChIP-seq

Fig. S9. Analysis of functional roles of PSF as an RNA binding protein and as a transcriptional regulator. (A) Overlaps of genes sets identified by PSF RIP-seq and ChIP-seq. We found gene sets with PSF bindings at genomic DNA regions (PSF ChIP-seq) and PSF-binding genes at RNA level (PSF RIP-seq). Ven diagrams show the overlaps of these two gene sets. (B) Analysis of the impacts of PSF bindings at DNA and RNA levels in LNCaP cells. LNCaP cells treated with siControl or siPSF #1. After 48 h incubation, cells were treated with vehicle or DHT for 24 h and then RNA-seq analysis was performed. The rate of genes regulated by PSF in the presence of DHT was calculated. *P*-values were obtained by chi-square test. (C) Effects of PSF bindings at DNA and RNA levels analyzed by RNA-seq analysis. We analyzed the roles of PSF in androgen dependent repression and induction of PSF-binding genes identified by RIP-seq and ChIP-seq in LNCaP cells. Androgen-dependent deacetylated regions, in which significant histone acetylation (AcH3) was detected only in the absence of androgen at the threshold of fold enrichment > 7.5, were found by ChIP-seq. *P*-values were obtained by chi-square test.

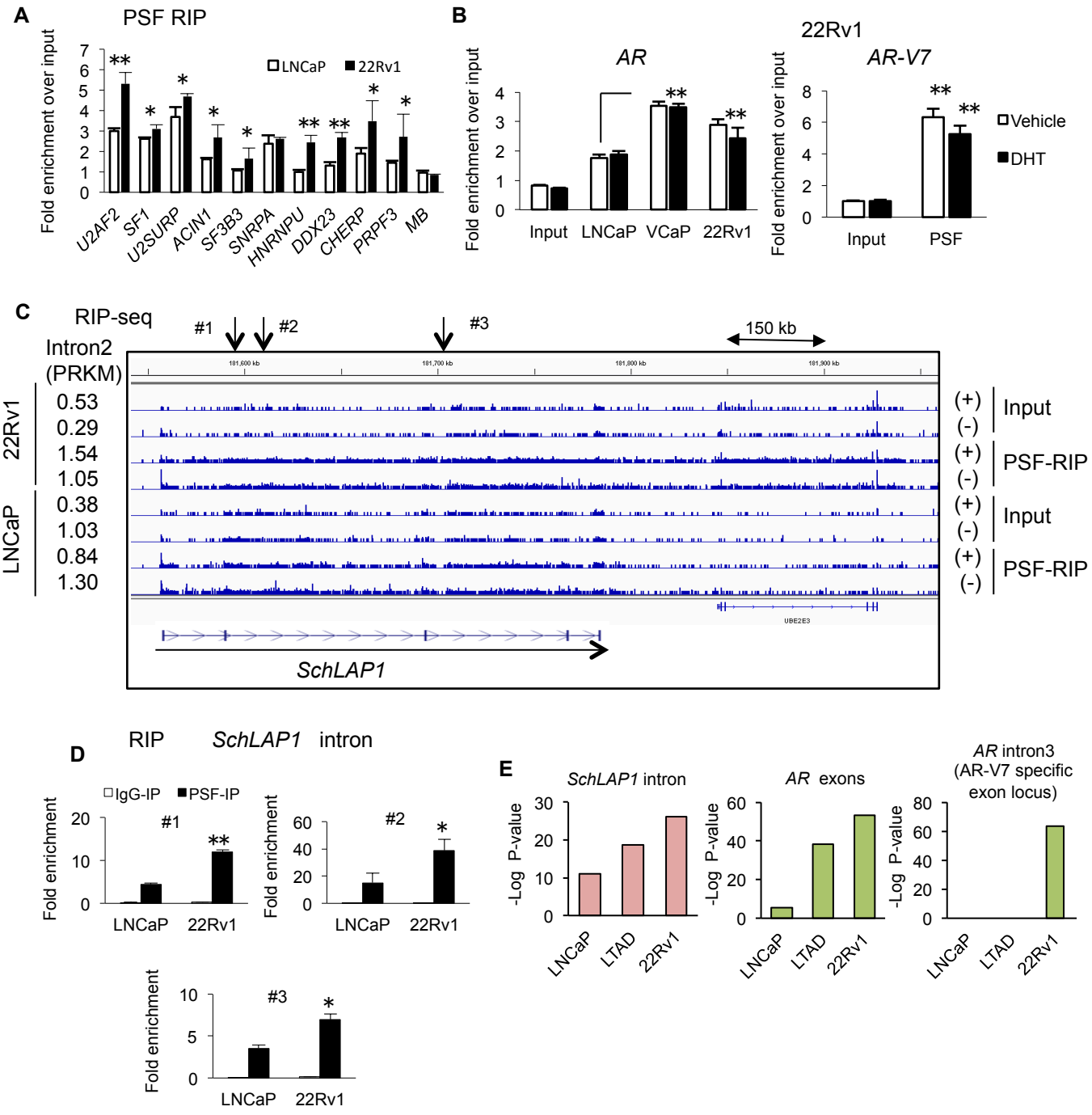
A
PSF RIP
Fold enrichment over input
LNCaP
22Rv1
U2AF2
SF1
U2SURP
ACIN1
SF3B3
SNRPA
HNRNPU
DDX23
CHERP
PRPF3
MB
B
AR
22Rv1
AR-V7
Fold enrichment over input
Input
LNCaP
VCaP
22Rv1
Input
PSF
Vehicle
DHT
C
RIP-seq
Intron2
(PRKM)
22Rv1
0.53
0.29
1.54
1.05
LNCaP
0.38
1.03
0.84
1.30
#1
#2
#3
150 kb
Input
PSF-RIP
Input
PSF-RIP
SchLAP1
D
RIP
SchLAP1
intron
IgG-IP
PSF-IP
Fold enrichment
#1
#2
#3
LNCaP
22Rv1
E
SchLAP1 intron
AR exons
AR intron3
(AR-V7 specific
exon locus)
-Log P-value
LNCaP
LTAD
22Rv1
F
LNCaP
#1
#2
PSF
CLIP
Vehicle
DHT
RIP
Input
PSF-IP
G
RIP
AR intron
IgG
PSF-IP
Exon
Fold enrichment
#1
#2
LNCaP
22Rv1

Fig. S10. Validation of PSF-bindings to the target genes identified by CLIP and RIP-seq.
(A) PSF-binding with transcripts of spliceosome genes is enhanced in CRPC model cells. RIP assay was performed in LNCaP and 22Rv1 cells (n = 3). (B) Binding of PSF with *AR* transcript is enhanced in AR-V7 positive cells. RIP assay was performed in LNCaP, LTAD, VCaP and 22Rv1 cells. Binding of PSF with *AR* and *AR-V7* relative to GAPDH was measured by qRT-PCR analysis (n = 3). (C) RIP-seq analysis of *SchLAP1* in prostate cancer cells. Sequence tags around *SchLAP1* are enriched by PSF-IP. Arrows indicate the location of primers used for RIP. (D) Binding of PSF with intron regions of *SchLAP1*. PSF RIP assay was performed in LNCaP and 22Rv1 cells. Association of PSF with *SchLAP1* intron regions was analyzed by qRT-PCR relative to *GAPDH* (n = 3). (E) Enhanced bindings of PSF to *SchLAP1, AR* exons and *AR* introns in CRPC cells. The results of RIP-seq in the intron region of *SchLAP1*, *AR* exons, and *AR* introns were shown. (F) The location of PSF-binding peaks. Mapping of CLIP tags around *AR* gene locus. The loci of primers are indicated with arrow. (G) Binding of PSF with transcripts of intron regions of *AR*. PSF RIP assay was performed in LNCaP and 22Rv1 cells. Binding of PSF with *AR* intron regions was analyzed by qRT-PCR relative to *GAPDH* (n = 3). Values represent the mean ± S.D. *$P < 0.05$, **$P < 0.01$.

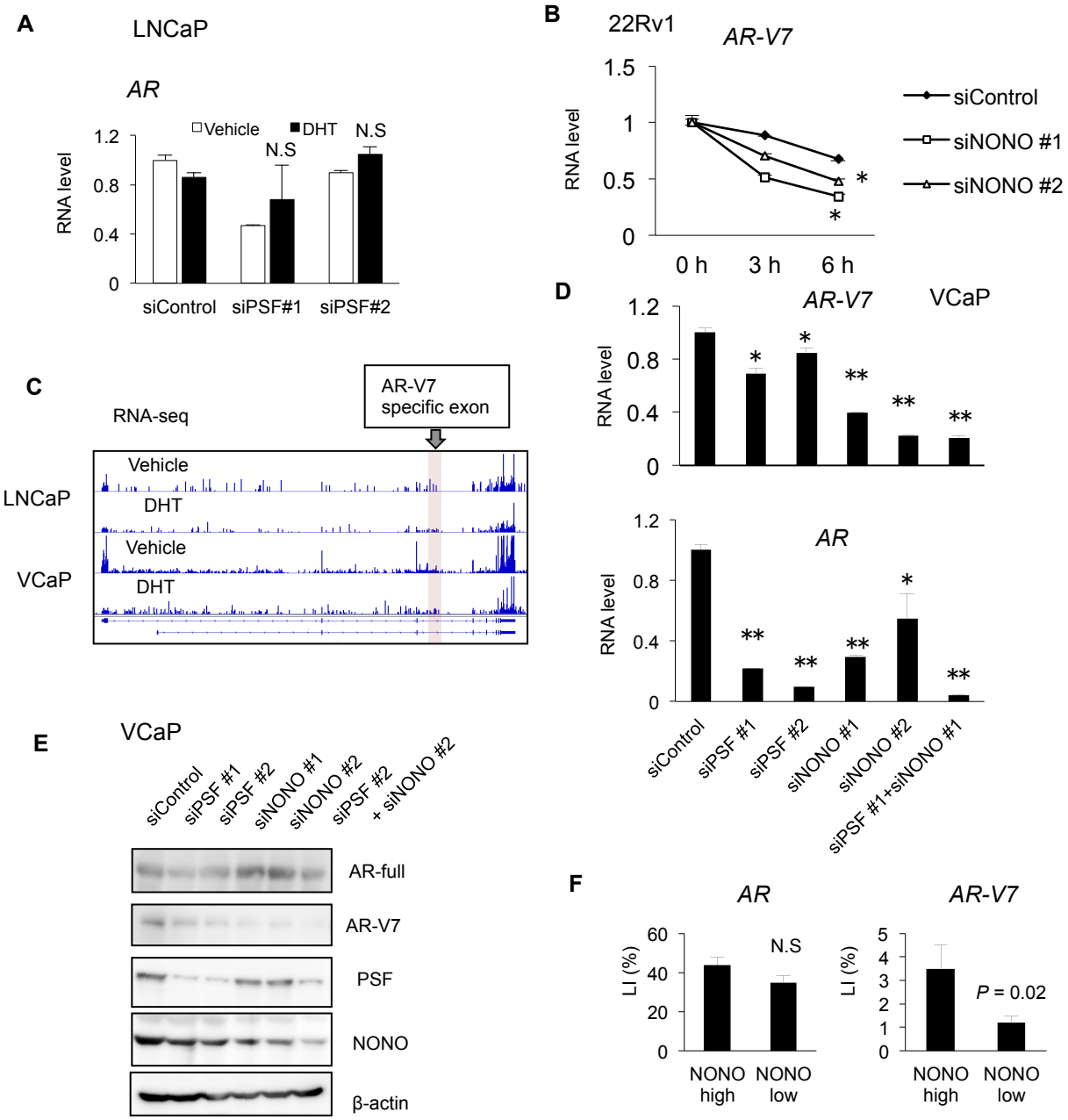

Fig. S11. AR and AR-V7 expression are regulated by PSF and NONO in prostate cancer cells.
(A) Regulation of *AR* mRNA expression by PSF in LNCaP cells. LNCaP cells were treated with siPSF for 72 h. *AR* mRNA level was measured by qRT-PCR (n = 3). Analysis at protein level has been shown in the previous report (11). (B) AR-V7 mRNA stability decreased by knockdown of NONO. 22Rv1 cells are treated with siControl, siNONO #1, or siNONO #2 for 48 h. Cells are treated with Actinomycin-D (1 uM) to inhibit transcription. mRNA expression levels were measured by qRT-PCR (n = 3). (C) RNA-seq analysis of *AR* transcripts in VCaP and LNCaP cells. RNAs of introns and *AR-V7* specific exon were elevated in VCaP cells. (D) Regulation of *AR* and *AR-V7* mRNA expression by PSF and NONO in VCaP cells. Cells were treated with siPSF or siNONO for 72 h. *AR* and *AR-V7* mRNA levels were measured by qRT-PCR (n = 3). (E) AR and AR-V7 expression are regulated by RNA-binding proteins in prostate cancer cells. Immunoblot of cell lysates from VCaP cells transfected with siPSF or siNONO to detect AR and AR-V7 protein levels. (F) Positive correlation of PSF expression with AR and AR-V7 levels in prostate cancer tissues (N = 102). Immunoreactivity index (LI) is shown. Values represent the mean ± S.D. *P < 0.05, **P < 0.01, N.S, not significant.

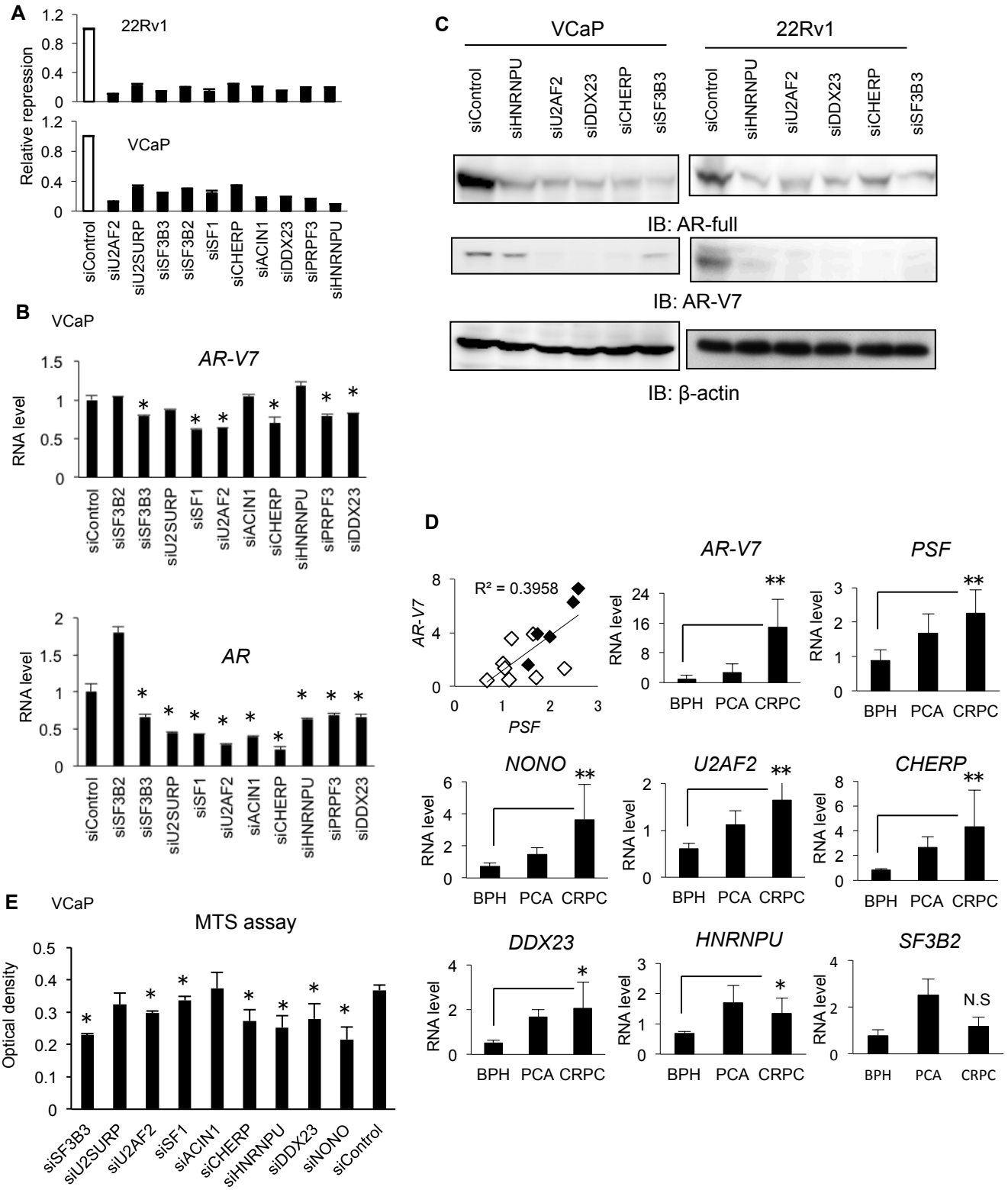
A
22Rv1
VCaP
Relative repression
siControl
siU2AF2
siU2SURP
siSF3B3
siSF3B2
siSF1
siCHERP
siACIN1
siDDX23
siPRPF3
siHNRNPU
B
VCaP
AR-V7
AR
RNA level
C
VCaP
22Rv1
IB: AR-full
IB: AR-V7
IB: β-actin
D
AR-V7
PSF
NONO
U2AF2
CHERP
DDX23
HNRNPU
SF3B2
R² = 0.3958
BPH PCA CRPC
N.S
E
VCaP
MTS assay
Optical density
siNONO

Fig. S12. Spliceosome genes targeted by PSF are responsible for AR splicing and cell proliferation in prostate cancer.

(A) Knockdown efficiency of siRNAs targeting spliceosome genes. Cells were transfected with each siRNA for 48 h. mRNA level of each gene was measured by qRT-PCR. Fold repression by siRNA relative to siControl treatment was shown. (B) Regulation of *AR* and *AR-V7* mRNA expression by PSF-targeted spliceosome genes in prostate cancer cells. VCaP cells were treated with siRNA targeting spliceosome genes for 72 h. *AR* and *AR-V7* mRNA levels were measured by qRT-PCR (n = 3). (C) Lysates from VCaP and 22Rv1 cells transfected with siControl or siRNA targeting spliceosome genes were analyzed by western blot analysis to detect AR and AR-V7. (D) Analysis of spliceosome genes in prostate cancer tissues. Total RNA was extracted from benign (N = 9) or prostate cancer tissue (localized n = 8, CRPC n = 5). BPH: benign prostatic hyperplasia, PCA: localized prostate cancer. We measured mRNA expression level of indicated genes by qRT-PCR (n = 3). (E) Spliceosome genes regulate prostate cancer cell growth. Growth of VCaP prostate cancer cells after transfection of siControl or siRNA targeting spliceosome genes (n = 4). Values represent the mean ± S.D. *P < 0.05, **P < 0.01, N.S, not significant.

References

1. Takayama K, Horie-Inoue K, Katayama S, Suzuki T, Tsutsumi S, et al (2013). Androgen-responsive long noncoding RNA CTBP1-AS promotes prostate cancer. EMBO J. 32:1665–1680.
2. Huang DW, Sherman BT, Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics rescorces. Nature Protoc. 4:44-57.
3. Takayama K, Suzuki T, Fujimura T, Urano T, Takahashi S, Homma Y, Inoue S (2014). CtBP2 modulates the androgen receptor to promote prostate cancer progression. Cancer Res. 2014;74(22):6542-6553.
4. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9: R137.
5. Takayama K, et al (2011). Integration of cap analysis of gene expression and chromatin immunoprecipitation analysis on array reveals genome-wide androgen receptor signaling in prostate cancer cells. Oncogene. 30(5):619-30.
6. Chen B, Yun J, Kim MS, Mendell JT, Xie Y (2014). PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. Genome Biol. 15(1):R18.
7. Fujimura T, Takahashi S, Urano T, Takayama K, Sugihara T, et al (2014). Expression of androgen and estrogen signaling components and stem cell markers to predict cancer progression and cancer-specific survival in patients with metastatic prostate cancer. Clin Cancer Res. 20:4625-4635.
8. Brusselmans K, De Schrijver E, Verhoeven G, Swinnen JV (2005). RNA interference-mediated silencing of the acetyl-CoA-carboxylase-alpha gene induces growth inhibition and apoptosis of prostate cancer cells. Cancer Res. 65(15):6719-25.