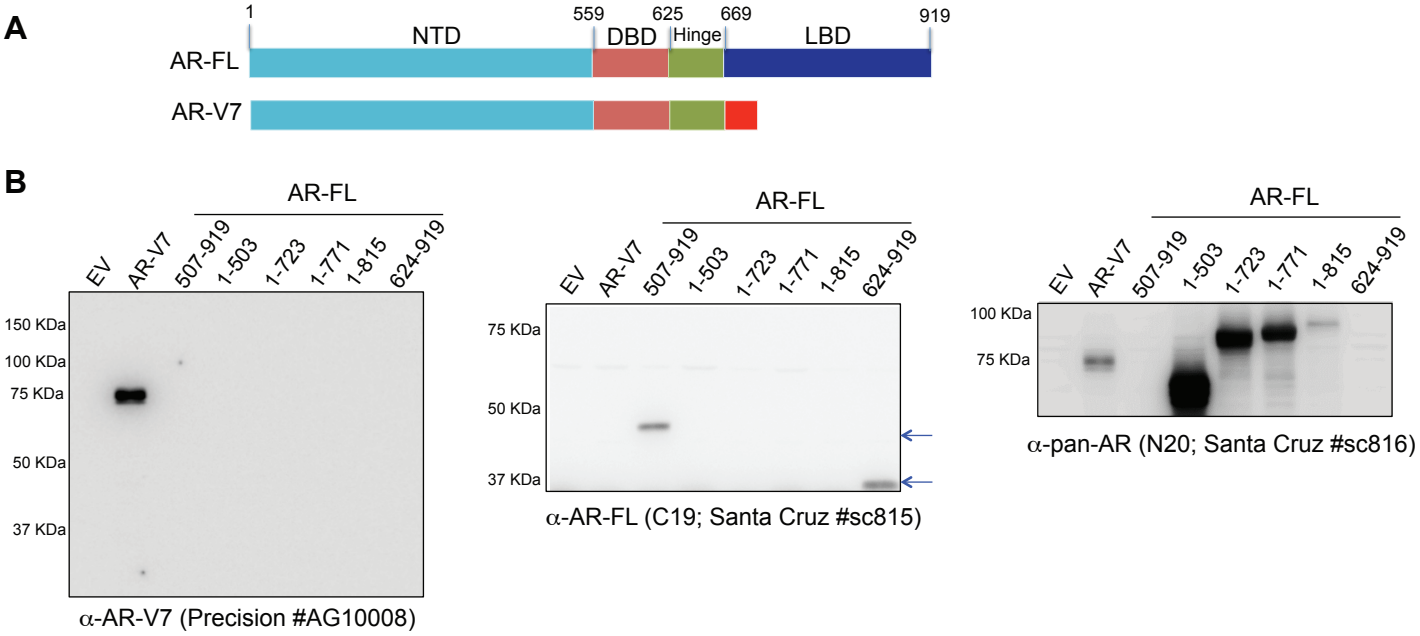
Supplementary Information

ZFX mediates non-canonical oncogenic functions of the androgen receptor splice variant 7 (AR-V7) in castrate-resistant prostate cancer

Ling Cai, Yi-Hsuan Tsai, Ping Wang, Jun Wang, Dongxu Li, Huitao Fan, Yilin Zhao, Rui Lu, Rohan Bareja, Elizabeth M. Wilson, Andrea Sboner, Young E. Whang, Deyou Zheng, Joel S. Parker, H. Shelton Earp, Gang Greg Wang

Fig. S1



C

Summary of ChIP-Seq read depth for the AR-FL, AR-V7 and BRD4 in 22RV1 and VCaP cells.

22Rv1		
Epitope	**Cell treatmet**	**Total # of raw Reads**
AR-FL C-terminus	vehicle (ligand-starved)	58,492,066
	DHT	39,764,623
	DHT plus JQ1	54,095,717
	DHT plus MDV3100	38,917,269
AR-V7 C-terminus	vehicle (ligand-starved)	66,521,128
	DHT	38,584,063
	DHT plus JQ1	54,418,919
	DHT plus MDV3100	45,804,204
BRD4	vehicle (ligand-starved)	18,682,596
	DHT	21,857,948
	DHT plus JQ1	20,854,481
	DHT plus MDV3100	24,138,480

VCaP		
Epitope	**Cell treatmet**	**Total # of raw Reads**
AR-FL	vehicle (ligand-starved)	37,382,769
	DHT	70,997,428
AR-V7	vehicle (ligand-starved)	36,267,228
	DHT	48,199,211

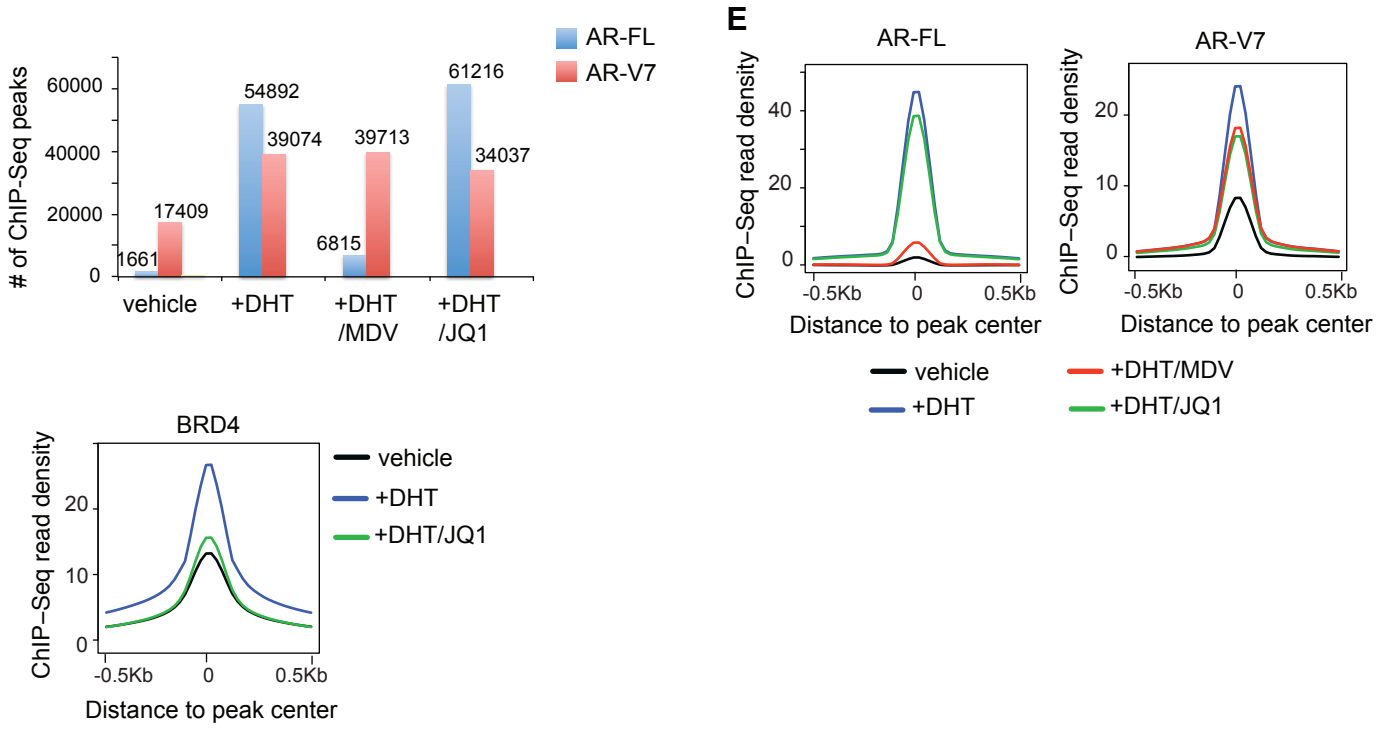


Figure S1. Summary and overview of ChIP-Seq peaks of endogenous AR-FL, AR-V7 and BRD4 in the 22Rv1 and VCaP cell models of castrate-resistant prostate cancer (CRPC), Related to Figure 1

(A) Protein domain structure of AR-FL and AR-V7. NTD, N-terminal transactivation domain; DBD, DNA-binding domain; LBD, ligand-binding domain.

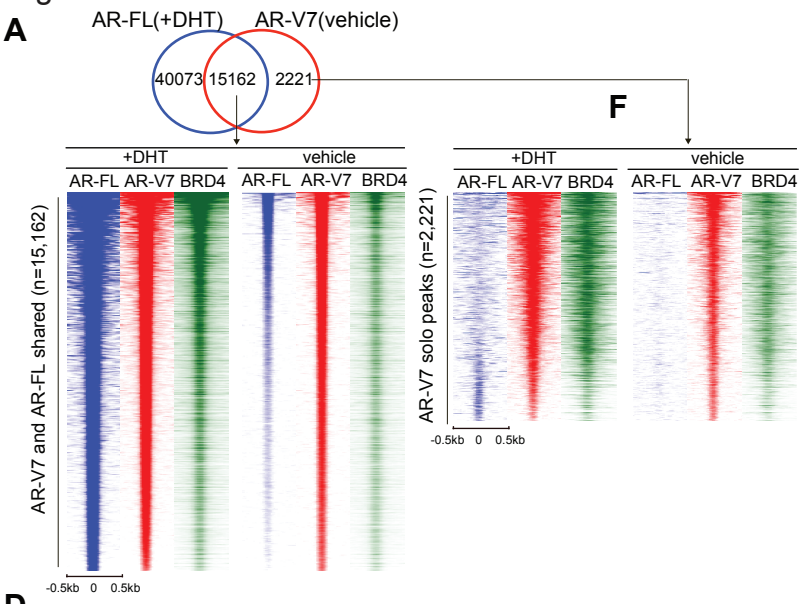
(B) Immunoblotting of total extracts of 293 cells transfected with 1 μg of each of the indicated AR-V7/FL wild-type or serial deletion plasmids with either AR-V7-specific (left), AR-FL-specific (middle) or pan-AR (right) antibody. EV, empty vector.

(C) Summary of the number of the ChIP-seq read tags for the AR-FL, AR-V7 and BRD4 in 22Rv1 (left) and VCaP (right) cells under the indicated treatment conditions for 6 hours.

(D) Total number of the called AR-FL and AR-V7 ChIP-Seq peaks in 22Rv1 cells that were cultured in the ligand-stripped medium followed by treatment with either vehicle, 10 nM of DHT, or DHT plus 10 μM of MDV3100 (+DHT/MDV), or DHT plus 500 nM of JQ1 (+DHT/JQ1).

(E-F) Averaged AR-FL, AR-V7 and BRD4 ChIP-Seq read density across the called peaks in 22Rv1 cells after the indicated treatment.

Fig. S2



B

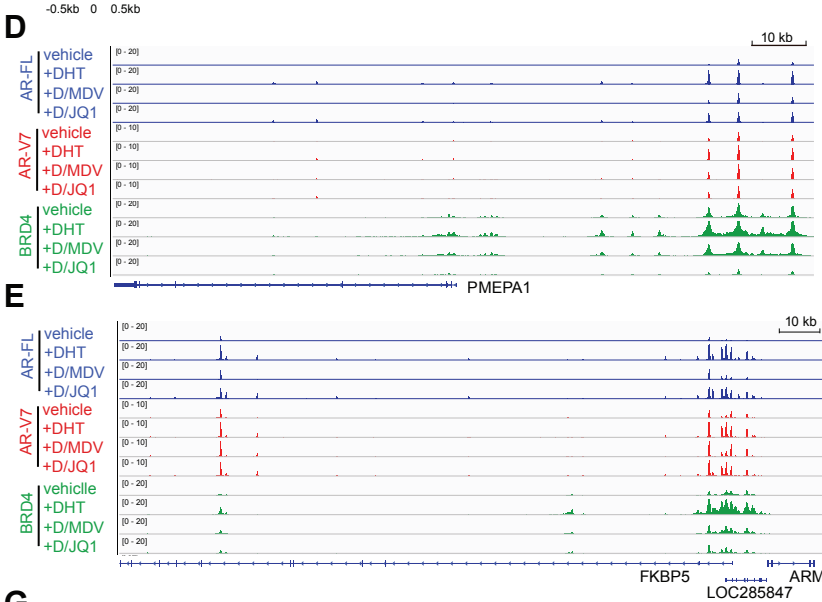
AR-FL(+DHT) sites by the Homer de novo motif search

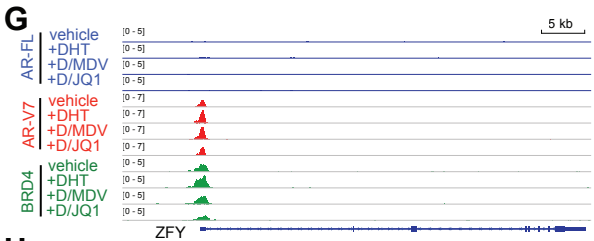
Rank	Seq of motif	p value	%target	% background	Best match/details
1		1e-611	42.39%	6.94%	MA0007.2_ARE /Jaspar 0.934
2		1e-301	45.57%	17.32%	FOXA1(Forkhead) /Homer 0.929
3		1e-100	31.90%	15.80%	FKH1/FKH2 /Yeast 0737
4		1e-83	38.16%	22.03%	MA0382.1_SKO1(dZIP) /Jaspar 0.745
5		1e-53	38.44%	25.26%	MA0161.1_NFIC /Jaspar 0.952
6		1e-47	40.50%	27.84%	MA0171.1_CG11085 /homeodomain/Jaspar 0.738
7		1e-29	11.40%	5.85%	ovo/dmmpmm(Bergman) /fly 0.736
8		1e-21	0.45%	0.01%	Antp/dmmpmm(SeSiMCMC) /fly 0.664
9		1e-29	11.40%	5.85%	Adf1/dmmpmm(Pollard) /fly 0.623
10		1e-21	0.45%	0.01%	dl/dmmpmm(Noyes) /fly 0.718

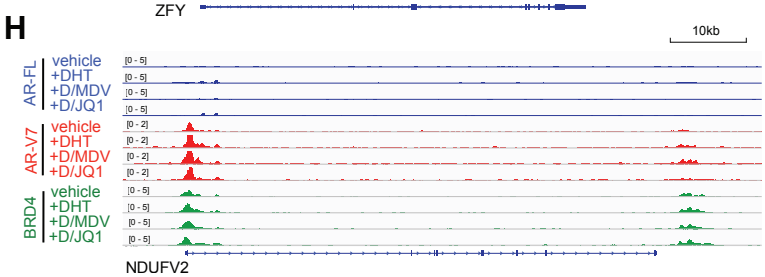
C

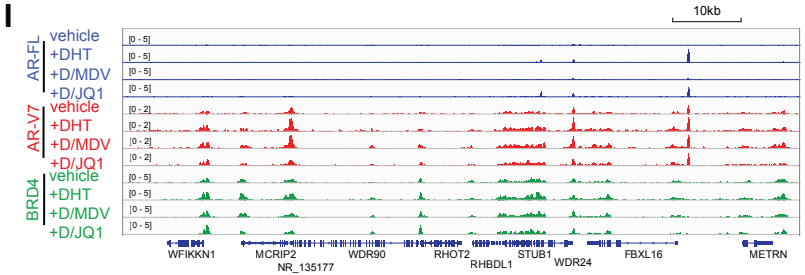
AR-V7(+DHT) sites by the Homer de novo motif search

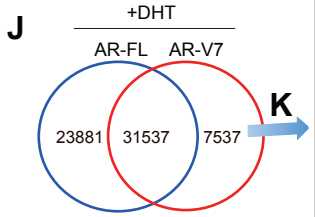
Rank	Seq of motif	p value	%target	% background	Best match/details
1		1e-505	20.28%	5.05%	MA0007.2_ARE /Jaspar 0.934
2		1e-425	31.14%	12.46%	FOXA1(Forkhead) /Homer 0.934
3		1e-182	22.56%	11.28%	ZNF711/ZFX(Zf) /Homer 0.876
4		1e-114	22.60%	13.32%	Pnt/dmmpmm(Papatsenko) /fly 0.878
5		1e-62	1.56%	0.22%	OPI1 /Yeast 0.817
6		1e-51	3.96%	1.50%	YY1(Zf) /Homer 0.902
7		1e-47	1.78%	0.40%	sd/dmmpmm(Bergman) /fly 0.683
8		1e-41	7.08%	3.84%	SWI6/(Harbison) /yeast 0.900
9		1e-41	18.48%	13.14%	ELF1(ETS) /Homer 0.753
10		1e-40	5.36%	2.65%	HOXD13(homeo domain) /Homer 0.851

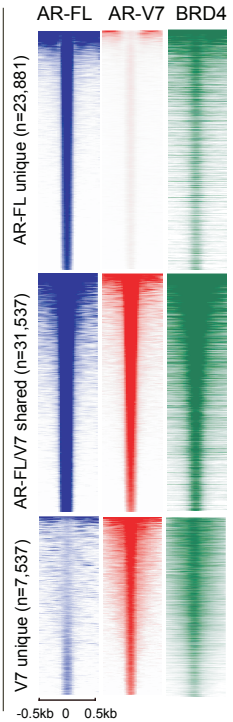












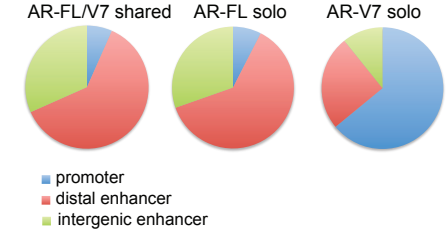


Figure S2. ChIP-Seq reveals both the AR-FL/AR-V7 common binding sites and those unique to AR-V7 in 22Rv1 cells, Related to Figure 1

(A) Venn diagram (top) and heatmap (bottom) showing ChIP-Seq peaks that are common to AR-V7 (red, middle column) in ligand-starved cells (vehicle) and AR-FL (blue, left column) in DHT-stimulated cells (+DHT), with the corresponding BRD4 binding (green) shown on the right column. Shown on each row of heatmap is ChIP-Seq signal density across ±0.5Kb from the center of a called peak, followed by ranking of the column based on density of AR-V7 binding in vehicle-treated cells.

(B-C) Top ranked motifs identified by Homer at either AR-FL (**B**) or AR-V7 (**C**) ChIP-seq peaks in 22Rv1 cells treated with DHT.

(D-E) ChIP-Seq profile reveals binding of AR-FL, AR-V7 and BRD4 at canonical AR target genes, PMEPA1 and FKBP5, in 22Rv1 cells, either ligand-starved (vehicle) or after the indicated treatment for 6 hours.

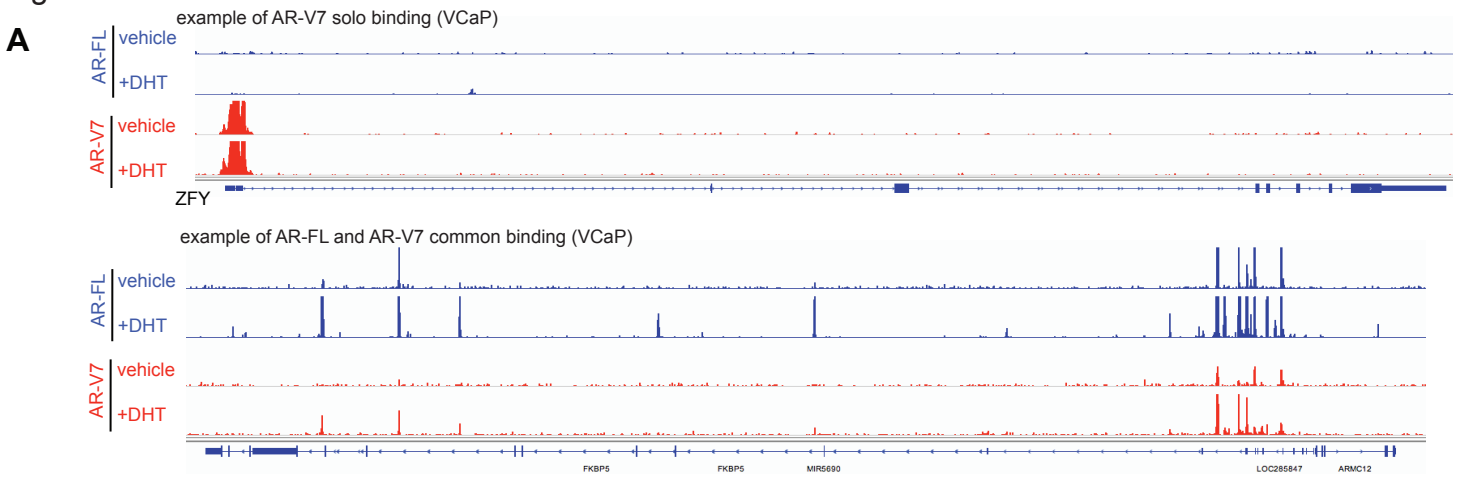
(F) Heatmap showing ChIP-Seq peaks unique to AR-V7 (red, middle column) in ligand-starved cells (vehicle) and lacking AR-FL binding (blue, left column) in the DHT-stimulated condition. Shown on each row of heatmap is ChIP-Seq signal density across ±0.5Kb from the center of a called peak, followed by ranking of the column based on density of AR-V7 binding. The corresponding BRD4 binding (green) at each peak is shown on the right.

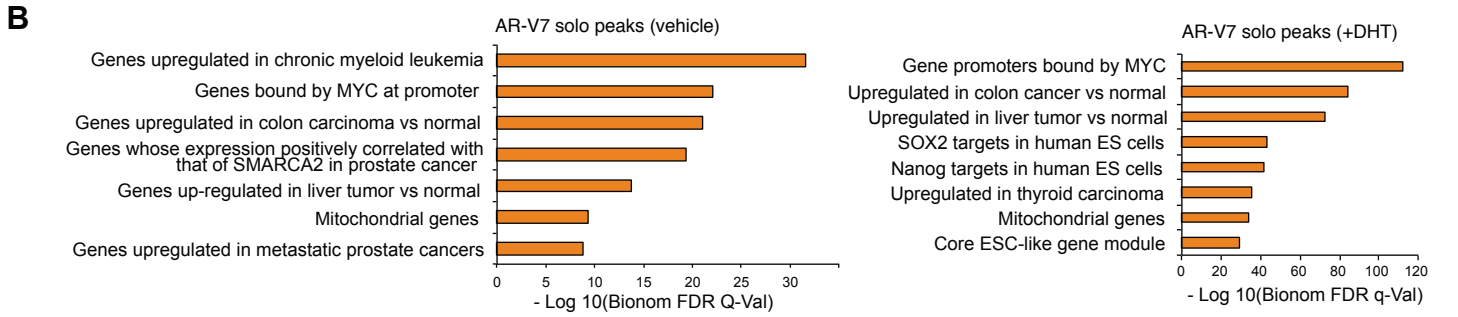
(G-I) ChIP-Seq profiles reveal the unique binding of AR-V7 (red) and lacking AR-FL (blue) binding at the indicated gene promoters in 22Rv1 cells after drug treatment.

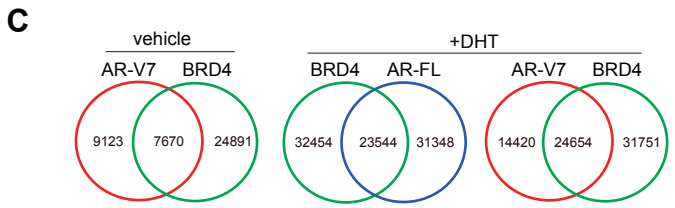
(J-K) Venn diagram (panel **J**) and heatmap (panel **K**) showing ChIP-Seq peaks defined as either AR-FL unique (**K**, top), common to AR-FL and AR-V7 (**K**, middle), or AR-V7 unique (**K**, bottom) in the 22Rv1 CRPC cells after DHT treatment. Shown on each row of heatmap is ChIP-Seq signal density across ±0.5Kb from the center of a called peak, followed by ranking of the column based on density of AR binding. The corresponding BRD4 binding at each peak is shown on the right column.

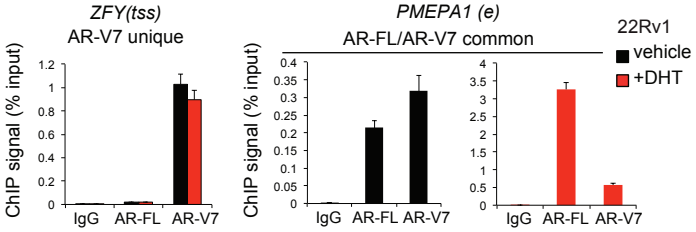
(L) Pie chart showing distribution of the indicated AR-FL and AR-V7 common and solo ChIP-Seq peaks as defined in panel **J** among promoters, distal enhancers or intergenic sites in the DHT-treated 22Rv1 cells.

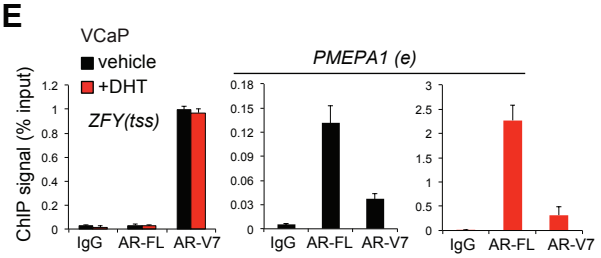
Fig. S3

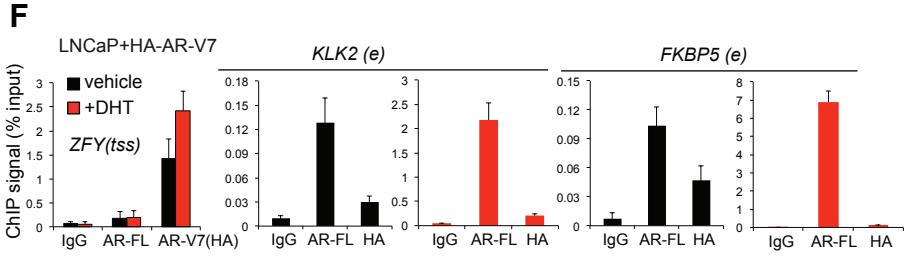

A
example of AR-V7 solo binding (VCaP)
AR-FL
vehicle
+DHT
AR-V7
vehicle
+DHT
ZFY
example of AR-FL and AR-V7 common binding (VCaP)
AR-FL
vehicle
+DHT
AR-V7
vehicle
+DHT
FKBP5
FKBP5
MIR5690
LOC285847
ARMC12

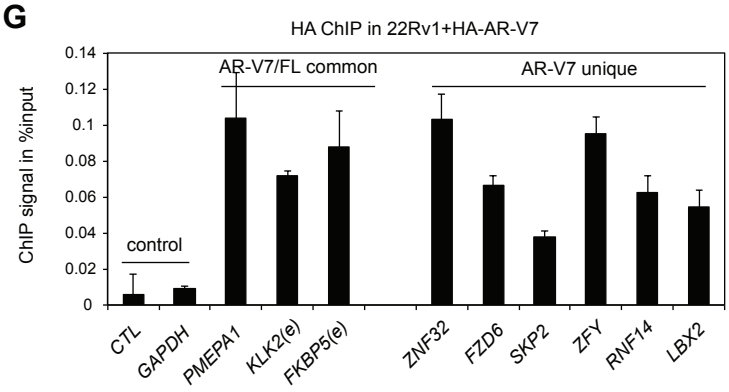

B
AR-V7 solo peaks (vehicle)
Genes upregulated in chronic myeloid leukemia
Genes bound by MYC at promoter
Genes upregulated in colon carcinoma vs normal
Genes whose expression positively correlated with that of SMARCA2 in prostate cancer
Genes up-regulated in liver tumor vs normal
Mitochondrial genes
Genes upregulated in metastatic prostate cancers
0 5 10 15 20 25 30
- Log 10(Bionom FDR Q-Val)
AR-V7 solo peaks (+DHT)
Gene promoters bound by MYC
Upregulated in colon cancer vs normal
Upregulated in liver tumor vs normal
SOX2 targets in human ES cells
Nanog targets in human ES cells
Upregulated in thyroid carcinoma
Mitochondrial genes
Core ESC-like gene module
0 20 40 60 80 100 120
- Log 10(Bionom FDR q-Val)


C
vehicle
AR-V7
BRD4
9123
7670
24891
+DHT
BRD4
AR-FL
32454
23544
31348
AR-V7
BRD4
14420
24654
31751


D
ZFY(tss)
AR-V7 unique
ChIP signal (% input)
IgG AR-FL AR-V7
PMEPA1 (e)
AR-FL/AR-V7 common
22Rv1
vehicle
+DHT


E
VCaP
vehicle
+DHT
ZFY(tss)
PMEPA1 (e)
ChIP signal (% input)
IgG AR-FL AR-V7


F
LNCaP+HA-AR-V7
vehicle
+DHT
ZFY(tss)
KLK2 (e)
FKBP5 (e)
ChIP signal (% input)
IgG AR-FL AR-V7(HA)
IgG AR-FL HA


G
HA ChIP in 22Rv1+HA-AR-V7
ChIP signal in %input
AR-V7/FL common
AR-V7 unique
control
CTL
GAPDH
PMEPA1
KLK2(e)
FKBP5(e)
ZNF32
FZD6
SKP2
ZFY
RNF14
LBX2

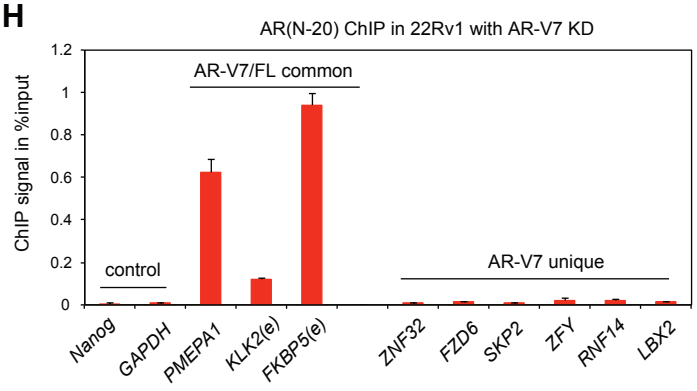

H
AR(N-20) ChIP in 22Rv1 with AR-V7 KD
ChIP signal in %input
AR-V7/FL common
AR-V7 unique
control
Nanog
GAPDH
PMEPA1
KLK2(e)
FKBP5(e)
ZNF32
FZD6
SKP2
ZFY
RNF14
LBX2

Figure S3. Validation of AR-FL and AR-V7 common and unique binding with multiple independent CRPC cell models, Related to Figure 1.

(A) ChIP-Seq profile examples show either the AR-FL (blue) and AR-V7 (red) common binding (bottom; FKBP5) or AR-V7 unique binding (top panel; the ZFY promoter) in VCaP cells.

(B) Genomic features of genes associated with AR-V7 solo binding. The Genomic Regions Enrichment of Annotations Tool (GREAT) ontology analysis reveals enrichment of the indicated signatures in genes associated with the AR-V7 solo binding peaks identified in either ligand-starved (left) or DHT-stimulated (right) 22Rv1 cells.

(C) Venn diagram showing overlap of BRD4 ChIP-Seq peaks with AR-FL or AR-V7 peaks in ligand-starved (left) or DHT-stimulated (right) 22Rv1 cells.

(D-F) ChIP-qPCR validation of the AR-V7 unique binding (left panel) and the AR-V7 and AR-FL common binding (right) at the indicated locus in three different CRPC cell lines, i.e., 22Rv1 (panel **D**), VCaP (**E**) and an HA-tagged AR-V7-overexpressing LNCaP cell line (**F**), under the ligand-starved (vehicle; black) or DHT-stimulated (+DHT; red) culture conditions. Plotted are ChIP signals from 3 independent experiments that are normalized to input and presented as mean ± SD. tss, transcription start site; e, ARE enhancer. Note that the AR ChIP signals increased from mock to DHT-stimulated conditions.

(G-H) Independent approaches for ChIP verification using either anti-HA antibodies in 22Rv1 cells with stable expression of an HA-tagged AR-V7 (panel **G**) or using a pan-AR antibody (AR N-20; Santa Cruz Bio) in 22Rv1 cells with specific knock-down of AR-V7 isoform (panel **H**; see also Fig 2B) at the indicated genomic loci. 22Rv1 cells were ligand-starved followed by a 6-hour treatment with 10 nM of DHT.

Fig. S4

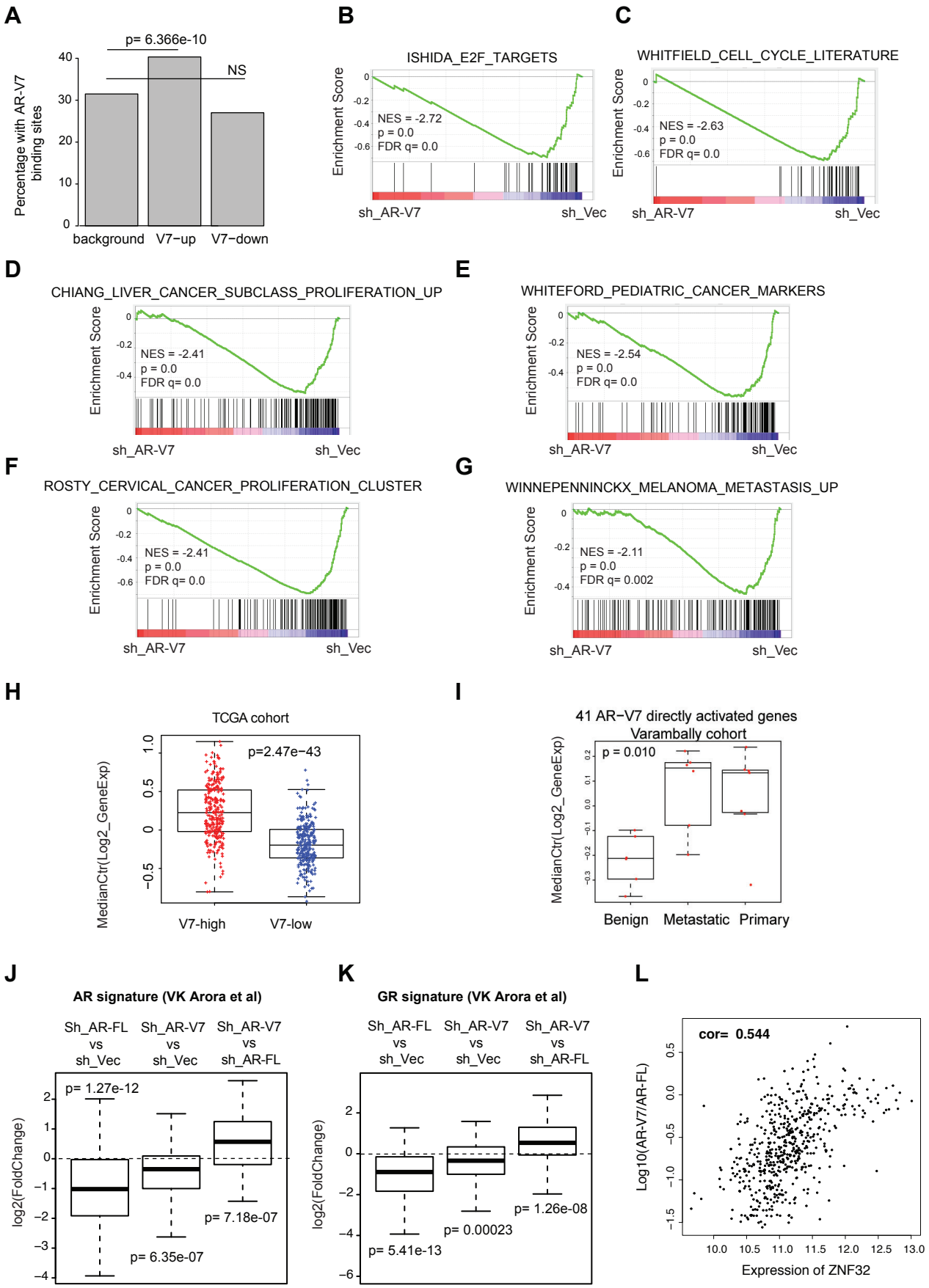

A
p= 6.366e-10
NS
Percentage with AR-V7 binding sites
background
V7−up
V7−down
B
ISHIDA_E2F_TARGETS
Enrichment Score
NES = -2.72
p = 0.0
FDR q= 0.0
sh_AR-V7
sh_Vec
C
WHITFIELD_CELL_CYCLE_LITERATURE
NES = -2.63
p = 0.0
FDR q= 0.0
D
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP
NES = -2.41
p = 0.0
FDR q= 0.0
E
WHITEFORD_PEDIATRIC_CANCER_MARKERS
NES = -2.54
p = 0.0
FDR q= 0.0
F
ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER
NES = -2.41
p = 0.0
FDR q= 0.0
G
WINNEPENNINCKX_MELANOMA_METASTASIS_UP
NES = -2.11
p = 0.0
FDR q= 0.002
H
TCGA cohort
p=2.47e−43
MedianCtr(Log2_GeneExp)
V7-high
V7-low
I
41 AR−V7 directly activated genes
Varambally cohort
p = 0.010
Benign
Metastatic
Primary
J
AR signature (VK Arora et al)
Sh_AR-FL vs sh_Vec
Sh_AR-V7 vs sh_Vec
Sh_AR-V7 vs sh_AR-FL
p= 1.27e-12
p= 6.35e-07
p= 7.18e-07
log2(FoldChange)
K
GR signature (VK Arora et al)
p= 5.41e-13
p= 0.00023
p= 1.26e-08
L
cor= 0.544
Log10(AR-V7/AR-FL)
Expression of ZNF32

Figure S4. Analysis of gene pathways affected by specific KD of either the AR-V7 or AR-FL isoform in 22Rv1 cells, Related to Figure 2-3.

(A) Integration of ChIP-Seq and RNA-Seq data reveals that significantly more of the genes up-regulated by AR-V7 (V7-up) show direct binding of AR-V7, in comparison to randomized control (background) or genes down-regulated by AR-V7 (V7-down).

(B-G) GSEA shows negative correlation of the indicated cell proliferative (**B-C**) and cancer-related (**D-G**) genes with selective KD of AR-V7 (sh_AR-V7), relative to mock (sh_Vec).

(H) Boxplot of mean expression of the select 41 genes in AR-V7-high and low groups, separated by the median of AR-V7 expression level. These 41 genes are identified in a separate analysis to be directly up-regulated by AR-V7 in 22Rv1 cells and also positively correlated (spearman rank r >0.2 and BH FDR <0.01) with the AR-V7 expression level in the TCGA prostate cancer cohort(Cancer Genome Atlas Research, 2015). Gene expression values in boxplots were upper-quantile normalized, log2 transformed and centered to its median. ANOVA test p-value was denoted on top.

(I) Box plots showing mean expression values of a 41-gene AR-V7 direct target signature (as defined in panel **H**) in a prostate cohort from the Varambally et al study (Varambally et al., 2008). Gene expression values (y-axis) were upper-quantile normalized, log2 transformed and centered to its median.

(J-K) Changes in expression of the previously defined AR or GR signature genes (Arora et al., 2013) in ligand-starved 22Rv1 with selective KD of AR-V7 or AR-FL versus mock treatment (sh_Vec). Plotted on y-axis are log2-converted fold-change values for the indicated comparison.

(L) Correlation analysis shows a positive relationship between ZNF32 and AR-V7 expression levels among the TCGA prostate cancer samples. Plotted on y-axis are the log10 (the ratio of RNA-Seq tag number for cryptic exon to that for the common NTD exons), which is the calculated relative expression level for AR-V7. X-axis is the log2 (up-quantile normalized read count of ZNF32).

Fig. S5

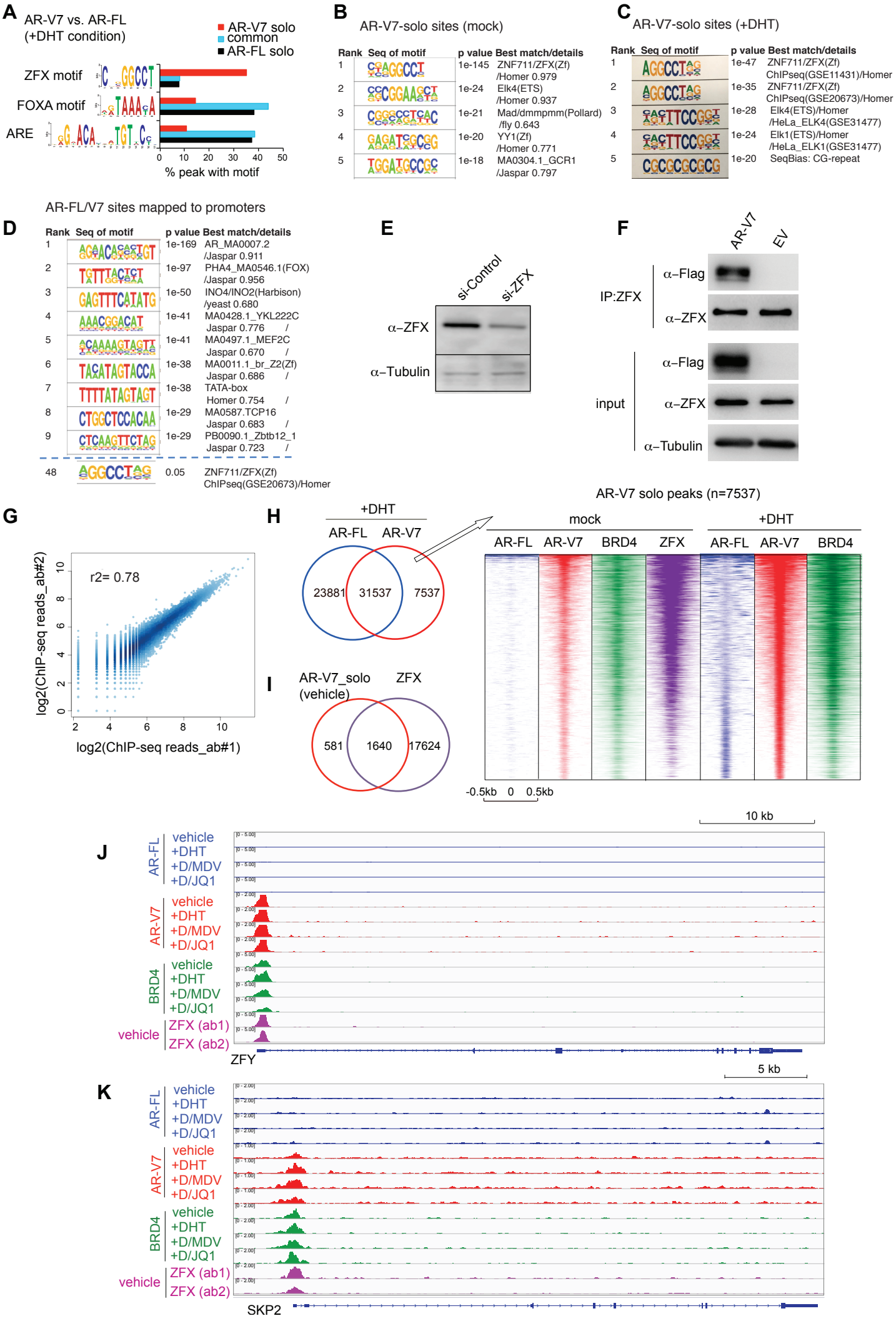

A
AR-V7 vs. AR-FL
(+DHT condition)
AR-V7 solo
common
AR-FL solo
ZFX motif
FOXA motif
ARE
0 10 20 30 40 50
% peak with motif
B
AR-V7-solo sites (mock)
Rank Seq of motif p value Best match/details
1 1e-145 ZNF711/ZFX(Zf) /Homer 0.979
2 1e-24 Elk4(ETS) /Homer 0.937
3 1e-21 Mad/dmmpmm(Pollard) /fly 0.643
4 1e-20 YY1(Zf) /Homer 0.771
5 1e-18 MA0304.1_GCR1 /Jaspar 0.797
C
AR-V7-solo sites (+DHT)
Rank Seq of motif p value Best match/details
1 1e-47 ZNF711/ZFX(Zf) ChIPseq(GSE11431)/Homer
2 1e-35 ZNF711/ZFX(Zf) ChIPseq(GSE20673)/Homer
3 1e-28 Elk4(ETS)/Homer /HeLa_ELK4(GSE31477)
4 1e-24 Elk1(ETS)/Homer /HeLa_ELK1(GSE31477)
5 1e-20 SeqBias: CG-repeat
D
AR-FL/V7 sites mapped to promoters
Rank Seq of motif p value Best match/details
1 1e-169 AR_MA0007.2 /Jaspar 0.911
2 1e-97 PHA4_MA0546.1(FOX) /Jaspar 0.956
3 1e-50 INO4/INO2(Harbison) /yeast 0.680
4 1e-41 MA0428.1_YKL222C Jaspar 0.776 /
5 1e-41 MA0497.1_MEF2C Jaspar 0.670 /
6 1e-38 MA0011.1_br_Z2(Zf) Jaspar 0.686 /
7 1e-38 TATA-box Homer 0.754 /
8 1e-29 MA0587.TCP16 Jaspar 0.683 /
9 1e-29 PB0090.1_Zbtb12_1 Jaspar 0.723 /
48 0.05 ZNF711/ZFX(Zf) ChIPseq(GSE20673)/Homer
E
si-Control
si-ZFX
α–ZFX
α–Tubulin
F
AR-V7
EV
IP:ZFX
α–Flag
α–ZFX
input
α–Flag
α–ZFX
α–Tubulin
G
r2= 0.78
log2(ChIP-seq reads_ab#2)
log2(ChIP-seq reads_ab#1)
H
+DHT
AR-FL AR-V7
23881 31537 7537
AR-V7 solo peaks (n=7537)
mock
+DHT
AR-FL AR-V7 BRD4 ZFX
AR-FL AR-V7 BRD4
-0.5kb 0 0.5kb
I
AR-V7_solo (vehicle)
ZFX
581 1640 17624
10 kb
J
AR-FL
vehicle
+DHT
+D/MDV
+D/JQ1
AR-V7
vehicle
+DHT
+D/MDV
+D/JQ1
BRD4
vehicle
+DHT
+D/MDV
+D/JQ1
vehicle
ZFX (ab1)
ZFX (ab2)
ZFY
5 kb
K
AR-FL
vehicle
+DHT
+D/MDV
+D/JQ1
AR-V7
vehicle
+DHT
+D/MDV
+D/JQ1
BRD4
vehicle
+DHT
+D/MDV
+D/JQ1
vehicle
ZFX (ab1)
ZFX (ab2)
SKP2

Figure S5. ChIP-Seq reveals co-recruitment of ZFX to the AR-V7 solo binding sites in 22Rv1 cells, Related to Figure 4.

(A) Percentage of AR-V7 unique (red), AR-FL/V7 common (cyan) or AR-FL unique (black) peaks that contain the indicated motif, as suggested by the HOMER motif search analysis using ChIP-Seq data from the DHT-stimulated 22Rv1 cells.

(B-D) Top ranked motifs identified by Homer at AR-V7 solo ChIP-seq peaks defined in either ligand-starved (**B**) or DHT-treated (**C**) 22Rv1 cells, as well as the 768 AR-FL/V7 common peaks mapped to promoter (**D**; the motif of ZFX in **D** here is ranked as 48^{th} and not significantly enriched).

(E) Immublotting verifies specificity of the used ZFX antibodies in the 22RV1 cells after siRNA-mediated transient KD of ZFX (right), relative to mock (left).

(F) CoIP in 293 cells detects interaction of ZFX to the transiently transfected Flag-tagged AR-V7 (right). EV, empty vector.

(G) Correlation plot showing consistence of ChIP-Seq data generated with the two different ZFX antibodies used in the study.

(H) Venn diagram (left) and heatmap (right) showing the indicated ChIP-Seq signals at the peaks defined as unique to AR-V7 among the DHT-starved (left) and DHT-stimulated 22Rv1 cells (right). Shown on each row is ChIP-Seq signal density across ±0.5Kb from the center of a called peak, followed by ranking of the column based on density of AR-V7 binding. The corresponding BRD4 binding (green) at the peak is shown on the right column.

(I) Venn diagram illustrating overlap of ZFX peaks with AR-V7 solo binding peaks in the ligand-starved 22Rv1 cells.

(J-K) ChIP-Seq profiles of AR-FL, AR-V7, BRD and ZFX binding at the ZFY and SKP2 gene in 22Rv1 cells under the indicated treatment conditions.

Fig. S6

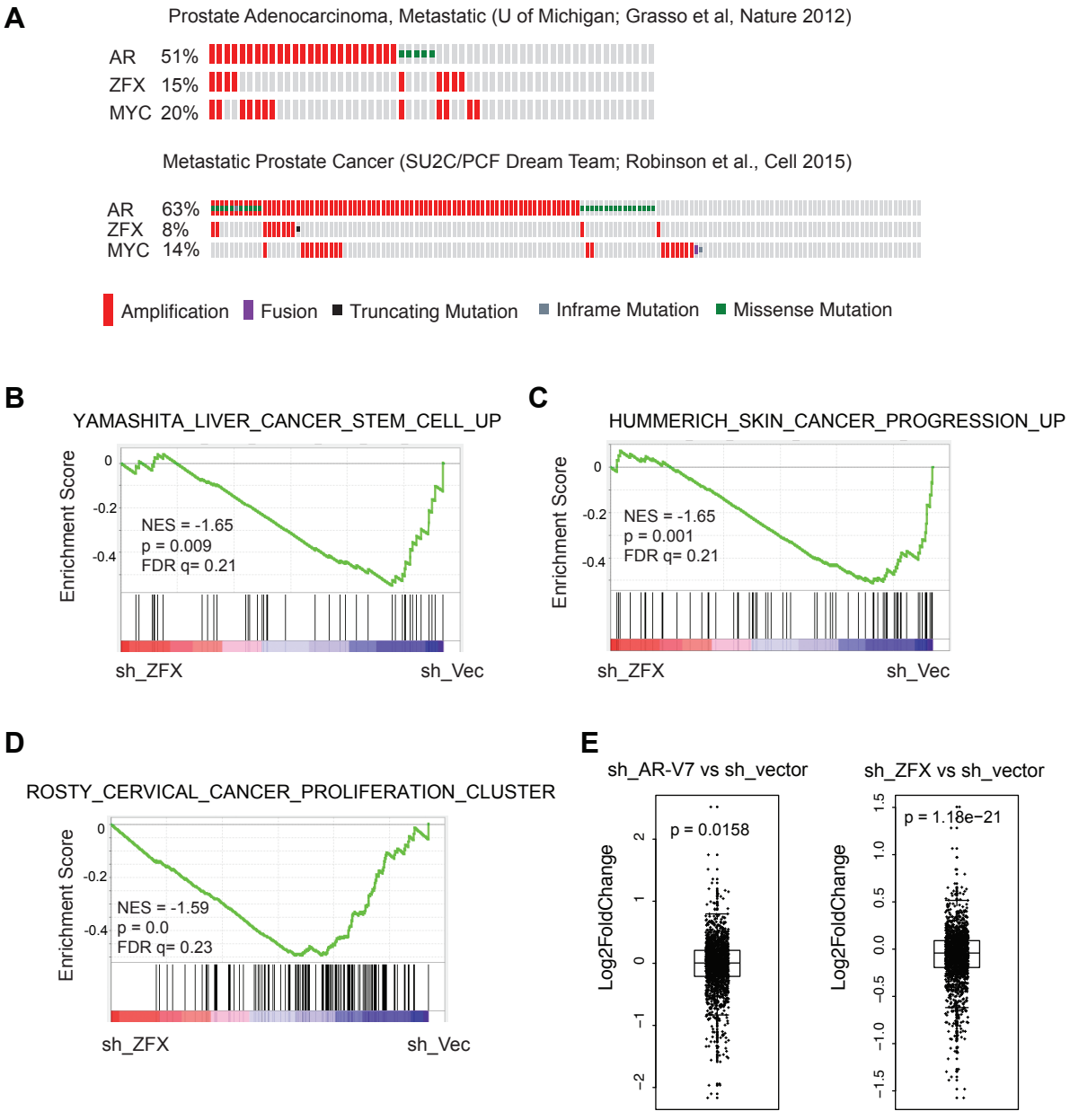

A
Prostate Adenocarcinoma, Metastatic (U of Michigan; Grasso et al, Nature 2012)
AR 51%
ZFX 15%
MYC 20%
Metastatic Prostate Cancer (SU2C/PCF Dream Team; Robinson et al., Cell 2015)
AR 63%
ZFX 8%
MYC 14%
Amplification
Fusion
Truncating Mutation
Inframe Mutation
Missense Mutation
B
YAMASHITA_LIVER_CANCER_STEM_CELL_UP
Enrichment Score
NES = -1.65
p = 0.009
FDR q= 0.21
sh_ZFX
sh_Vec
C
HUMMERICH_SKIN_CANCER_PROGRESSION_UP
Enrichment Score
NES = -1.65
p = 0.001
FDR q= 0.21
sh_ZFX
sh_Vec
D
ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER
Enrichment Score
NES = -1.59
p = 0.0
FDR q= 0.23
sh_ZFX
sh_Vec
E
sh_AR-V7 vs sh_vector
Log2FoldChange
p = 0.0158
sh_ZFX vs sh_vector
Log2FoldChange
p = 1.18e−21

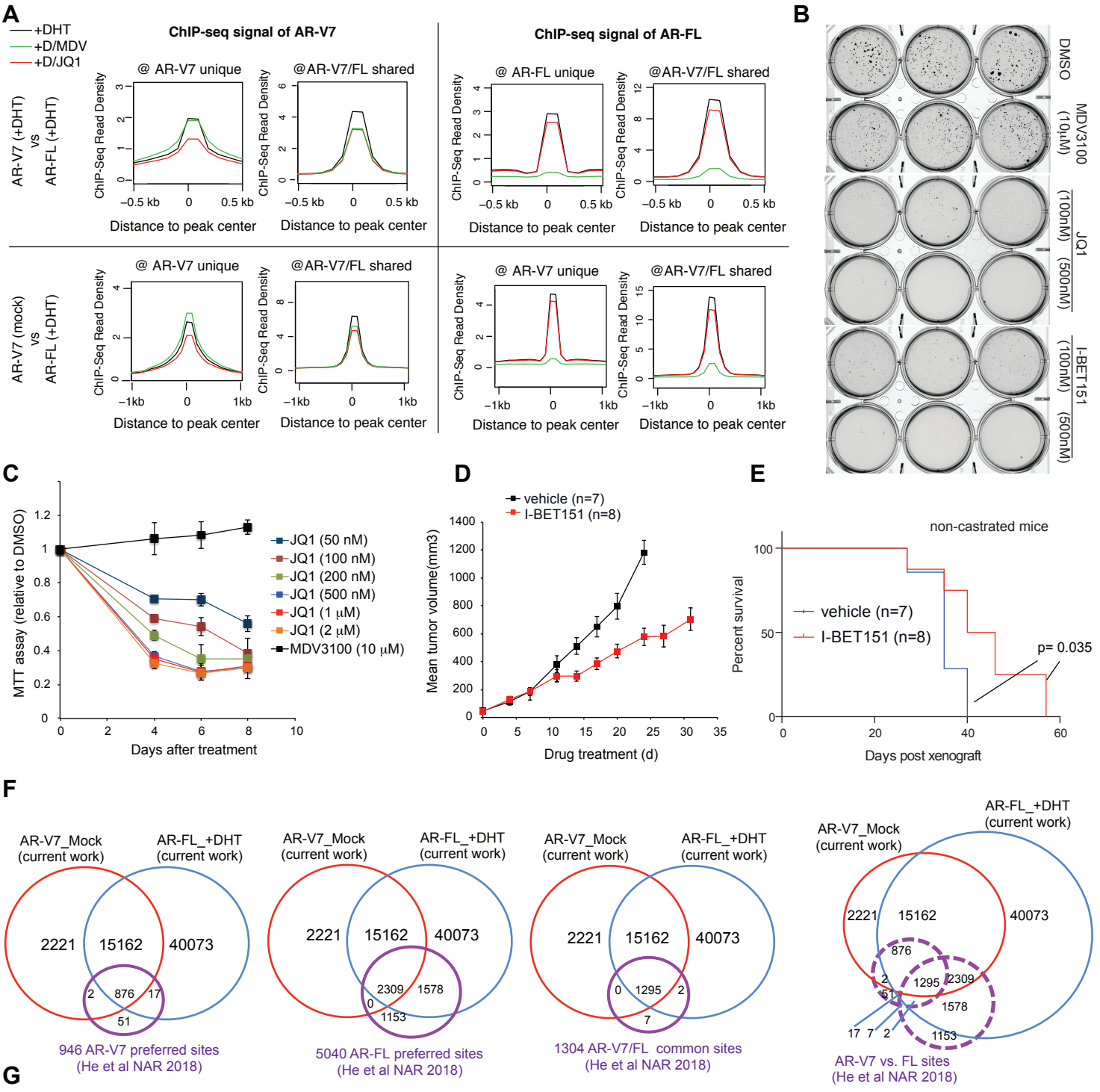
Figure S6. ZFX shows gene amplification in prostate cancers and is required for AR-V7-mediated gene expression programs in the 22Rv1 cells, Related to Figure 5-6.

(A) Amplification of AR, ZFX and MYC among various prostate cancer cohorts in the previous studies. Datasets of the indicated studies (Grasso et al., 2012; Kumar et al., 2016; Robinson et al., 2015) are extracted from cBioportal for Cancer Genomics (http://www.cbioportal.org).

(B-D) GSEA shows negative correlation of the indicated gene set to ZFX KD (sh_ZFX), relative to mock (sh_Vec).

(E) Box plots showing fold-change in expression of the 1603 genes associated with the defined AR-V7-solo peaks (see Table S5), as revealed by RNA-Seq profiles of 22Rv1 cells after KD of AR-V7 (left) or ZFX (right), relative to mock. Y-axis shows log2(fold change). Each dot in plot represents a gene.

Fig. S7



G

	# of pan-AR peaks	# of AR-FL peaks	cell line	Fold (vs. He et al)	antibody	epitope	GSMID (GEO)
He et al NAR 2018	ND	6,344	22Rv1 (R1881)	**1.0**	anti-AR, C-19 (Santa Cruz , SC-815)	AR-FL specific	GSE80743
current work	ND	54,892	22Rv1 (DHT)	**8.7**	anti-AR, C-19 (Santa Cruz , SC-815)	AR-FL specific	GSE94013
Arul M. Chinnaiyan et al Nature 2014	34,606	ND	VCaP (vehicle)	**5.5**	AR_PG-21 (Millipore Cat. # 06-680)	AR N-terminus	GSM1328947
	59,465	ND	VCaP (DHT)	**9.4**	AR_PG-21 (Millipore Cat. # 06-680)	AR N-terminus	GSM1328952
	61,618	ND	VCaP (DHT Bicalutamide)	**9.7**	AR_PG-21 (Millipore Cat. # 06-680)	AR N-terminus	GSM1328958

	# of pan-AR peaks	# of AR-V7 peaks	cell line	Fold (vs. He et al)	antibody	epitope	GSMID (GEO)
He et al NAR 2018	ND	2,250	22Rv1 (vehicle; KD of AR-FL)	**1.0**	AR N terminus, N-20 (santa cruz cat# SC-816)	AR N-terminus	GSE80743
current work	ND	17,409	22Rv1 (vehicle)	**7.7**	AR-V7 specific (Precision # AG10008)	AR-V7 specific	GSE94013
	ND	39,074	22Rv1 (DHT)	**17.4**	AR-V7 specific (Precision # AG10008)	AR-V7 specific	GSE94013

Figure S7. Effects of compound treatment on AR-V7/FL binding, expression of AR-V7-associated signature genes and growth of 22Rv1 cells in vitro and in vivo, Related to Figure 7.

(A) Averaged ChIP-Seq read density of AR-V7 (two panels on the left) or AR-FL (two panels on the right) at the AR-V7-solo peaks (1st and 3rd columns) or AR-FL/V7 shared peaks (2nd and 4th columns). The determination of AR-V7 solo binding is based on AR-V7 binding identified in either DHT-treated (top panel) or ligand-starved (bottom) 22Rv1 cells, relative to AR-FL peaks in DHT-treated 22Rv1 cells.

(B) Colony formation of 22Rv1 cells under the indicated treatment condition.

(C) MTT assays measuring effect of the indicated compound treatment on in vitro proliferation of 22Rv1 cells cultured under ligand-starved conditions. Plotted are data sets of 3 independent experiments after normalization to the corresponding vehicle controls.

(D) In vivo growth of 22Rv1 cells xenografted in non-castrated NSG mice, followed by the indicated treatment. n, size of cohort.

(E) Kaplan-Meier curve of non-castrated NSG mice after subcutaneous inoculation of 22Rv1 cells and the indicated treatment. n, size of cohort; NS, not significant.

(F) Venn diagrams comparing the ChIP-Seq peaks of AR-V7 (ligand-starved, top/red) and AR-FL (DHT-treated, top/blue) identified in 22Rv1 cells by the current work with the AR-V7-preferred (1st panel, bottom/purple), AR-FL-preferred (2nd panel, bottom/purple) or common peaks (3rd panel, bottom/purple) as reported by a recent work (He et al., 2018) in 22Rv1 cells. Right panel shows overall comparison for AR-V7/FL sites reported by the two studies.

(G) Summary of the total number of pan-AR, AR-FL-specific and AR-V7-specific ChIP-Seq peaks reported in recent studies (Asangani et al., 2014; He et al., 2018) and this current work, with the information of the used cells, antibody and dataset listed.

Table S3. 41 AR-V7 directly activated genes in 22Rv1 cell that positively correlate to the AR-V7 level in the TCGA prostate cancer patient cohort, Related to Figure 2.

* V7.log2FC, log2 value of fold change in expression for sh_AR_V7 vs sh_Vector in 22Rv1 cells;
padj, adjusted p value; V7peaks, existence of a called AR-V7 ChIP-seq peak in 22Rv1 cells.

Symbol	entrez ID	description	V7.log2FC*	V7.pvalue	V7.padj	V7peaks
ALB	213	albumin [Source:HGNC Symbol;Acc:HGNC:399]	-1.834732153	1.51E-93	7.56E-91	yes
ANKRD16	54522	ankyrin repeat domain 16 [Source:HGNC Symbol;Acc:HGI	-0.839958619	3.86E-05	0.000229658	yes
BCAS4	55653	breast carcinoma amplified sequence 4 [Source:HGNC Sy	-0.908526197	0.00079286	0.003341616	yes
BIRC5	332	baculoviral IAP repeat containing 5 [Source:HGNC Symbo	-1.01325133	2.13E-17	6.45E-16	yes
C14orf166	51637	chromosome 14 open reading frame 166 [Source:HGNC S	-1.150385031	2.49E-26	1.48E-24	yes
CCBL1	883	cysteine conjugate-beta lyase, cytoplasmic [Source:HGN	-0.665250828	0.002383626	0.008698995	yes
CHKA	1119	choline kinase alpha [Source:HGNC Symbol;Acc:HGNC:19	-0.671872884	9.86E-05	0.000531377	yes
CMTM7	112616	CKLF-like MARVEL transmembrane domain containing 7 [	-0.78996647	0.002755252	0.009923438	yes
CNIH4	29097	cornichon family AMPA receptor auxiliary protein 4 [Sou	-0.627393764	6.71E-09	7.71E-08	yes
COPS7B	64708	COP9 signalosome subunit 7B [Source:HGNC Symbol;Acc	-0.994714761	6.70E-17	1.95E-15	yes
DSN1	79980	DSN1 homolog, MIS12 kinetochore complex component	-0.643876483	1.45E-06	1.15E-05	yes
ETHE1	23474	ethylmalonic encephalopathy 1 [Source:HGNC Symbol;Ac	-0.980144463	1.89E-05	0.000120547	yes
EVL	51466	Enah/Vasp-like [Source:HGNC Symbol;Acc:HGNC:20234]	-1.116843248	9.11E-10	1.19E-08	yes
FAM174B	400451	family with sequence similarity 174, member B [Source:H	-0.60782139	0.001157403	0.004652853	yes
FAM200B	285550	family with sequence similarity 200, member B [Source:H	-0.582412747	0.000560362	0.0024625	yes
FDX1	2230	ferredoxin 1 [Source:HGNC Symbol;Acc:HGNC:3638]	-0.644009545	2.62E-05	0.000162299	yes
HMGB1	3146	high mobility group box 1 [Source:HGNC Symbol;Acc:HGI	-0.711584313	1.18E-22	5.34E-21	yes
HPN	3249	hepsin [Source:HGNC Symbol;Acc:HGNC:5155]	-0.621739442	3.00E-07	2.66E-06	yes
IGFBP2	3485	insulin-like growth factor binding protein 2, 36kDa [Sourc	-0.705227281	4.61E-17	1.37E-15	yes
ITPK1	3705	inositol-tetrakisphosphate 1-kinase [Source:HGNC Symbc	-0.627102081	2.26E-13	4.66E-12	yes
KIF12	113220	kinesin family member 12 [Source:HGNC Symbol;Acc:HGI	-0.629184444	1.75E-09	2.20E-08	yes
KLKP1	606293	kallikrein pseudogene 1 [Source:HGNC Symbol;Acc:HGNC	-1.934951544	6.03E-09	7.00E-08	yes
MEGF6	1953	multiple EGF-like-domains 6 [Source:HGNC Symbol;Acc:H	-0.757042584	3.06E-07	2.71E-06	yes
MESP1	55897	mesoderm posterior bHLH transcription factor 1 [Source:	-1.441258524	4.44E-10	6.06E-09	yes
MMAB	326625	methylmalonic aciduria (cobalamin deficiency) cblB type	-0.830484514	2.18E-16	6.12E-15	yes
MMP24	10893	matrix metallopeptidase 24 (membrane-inserted) [Sourc	-0.767168041	9.35E-12	1.55E-10	yes
MSI1	4440	musashi RNA-binding protein 1 [Source:HGNC Symbol;Ac	-0.861063611	1.05E-10	1.54E-09	yes
MTMR7	9108	myotubularin related protein 7 [Source:HGNC Symbol;Ac	-1.241774191	0.000210603	0.00103521	yes
PIGY	84992	phosphatidylinositol glycan anchor biosynthesis, class Y [	-0.719569794	7.68E-11	1.15E-09	yes
PKDCC	91461	protein kinase domain containing, cytoplasmic [Source:H	-0.832289195	0.002472854	0.009000332	yes
RFC2	5982	replication factor C (activator 1) 2, 40kDa [Source:HGNC	-0.721630136	3.29E-14	7.48E-13	yes

RPL27A	6157	ribosomal protein L27a [Source:HGNC Symbol;Acc:HGNC	-0.670435785	8.37E-22	3.64E-20	yes
SERPINB6	5269	serpin peptidase inhibitor, clade B (ovalbumin), member	-0.663143102	1.33E-13	2.82E-12	yes
SETMAR	6419	SET domain and mariner transposase fusion gene [Source	-0.829128454	8.20E-06	5.63E-05	yes
SNRPB	6628	small nuclear ribonucleoprotein polypeptides B and B1 [S	-0.652632268	8.88E-15	2.13E-13	yes
SPINK1	6690	serine peptidase inhibitor, Kazal type 1 [Source:HGNC Sy	-0.764530062	5.52E-16	1.46E-14	yes
TMEM70	54968	transmembrane protein 70 [Source:HGNC Symbol;Acc:H	-0.671226218	5.44E-06	3.89E-05	yes
TP53TG1	11257	NA	-1.004931743	2.47E-15	6.21E-14	yes
TPRG1L	127262	tumor protein p63 regulated 1-like [Source:HGNC Symbo	-0.867837728	3.12E-19	1.11E-17	yes
VRK1	7443	vaccinia related kinase 1 [Source:HGNC Symbol;Acc:HGN	-0.611836795	0.0026273	0.009509802	yes
ZNF32	7580	zinc finger protein 32 [Source:HGNC Symbol;Acc:HGNC:1	-0.966237289	9.79E-07	7.95E-06	yes

Table S6. Information of oligos or shRNA used in the study, Related to Figure 1-7 and STAR Methods.

ChIP-qPCR Primer	Sequence	Reference
control region (Irrelevant)-forward	AATCTAGCTGATATAGTGTGGCTC	Guo et al (2009)
control region (Irrelevant)-reverse	AAGCATACACTTACACGGCACTCC	
FKBP5-enhancer (e)-forward	GCTGCTGCAGGATAAAGGTC	
FKBP5-enhancer (e)-reverse	ATGGAGTGCAGGTTGGAGAG	
FZD6-tss-forward	CGCGGGTAATTCGTAATCC	
FZD6-tss-reverse	AGGAAAGTGCGCAAATCTGT	
hGAPDH-tss-f	TCTCCCCACACACATGCACTT	
hGAPDH-tss-r	CCTAGTCCCAGGGCTTTGATT	
KLK2 enhancer ARE(e)-forward	GTTGAAAGCAGACCTACTCTGGA	Guo et al (2009)
KLK2 enhancer ARE(e)-reverse	CTGGACCATCTTTTCAAGCAT	
KLK2 promoter ARE-f	GAGAATGCCTCCAGACTGAT	Guo et al (2009)
KLK2 promoter ARE-r	CTTGCCCTGTTGGCACCTA	
KLK3 enhancer (e) -forward	TGAGAAACCTGAGATTAGGA	
KLK3 enhancer (e) -reverse	ATCTCTCTCAGATCCAGGCT	
LBX2-f	CAGTGCTAAGCTGCACAGGA	
LBX2-r	AGGTGGTCACTTGGTTCCAG	
NANOG promoter/tss-f	GAGGGGTGGGTCTAAGGTGA	
NANOG promoter/tss-r	ATGAGGCAACCAGCTCAGTC	
PMEPA1 enhancer (e)-forward	AAACAGAAGGTGGGAGACAAA	
PMEPA1 enhancer (e)-reverse	TACCCTGGCTAAAGCAGTTTC	
SKP2-promoter(tss)-f	AATTCTTTTAGGCCGCGAAT	
SKP2-promoter(tss)-f	GCACTAAGGCATCCATCCAT	
ZFY-tss-forward	AAGGCAGTTGATGGACCTGG	
ZFY-tss-reverse	TCCCCTTCGCGGACATTTAC	
ZNF32-tss-forward	GTCCTGAACCTCGGCATC	
ZNF32-tss-reverse	CGCCAAGGAAATGAGCAG	

RT-qPCR Primer	Sequence	Reference
AR-V7-RT-forward	CTACTCCGGACCTTACGGGGACATGCG	Guo et al (2009)
AR-V7-RT-reverse	TGCCAACCCGGAATTTTTCTCCC	
AR(3'UTR)-RT-f	CCATGGCACCTTCAGACTTT	
AR(3'UTR)-RT-r	ACTGGGCCATATGAGGATCA	
FKBP5-RT-f	CAGATCTCCATGTGCCAGAA	
FKBP5-RT-r	CTTGCCCATTGCTTTATTGG	
FOXA1-RT-f	GAAGATGGAAGGGCATGAAA	
FOXA1-RT-r	GCCTGAGTTCATGTTGCTGA	
FZD6-RT-f	TTGTTGGCATCTCTGCTGTC	
FZD6-RT-r	TCTTCGACTTTCACTGATTGGA	
GAPDH-RT-f	GAAGGTGAAGGTCGGAGTCA	
GAPDH-RT-r	AATGAAGGGGTCATTGATGG	
HDAC3-RT-f	CTGGTCCTGCATTACGGTCT	
HDAC3-RT-r	TATTGGTGGGGCTGACTCTC	
IDH2-RT-f	TGGTGATGAGATGACCCGTA	

IDH2-RT-r	GGAGCCCGAGGTCAAAATAC
KLK2-RT-f	GGCTCTGGACAGGTGGTAAA
KLK2-RT-r	GCCCCATGATGTGATACCTT
KLK3/PSA-RT-f	ACTTCAGTGTGTGGACCTCCATGT
KLK3/PSA-RT-r	AGCACACAGCATGAACTTGGTCAC
PHF21B-RT-f	AAAGAAAGACGAGGGTGTGC
PHF21B-RT-r	GCAGCTTCTGCTTCTCCTCT
PMEPA1-RT-f	TGCAACTGCAAACGCTCTTT
PMEPA1-RT-r	CCACCACCATCACCATCATC
SKP2-RT-f	TCCCTGAGCTGCTAAAGGTC
SKP2-RT-r	CAACCGACCAGTCACATCC
ZFX-RT-f	CGTAGGAGAGGAGGATGCTG
ZFX-RT-r	TTCCGGTTTTCAATTCCATC
ZNF32-RT-f	TGGATTTCCAACAGCTACCC
ZNF32-RT-r	ATTTGTGGTGGGCTTCAGTC

shRNA in pLKO.1-puro	**Target sequence**	**Clone ID (Sigma)**
sh-FZD6#1	CCACCCATTGATTGTATTATA	TRCN0000008338
sh-FZD6#2	CCCATGTCCTTATCAGGCAAA	TRCN0000008339
sh-ZFX#1	CCAATCAGTCTCATTCACATA	TRCN0000017308
sh-ZFX#2 **	GTCGGAAATTGATCCTTGTAA	TRCN0000017309
sh-ZFX#3	AGTGATTTGAAACGACACATA	TRCN0000017310
sh-ZNF32#1	CCTGAAGAGAAGTTCTCTTTA	TRCN0000014923
sh-ZNF32#2	GAAACCCTACGAGTGTGCTAT	TRCN0000014926

**, sh2 of ZFX used for functional work