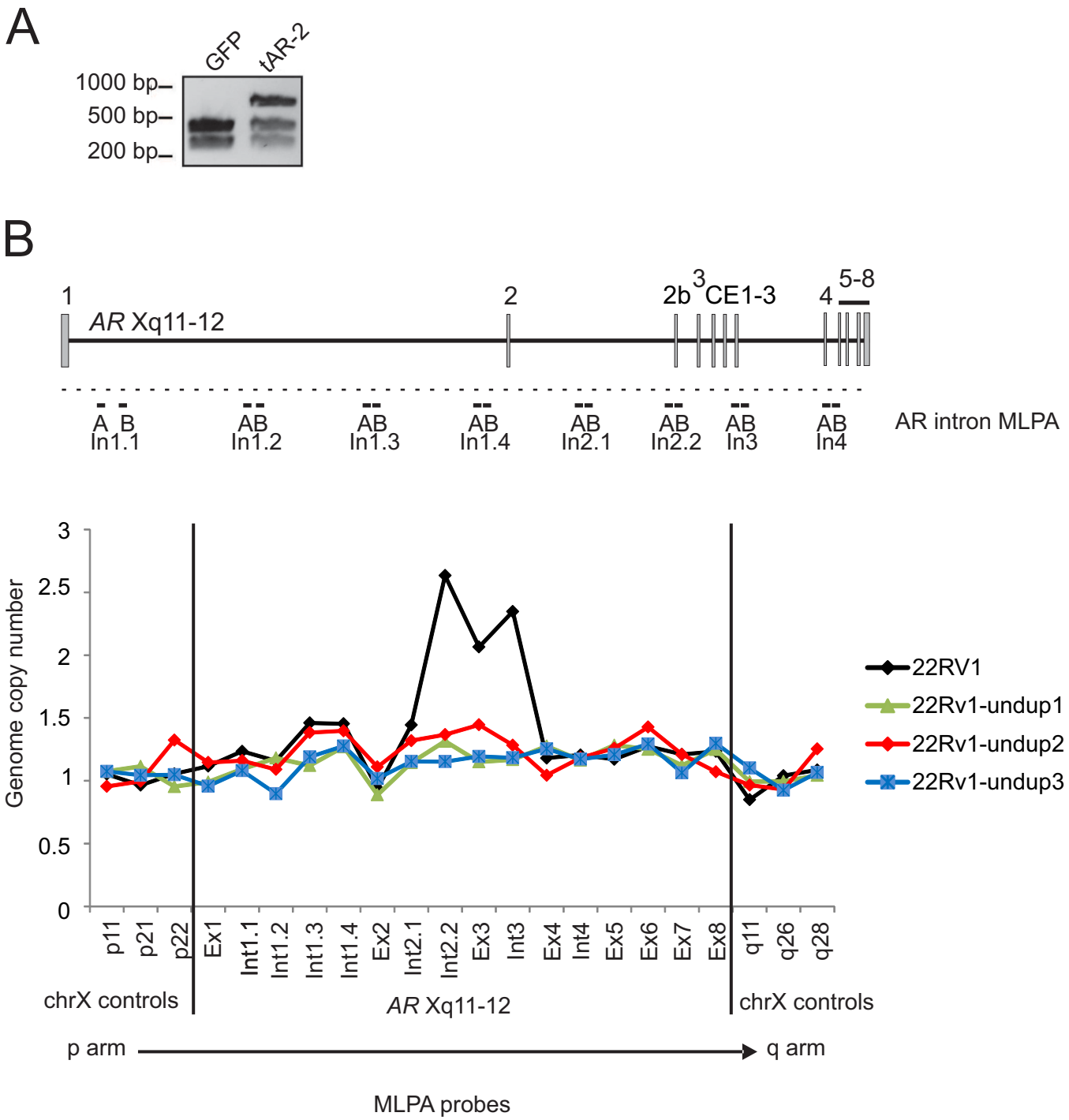


Figure S1: TALEN activity and genetic correction of 22Rv1 sub-lines. A. tAR-2 activity was confirmed with a PCR digest assay that measures nonhomologous end joining (NHEJ) mutations induced within the cleavage site in the spacer sequence between TALEN binding sites. An AflIII endonuclease site located in the spacer sequence; thus, error prone NHEJ from tAR-2 induced dsDNA breaks, induced mutations in the AflIII site, rendering the site resistant to AflIII digestion. B. MLPA of 22Rv1 and undup1-3 sub-lines. Probes were designed to target AR exons 1-8, and introns at the specified genomic locations (top panel).

3’
Undup3: TATATAACAATTTCAGAGAGTCCAC GGACTCTCTGAAATTGTTATAAGGT
Undup2: ATAACAATTTCAGAGAGTCCACATA TTGTTATAAGGTCTTTTTCTTTGTT
Undup1: TATATAACAATTTCAGAGAGTCCAC GGACTCTCTGAAATTGTTATAAGGT

SSA
TALEN
AR

5’
Undup3: ATGAACATTCCTGCCTGGCTGACAT CTTTACTCATATATACTTTAGATTC
Undup2: GGATGAACATTCCTGCCTGGCTGAC GGTGGCGGCAAGCAAGCGCTCGAAA
Undup1: ATGAACATTCCTGCCTGGCTGACAT CTTTACTCATATATACTTTAGATTC

Figure S2: 5' and 3' breakpoint sequence junctions of 22Rv1 undup1-3 sub-lines. Different colors represent sequence origin; SSA reporter sequence (blue), TALEN (purple), AR gene (red).

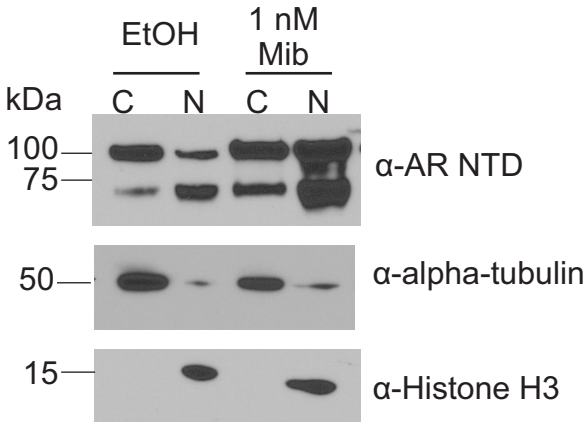


Figure S3: AR localizes to the nucleus upon stimulation with 1 nM mibolerone in 22Rv1-undup3 cells. AR variants are constitutively nuclear. Cells were cultured under androgen-free conditions for 24 hours and then treated for 24 hours with medium containing 1 nM mibolerone (Mib, synthetic androgen) or ethanol (EtOH, vehicle control). Cell lysates were fractionated into cytoplasmic (C) or nuclear (N) fractions and analyzed by Western blot with AR-NTD, alpha-tubulin (cytoplasmic marker), and Histone-H3 (nuclear marker)- specific antibodies.

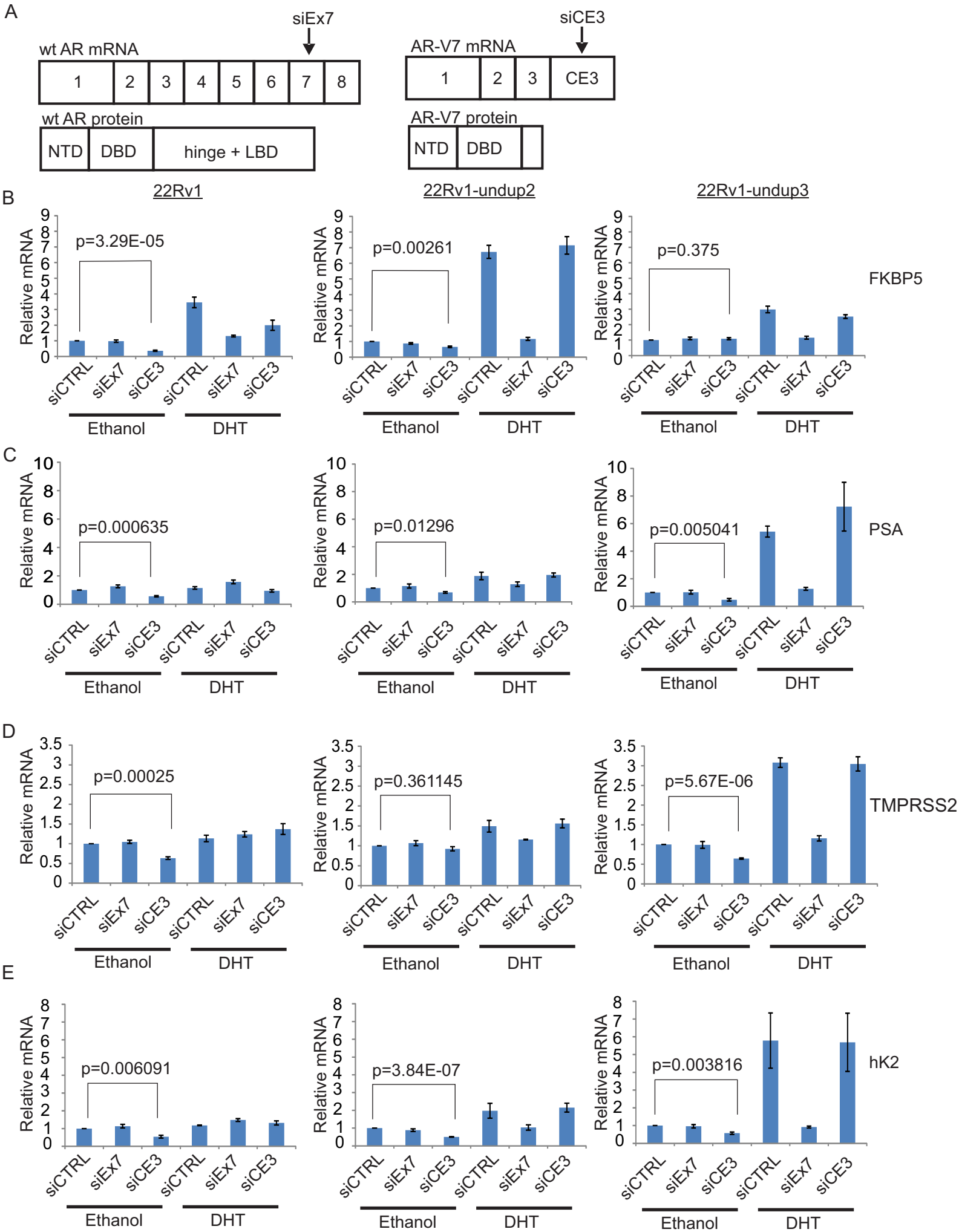

A
siEx7
siCE3
wt AR mRNA
1
2
3
4
5
6
7
8
AR-V7 mRNA
1
2
3
CE3
wt AR protein
NTD
DBD
hinge + LBD
AR-V7 protein
NTD
DBD
B
22Rv1
22Rv1-undup2
22Rv1-undup3
Relative mRNA
p=3.29E-05
p=0.00261
p=0.375
FKBP5
siCTRL
siEx7
siCE3
Ethanol
DHT
C
p=0.000635
p=0.01296
p=0.005041
PSA
D
p=0.00025
p=0.361145
p=5.67E-06
TMPRSS2
E
p=0.006091
p=3.84E-07
p=0.003816
hK2

Figure S4: AR target gene expression panel. FL-AR and AR-V7 were knocked down with siRNAs targeting exon 7 or CE3 (A). Relative mRNA levels of AR target genes, FKBP5 (B), PSA (C), TMPRSS2 (D), and hK2 (E), were assessed by RT-qPCR. Error bars represent the standard error of the mean (n=8). Significance was determined using a two-tailed t- test. GAPDH was used as an internal control.

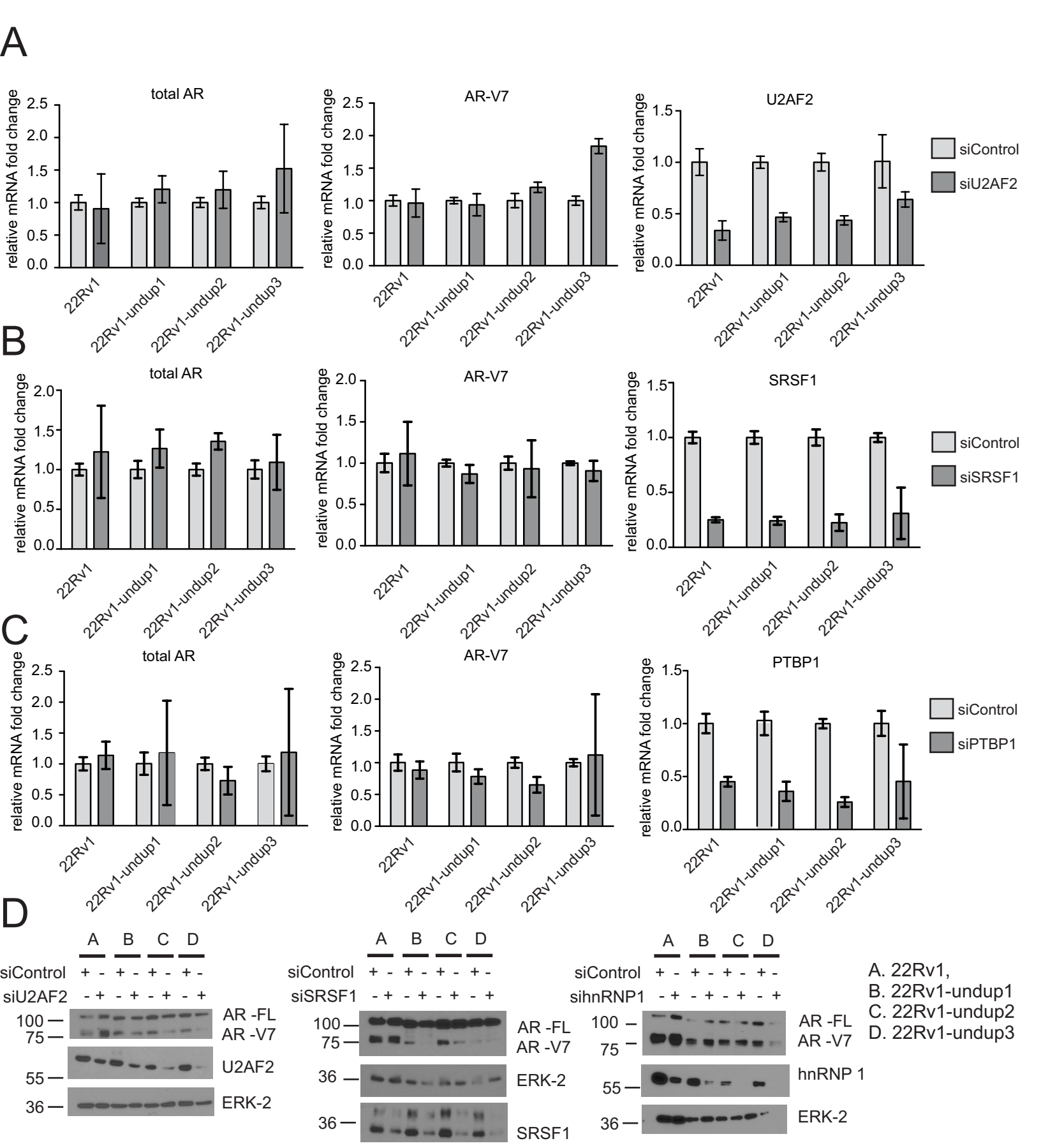
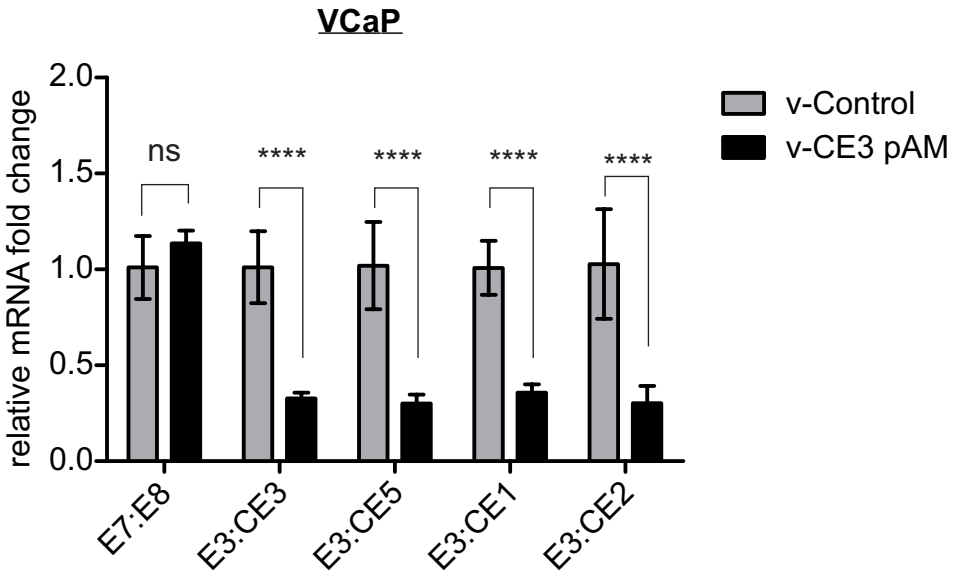


Figure S5: Knockdown of splicing factors U2AF2, SRSF1, and PTBP1. The effect of splicing factor knockdown on levels of total AR and AR-V7 relative to nontargeting controls was assessed by RT-qPCR (mRNA, n=4, A-C) and Western blot (protein, D). Knockdown of U2AF2, SRSF1, and PTBP1 was confirmed by RT-qPCR (A, B, C) and Western blot (D). GAPDH was used as an internal control (calibrator) for RT-qPCR experiments. Error bars represent 95% confidence intervals in A-C.

A



B

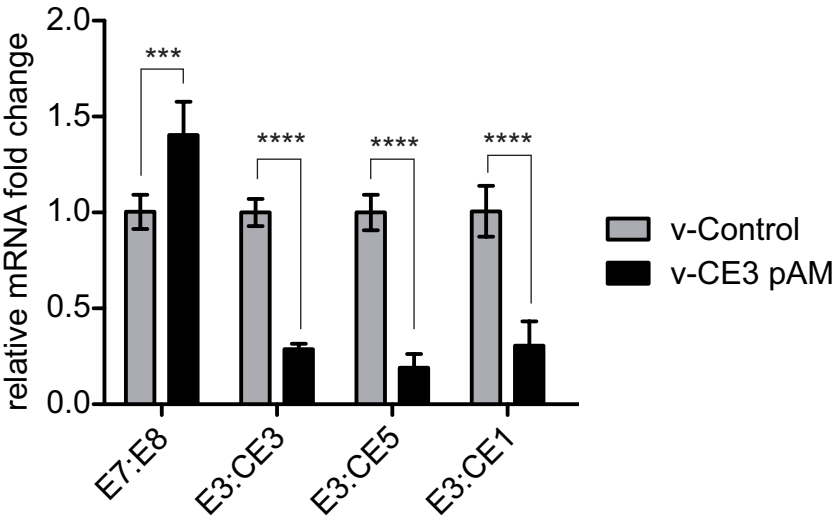


Figure S6: Blockade of the AR CE3 PAS blocks AR-V expression in A. VCaP and B. LNCaP 95 cells. Cells were incubated in the presence of 1 μM vivo-control morpholino (grey bars) or 1 μM vivo-CE3 pAM for 72 hours in androgen replete, reduced serum culture medium. Mean mRNA fold change of AR and AR-Vs relative to actin control from 3 biological replicates is shown (n=6, error bars= 95% confidence intervals). **** $p \leq 0.0001$, *** $p \leq 0.001$, ns $p \geq 0.05$, unpaired t-tests.

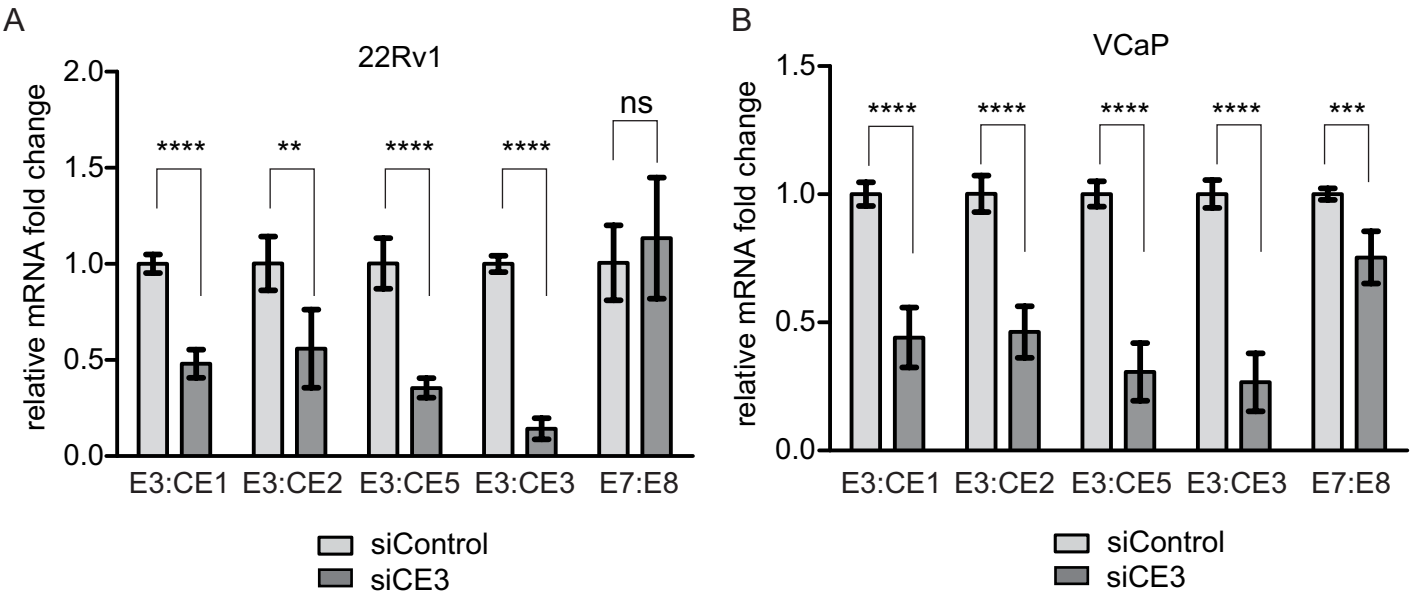


Figure S7: Knockdown of CE3 reduces relative levels of mRNAs harboring AR exon 3 spliced to CE1, CE2, and CE5 in 22Rv1 (A, n=4) and VCaP (B, n=6) cells. Bars represent mean mRNA fold changes relative to nontargeting control, error bars represent 95% confidence intervals. Asterisks indicate p values calculated by a two tailed t-test (p values; **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, ns $p > 0.05$). Actin was used as an internal control (calibrator) for qPCR studies.

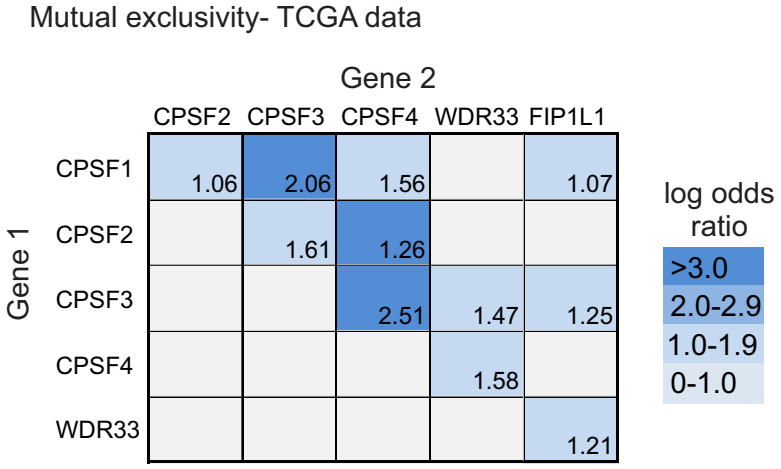


Figure S8: A. Mutual exclusivity matrix depicting co-occurrence of alterations in the core CPSF complex. Analysis was performed in cBioPortal on TCGA whole genome and RNA-Seq datasets. Positive log odds ratios from Fisher's exact test are depicted in blue.

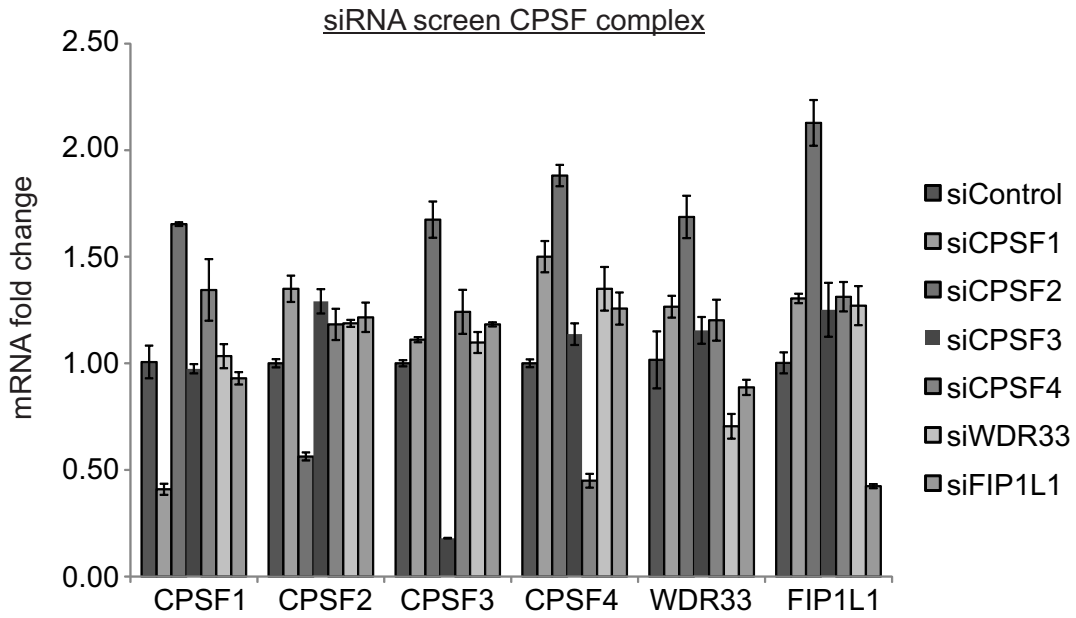


Figure S9: Confirmation of CPSF complex gene knockdown. siRNA pools targeting each CPSF core complex gene were transfected into 22Rv1 cells. Knockdown was assessed at the mRNA level by RT-qPCR. The comparative ΔΔCt method was used to calculate mRNA fold changes relative to 18S rRNA. Bars represent mean mRNA fold change. Error bars indicate standard deviation, n=3.

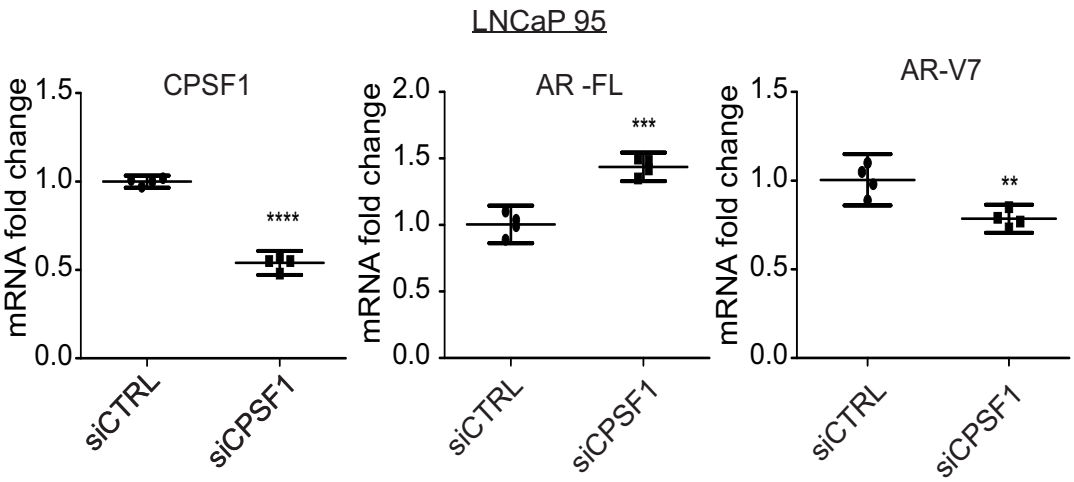


Figure S10. Relative mRNA levels of AR and AR-V7 after knockdown of CPSF1 in LNCaP 95 cells (siControl: nontargeting siRNA, siCPSF1: CPSF1 smartpool). Relative mRNA fold change from 2 biological and 2 technical replicates are plotted (n=4); the center line represents the mean, error bars represent 95% confidence intervals. 18S served as an internal control. **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, unpaired t-tests.