Cancer Cell, Volume *35*

Supplemental Information

ARv7 Represses Tumor-Suppressor Genes in Castration-Resistant Prostate Cancer

Laura Cato, Jonas de Tribolet-Hardy, Irene Lee, Jaice T. Rottenberg, Ilsa Coleman, Diana Melchers, René Houtman, Tengfei Xiao, Wei Li, Takuma Uo, Shihua Sun, Nane C. Kuznik, Bettina Göppert, Fatma Ozgun, Martin E. van Royen, Adriaan B. Houtsmuller, Raga Vadhi, Prakash K. Rao, Lewyn Li, Steven P. Balk, Robert B. Den, Bruce J. Trock, R. Jeffrey Karnes, Robert B. Jenkins, Eric A. Klein, Elai Davicioni, Friederike J. Gruhl, Henry W. Long, X. Shirley Liu, Andrew C.B. Cato, Nathan A. Lack, Peter S. Nelson, Stephen R. Plymate, Anna C. Groner, and Myles Brown

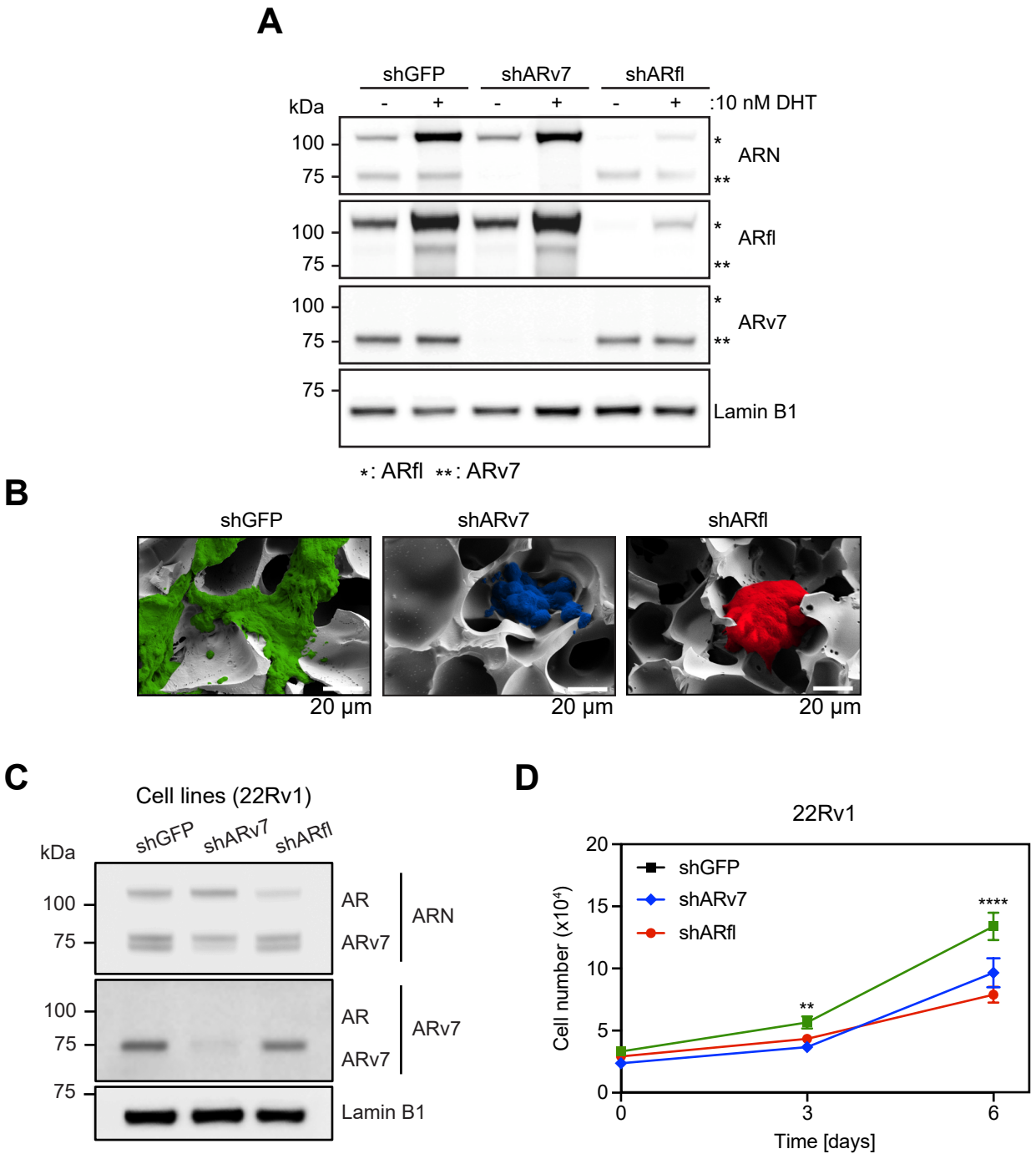

Figure S1 (related to Figure 1). Efficiency of ARv7 and ARfl knockdown in LNCaP95 and 22Rv1 cells and its effect on 22Rv1 cell growth

A. Western blot of nuclear lysates of LNCaP95 shGFP, shARv7 and shARfl cells induced with dox for 3 days and treatment with vehicle (-) or 10 nM DHT (+) for 4 hr. AR levels were detected (and specificity of the respective antibodies confirmed) using antibodies against the AR N-terminus (ARN; which recognized both ARfl and ARv7) or antibodies specific for ARfl or ARv7. Equal protein loading was confirmed with Lamin B1. **B.** Representative scanning electron microscope images of LNCaP95 cell lines grown in 3D/PEGda cryogels and in the presence of dox (as in Figure 1 C). Images were acquired after 7 days of growth. Scale bar = 20 µm. **C.** Western blot of nuclear lysates from 22Rv1 cells with dox-inducible sh constructs against GFP (shGFP), ARv7 (shARv7) or ARfl (shARfl). Cells were hormone starved and treated with dox and vehicle (ETOH) for 6-days. Proteins were detected as described in A, using the ARN and ARv7 antibodies. **D.** Cell growth assay of 22Rv1 cells, as in C, grown over a 6 day period in the presence of dox and vehicle (ETOH). Data is the mean of 3 independent experiments ± SD. p values were calculated using standard t test; **: $p \leq 0.01$; ****: $p \leq 0.0001$.

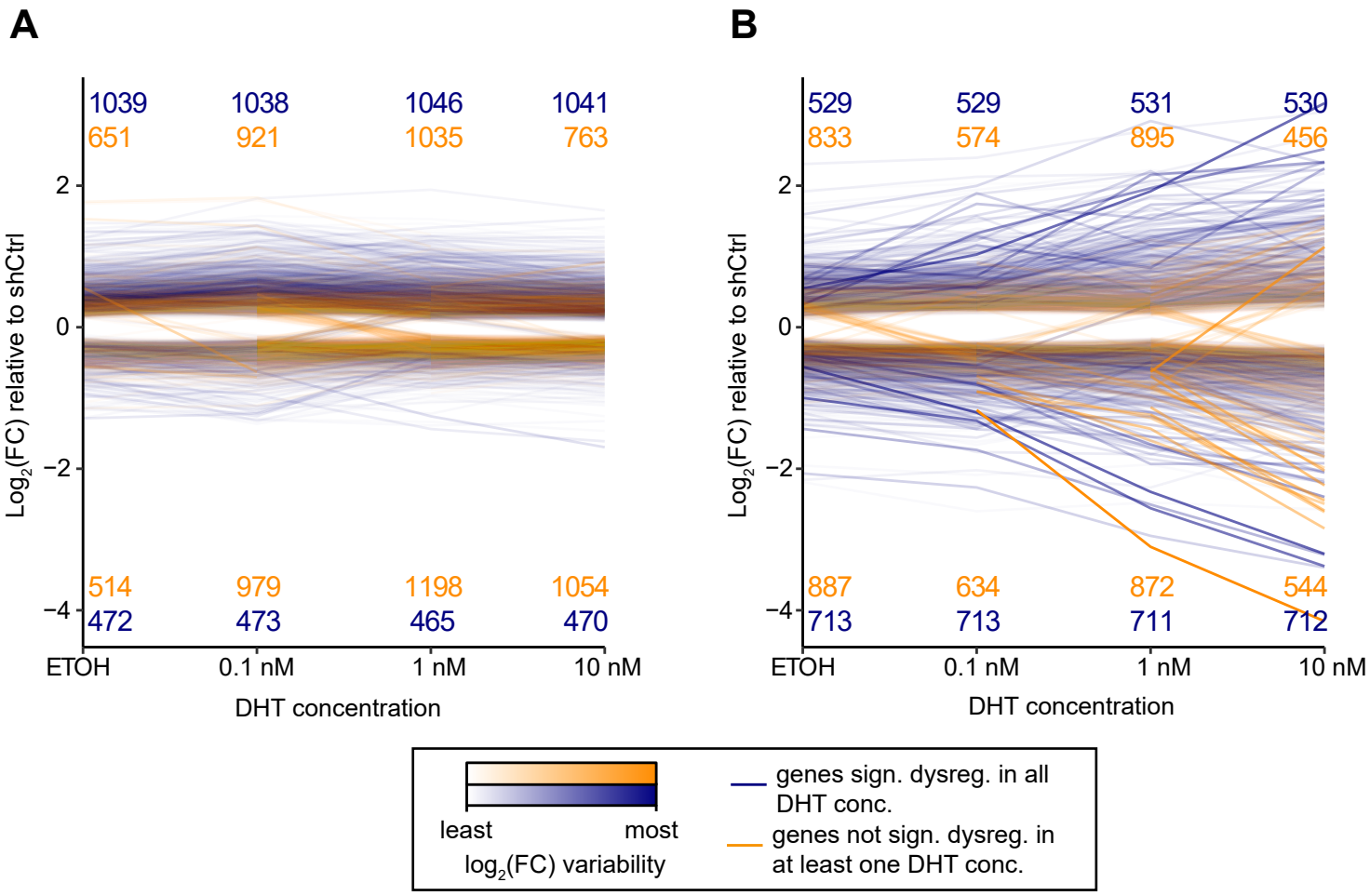

Figure S2 (related to Figure 2). DHT dose-dependent responses in shARv7 or shARfl LNCaP95 cells

A, B. Lineplots of differentially-expressed genes in the shARv7 (A) or shARfl (B) LNCaP95 cell lines, compared to the shGFP control. Cells were grown in the presence of dox for 3 days, with different doses of DHT (0.1, 1, and 10 nM) or with vehicle (ETOH) for 24 hr. Genes significantly dysregulated relative to the control (shGFP) for all DHT concentrations are drawn in blue. Genes with missing values are drawn in orange. The opacity of the lines is proportional to their variance over the four concentrations: i.e. higher standard deviation equals higher opacity. The numbers indicate the number of points with a positive (top) or negative log2 fold change (log2FC) (bottom) for a given DHT concentration.

Table S1 (related to Figure 2). Hallmark Gene Set Enrichment Analysis of ARfl- and ARv7-regulated genes

NAME	NES	NOM p val	CELL LINE
E2F TARGETS	1.875	0.0000	LN95 shARv7 ETOH
G2M CHECKPOINT	1.862	0.0000	LN95 shARv7 ETOH
ESTROGEN RESPONSE EARLY	1.752	0.0000	LN95 shARv7 ETOH
ESTROGEN RESPONSE LATE	1.720	0.0000	LN95 shARv7 ETOH
IL2 STAT5 SIGNALING	1.695	0.0000	LN95 shARv7 ETOH
MITOTIC SPINDLE	1.605	0.0000	LN95 shARv7 ETOH
EMT	1.488	0.0000	LN95 shARv7 ETOH
ADIPOGENESIS	1.519	0.0015	LN95 shARv7 ETOH
NOTCH SIGNALING	1.651	0.0018	LN95 shARv7 ETOH
UV RESPONSE UP	1.446	0.0030	LN95 shARv7 ETOH
UV RESPONSE DN	1.498	0.0032	LN95 shARv7 ETOH
INFLAMMATORY RESPONSE	1.507	0.0046	LN95 shARv7 ETOH
APICAL JUNCTION	1.416	0.0046	LN95 shARv7 ETOH
APICAL SURFACE	1.757	0.0054	LN95 shARv7 ETOH
IL6 JAK STAT3 SIGNALING	1.656	0.0086	LN95 shARv7 ETOH
PEROXISOME	1.453	0.0100	LN95 shARv7 ETOH
OXIDATIVE PHOSPHORYLATION	1.359	0.0119	LN95 shARv7 ETOH
ANGIOGENESIS	1.570	0.0124	LN95 shARv7 ETOH
FATTY ACID METABOLISM	1.370	0.0216	LN95 shARv7 ETOH
HEDGEHOG SIGNALING	1.536	0.0295	LN95 shARv7 ETOH
BILE ACID METABOLISM	1.373	0.0356	LN95 shARv7 ETOH
WNT BETA CATENIN SIGNALING	1.443	0.0359	LN95 shARv7 ETOH
CHOLESTEROL HOMEOSTASIS	1.385	0.0403	LN95 shARv7 ETOH
MTORC1 SIGNALING	1.286	0.0425	LN95 shARv7 ETOH
DNA REPAIR	-1.336	0.0258	LN95 shARv7 ETOH
MYC TARGETS V1	2.112	0.0000	LN95 shARfl ETOH
E2F TARGETS	1.970	0.0000	LN95 shARfl ETOH
G2M CHECKPOINT	1.941	0.0000	LN95 shARfl ETOH
TGF BETA SIGNALING	1.677	0.0000	LN95 shARfl ETOH
OXIDATIVE PHOSPHORYLATION	1.661	0.0000	LN95 shARfl ETOH
MYC TARGETS V2	1.777	0.0023	LN95 shARfl ETOH
MITOTIC SPINDLE	1.278	0.0305	LN95 shARfl ETOH
FATTY ACID METABOLISM	1.325	0.0355	LN95 shARfl ETOH
ANDROGEN RESPONSE	-1.934	0.0000	LN95 shARfl ETOH
UNFOLDED PROTEIN RESPONSE	-1.666	0.0000	LN95 shARfl ETOH
INFLAMMATORY RESPONSE	-1.594	0.0000	LN95 shARfl ETOH
HEME METABOLISM	-1.478	0.0016	LN95 shARfl ETOH
EMT	-1.565	0.0017	LN95 shARfl ETOH
CHOLESTEROL HOMEOSTASIS	-1.677	0.0017	LN95 shARfl ETOH
TNFA SIGNALING VIA NFKB	-1.416	0.0051	LN95 shARfl ETOH
PROTEIN SECRETION	-1.563	0.0053	LN95 shARfl ETOH
MTORC1 SIGNALING	-1.377	0.0081	LN95 shARfl ETOH
COAGULATION	-1.497	0.0085	LN95 shARfl ETOH

ESTROGEN RESPONSE EARLY	-1.327	0.0237	LN95 shARfl ETOH
KRAS SIGNALING UP	-1.332	0.0244	LN95 shARfl ETOH
BILE ACID METABOLISM	-1.391	0.0389	LN95 shARfl ETOH

NES: (Normalized Enrichment Scores) the degree to which a given gene set is overrepresented compared to all analyzed gene sets. NOM p value: (Nominal p values) the statistical significance of the enrichment score. Pathways highlighted in blue are shared between the shARv7 and shARfl cell lines.

Table S2 (related to Figure 2C). Detailed statistics of the DHT dose-dependent transcriptional responses in shARv7 or shARfl LNCaP95 cells

Set A			Set B			tHSD		Data subsets
Transf.	DHT conc.	mean log2FC	Transf.	DHT conc.	mean log2FC	p adj	sign.	
shARv7	ETOH	0.101015802	shARv7	0.1 nM	0.065699762	**0.058050204**	**n.s.**	All
shARv7	ETOH	0.101015802	shARv7	1 nM	0.059176828	**0.008472829**	**	All
shARv7	ETOH	0.101015802	shARv7	10 nM	0.052469159	**0.001722733**	**	All
shARv7	ETOH	0.101015802	shARfl	ETOH	-0.024528024	**6.57E-14**	****	All
shARv7	ETOH	0.101015802	shARfl	0.1 nM	-0.050549201	**0**	****	All
shARv7	ETOH	0.101015802	shARfl	1 nM	-0.018436402	**9.19E-14**	****	All
shARv7	ETOH	0.101015802	shARfl	10 nM	-0.06131223	**0**	****	All
shARv7	0.1 nM	0.065699762	shARv7	1 nM	0.059176828	**0.998646135**	**n.s.**	All
shARv7	0.1 nM	0.065699762	shARv7	10 nM	0.052469159	**0.933327401**	**n.s.**	All
shARv7	0.1 nM	0.065699762	shARfl	ETOH	-0.024528024	**9.46E-14**	****	All
shARv7	0.1 nM	0.065699762	shARfl	0.1 nM	-0.050549201	**5.96E-14**	****	All
shARv7	0.1 nM	0.065699762	shARfl	1 nM	-0.018436402	**1.43E-12**	****	All
shARv7	0.1 nM	0.065699762	shARfl	10 nM	-0.06131223	**8.20E-14**	****	All
shARv7	1 nM	0.059176828	shARv7	10 nM	0.052469159	**0.99869698**	**n.s.**	All
shARv7	1 nM	0.059176828	shARfl	ETOH	-0.024528024	**2.77E-13**	****	All
shARv7	1 nM	0.059176828	shARfl	0.1 nM	-0.050549201	**9.47E-14**	****	All
shARv7	1 nM	0.059176828	shARfl	1 nM	-0.018436402	**5.16E-11**	****	All
shARv7	1 nM	0.059176828	shARfl	10 nM	-0.06131223	**6.21E-14**	****	All
shARv7	10 nM	0.052469159	shARfl	ETOH	-0.024528024	**2.52E-10**	****	All
shARv7	10 nM	0.052469159	shARfl	0.1 nM	-0.050549201	**9.00E-14**	****	All
shARv7	10 nM	0.052469159	shARfl	1 nM	-0.018436402	**2.03E-08**	****	All
shARv7	10 nM	0.052469159	shARfl	10 nM	-0.06131223	**1.01E-13**	****	All
shARfl	ETOH	-0.024528024	shARfl	0.1 nM	-0.050549201	**0.405401058**	**n.s.**	All
shARfl	ETOH	-0.024528024	shARfl	1 nM	-0.018436402	**0.999476191**	**n.s.**	All
shARfl	ETOH	-0.024528024	shARfl	10 nM	-0.06131223	**0.059440793**	**n.s.**	All
shARfl	0.1 nM	-0.050549201	shARfl	1 nM	-0.018436402	**0.168211728**	**n.s.**	All
shARfl	0.1 nM	-0.050549201	shARfl	10 nM	-0.06131223	**0.993058836**	**n.s.**	All
shARfl	1 nM	-0.018436402	shARfl	10 nM	-0.06131223	**0.015163179**	*	All

shARv7	ETOH	0.412119564	shARv7	0.1 nM	0.450191355	**2.05737E-05**	****	Up
shARv7	ETOH	0.412119564	shARv7	1 nM	0.409918282	**0.999991801**	**n.s.**	Up
shARv7	ETOH	0.412119564	shARv7	10 nM	0.393006619	**0.237353559**	**n.s.**	Up
shARv7	ETOH	0.412119564	shARfl	ETOH	0.363676569	**5.71E-08**	****	Up
shARv7	ETOH	0.412119564	shARfl	0.1 nM	0.490935854	**8.82608E-09**	****	Up
shARv7	ETOH	0.412119564	shARfl	1 nM	0.459979097	**1.95E-07**	****	Up
shARv7	ETOH	0.412119564	shARfl	10 nM	0.645198704	**8.82601E-09**	****	Up
shARv7	0.1 nM	0.450191355	shARv7	1 nM	0.409918282	**4.57594E-07**	****	Up
shARv7	0.1 nM	0.450191355	shARv7	10 nM	0.393006619	**8.82664E-09**	****	Up
shARv7	0.1 nM	0.450191355	shARfl	ETOH	0.363676569	**8.83E-09**	****	Up
shARv7	0.1 nM	0.450191355	shARfl	0.1 nM	0.490935854	**5.96E-05**	****	Up
shARv7	0.1 nM	0.450191355	shARfl	1 nM	0.459979097	**9.15E-01**	**n.s.**	Up
shARv7	0.1 nM	0.450191355	shARfl	10 nM	0.645198704	**8.83E-09**	****	Up
shARv7	1 nM	0.409918282	shARv7	10 nM	0.393006619	**0.300080308**	**n.s.**	Up
shARv7	1 nM	0.409918282	shARfl	ETOH	0.363676569	**2.90E-08**	****	Up
shARv7	1 nM	0.409918282	shARfl	0.1 nM	0.490935854	**8.83E-09**	****	Up
shARv7	1 nM	0.409918282	shARfl	1 nM	0.459979097	**1.15E-08**	****	Up
shARv7	1 nM	0.409918282	shARfl	10 nM	0.645198704	**8.83E-09**	****	Up
shARv7	10 nM	0.393006619	shARfl	ETOH	0.363676569	**4.48E-03**	**	Up
shARv7	10 nM	0.393006619	shARfl	0.1 nM	0.490935854	**8.83E-09**	****	Up
shARv7	10 nM	0.393006619	shARfl	1 nM	0.459979097	**8.83E-09**	****	Up
shARv7	10 nM	0.393006619	shARfl	10 nM	0.645198704	**8.83E-09**	****	Up
shARfl	ETOH	0.363676569	shARfl	0.1 nM	0.490935854	**8.82601E-09**	****	Up
shARfl	ETOH	0.363676569	shARfl	1 nM	0.459979097	**8.82601E-09**	****	Up
shARfl	ETOH	0.363676569	shARfl	10 nM	0.645198704	**8.82601E-09**	****	Up
shARfl	0.1 nM	0.490935854	shARfl	1 nM	0.459979097	**0.015520346**	*	Up
shARfl	0.1 nM	0.490935854	shARfl	10 nM	0.645198704	**8.82601E-09**	****	Up
shARfl	1 nM	0.459979097	shARfl	10 nM	0.645198704	**8.82601E-09**	****	Up
shARv7	ETOH	-0.391452534	shARv7	0.1 nM	-0.382616663	**0.972680671**	**n.s.**	Down
shARv7	ETOH	-0.391452534	shARv7	1 nM	-0.338362766	**2.36307E-08**	****	Down
shARv7	ETOH	-0.391452534	shARv7	10 nM	-0.335184282	**1.1808E-08**	****	Down
shARv7	ETOH	-0.391452534	shARfl	ETOH	-0.387748216	**1.00E+00**	**n.s.**	Down
shARv7	ETOH	-0.391452534	shARfl	0.1 nM	-0.499942533	**3.82667E-09**	****	Down
shARv7	ETOH	-0.391452534	shARfl	1 nM	-0.445998655	**1.82E-08**	****	Down
shARv7	ETOH	-0.391452534	shARfl	10 nM	-0.648881929	**3.82667E-09**	****	Down
shARv7	0.1 nM	-0.382616663	shARv7	1 nM	-0.338362766	**4.78492E-08**	****	Down
shARv7	0.1 nM	-0.382616663	shARv7	10 nM	-0.335184282	**2.33885E-08**	****	Down
shARv7	0.1 nM	-0.382616663	shARfl	ETOH	-0.387748216	**9.97E-01**	**n.s.**	Down
shARv7	0.1 nM	-0.382616663	shARfl	0.1 nM	-0.499942533	**3.83E-09**	****	Down
shARv7	0.1 nM	-0.382616663	shARfl	1 nM	-0.445998655	**3.83E-09**	****	Down
shARv7	0.1 nM	-0.382616663	shARfl	10 nM	-0.648881929	**3.83E-09**	****	Down
shARv7	1 nM	-0.338362766	shARv7	10 nM	-0.335184282	**0.999894208**	**n.s.**	Down
shARv7	1 nM	-0.338362766	shARfl	ETOH	-0.387748216	**4.15E-09**	****	Down
shARv7	1 nM	-0.338362766	shARfl	0.1 nM	-0.499942533	**3.83E-09**	****	Down
shARv7	1 nM	-0.338362766	shARfl	1 nM	-0.445998655	**3.83E-09**	****	Down
shARv7	1 nM	-0.338362766	shARfl	10 nM	-0.648881929	**3.83E-09**	****	Down
shARv7	10 nM	-0.335184282	shARfl	ETOH	-0.387748216	**4.00E-09**	****	Down
shARv7	10 nM	-0.335184282	shARfl	0.1 nM	-0.499942533	**3.83E-09**	****	Down

shARv7	10 nM	-0.335184282	shARfl	1 nM	-0.445998655	**3.83E-09**	****	Down
shARv7	10 nM	-0.335184282	shARfl	10 nM	-0.648881929	**3.83E-09**	****	Down
shARfl	ETOH	-0.387748216	shARfl	0.1 nM	-0.499942533	**3.82667E-09**	****	Down
shARfl	ETOH	-0.387748216	shARfl	1 nM	-0.445998655	**3.82692E-09**	****	Down
shARfl	ETOH	-0.387748216	shARfl	10 nM	-0.648881929	**3.82667E-09**	****	Down
shARfl	0.1 nM	-0.499942533	shARfl	1 nM	-0.445998655	**4.53458E-09**	****	Down
shARfl	0.1 nM	-0.499942533	shARfl	10 nM	-0.648881929	**3.82667E-09**	****	Down
shARfl	1 nM	-0.445998655	shARfl	10 nM	-0.648881929	**3.82667E-09**	****	Down

Set A and Set B: Values used for ANOVA. tHSD: Tukey's honest significance test. Transf.: Transfection of the cell line. DHT conc.: DHT concentration in which the cells were grown. Mean log2FC: Averages of logarithm to the base 2 of the fold changes for the cell lines relative to control (shGFP). P adj: adjusted p values for the tHSD. Sign: Significance of the adjusted p values. Data subsets: Indication which data points were used in the analyses. All: Using all differentially expressed genes (adj $p < 0.05$). Up: Using up regulated differentially expressed genes ($log2FC > 0$, adj $p < 0.05$). Down: Using down regulated differentially expressed genes ($log2FC < 0$, adj $p < 0.05$).

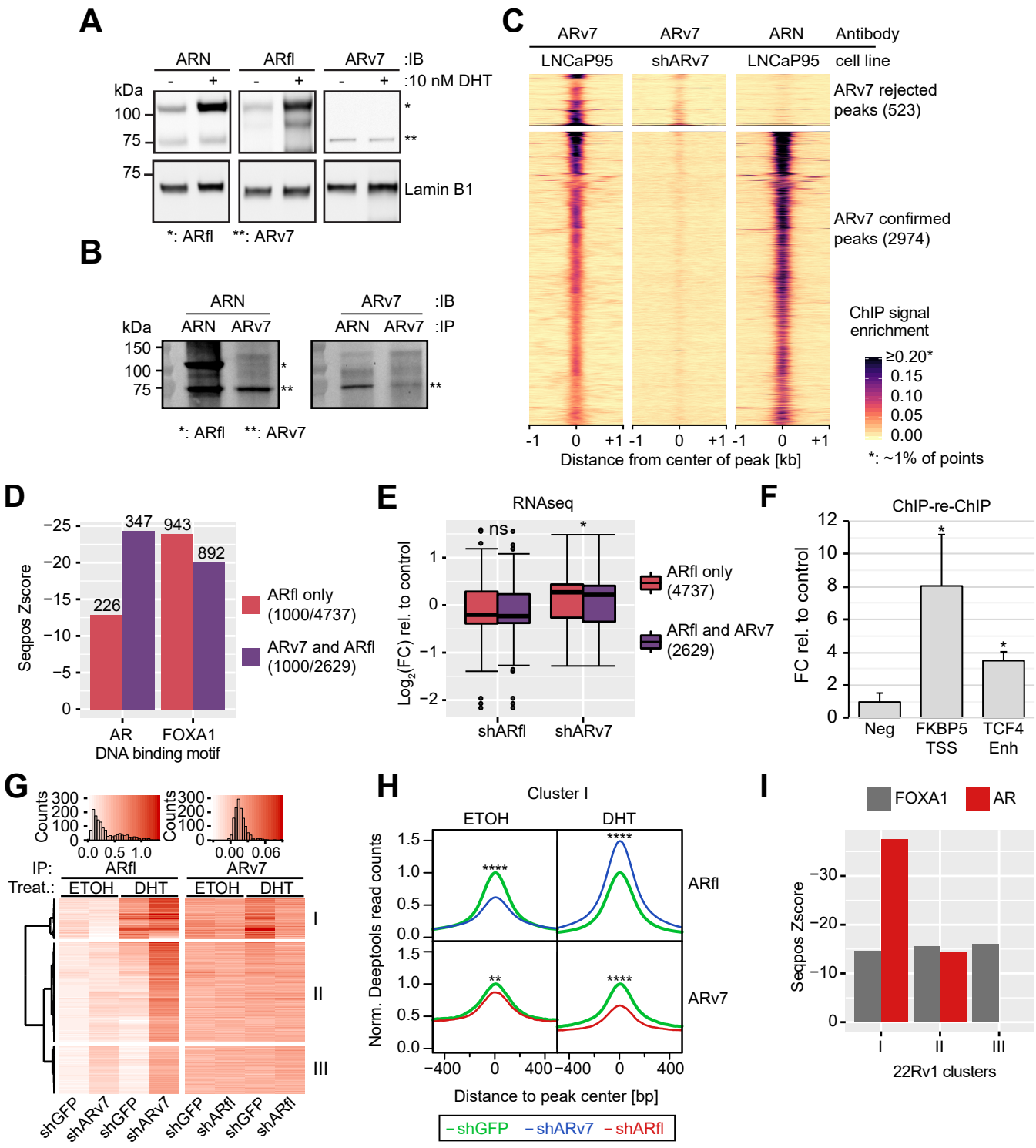
A
ARN
ARfl
ARv7
:IB
:10 nM DHT
kDa
100
75
75
Lamin B1
*: ARfl
**: ARv7
B
ARN
ARv7
:IB
:IP
kDa
150
100
75
*: ARfl
**: ARv7
C
ARv7
ARv7
ARN
Antibody
LNCaP95
shARv7
LNCaP95
cell line
ARv7 rejected peaks (523)
ARv7 confirmed peaks (2974)
ChIP signal enrichment
≥0.20*
0.15
0.10
0.05
0.00
*: ~1% of points
-1
0
+1
Distance from center of peak [kb]
D
Seqpos Zscore
347
943
892
226
ARfl only (1000/4737)
ARv7 and ARfl (1000/2629)
AR
FOXA1
DNA binding motif
E
RNAseq
ns
*
Log2(FC) rel. to control
ARfl only (4737)
ARfl and ARv7 (2629)
shARfl
shARv7
F
ChIP-re-ChIP
FC rel. to control
Neg
FKBP5 TSS
TCF4 Enh
G
Counts
IP:
ARfl
ARv7
Treat.:
ETOH
DHT
I
II
III
shGFP
shARv7
shGFP
shARv7
shGFP
shARfl
shGFP
shARfl
H
Cluster I
ETOH
DHT
ARfl
ARv7
Norm. Deeptools read counts
Distance to peak center [bp]
shGFP
shARv7
shARfl
I
FOXA1
AR
Seqpos Zscore
22Rv1 clusters

Figure S3 (related to Figure 3). ARv7 and ARfl ChIP-seq in LNCaP95 and 22Rv1 and AR isoform-specific fluorescence recovery after photobleaching (FRAP) microscopy

A. Western blot of nuclear lysates of LNCaP95 cells treated with vehicle (-) or 10 nM DHT (+) for 4 h. AR levels were detected (and specificity of the antibody confirmed) using ARN, ARfl or ARv7 antibodies. Equal protein loading was confirmed by probing with Lamin B1. **B.** Western blot of AR N-terminus (ARN) or ARv7 co-IPs, using formaldehyde cross-linked and sonicated lysates from LNCaP95 cells, in the absence of hormone. Blots were probed (IB) with either ARN or ARv7 antibodies, as indicated. **C.** Heatmap of ARv7 or ARN ChIP-seq signals centered at ARv7 peaks in LNCaP95 parental or shARv7 cell lines, as indicated. Each ChIP-seq dataset was normalized separately. ARv7 peak signals are hierarchically clustered by Euclidian distance and ward functions in R. 523 ARv7 peaks clustered separately. These peaks were defined as "ARv7 rejected peaks" and were removed from further downstream analyses, because they did not show signals in an additional N-terminal AR-specific cistrome, and did not show a reduction of signal with shARv7. **D.** Comparison of DNA binding motifs of ARfl-only (n=4737) and ARv7-ARfl-shared ChIP-seq peaks (n=2629) peaks, as defined in Figure 3B. Peaks were randomly subsampled for this analysis to 1000 each. The height of the bars represents the Zscore from SeqPos which is proportional to the p value of the motif, as well as its position within the peaks. The numbers indicate the number of times the motif was found within the 1000 peaks; multiple hits per peak are possible. **E.** Expression of significantly dysregulated genes (adj. $p < 0.05$) associated with ChIP-seq peaks as defined in D (50 kb interval around the transcription start site). The boxplots show mRNA fold changes (FC) after knock-down of either ARfl (shARfl) or ARv7 (shARv7) relative to control (shGFP). P values between groups were calculated using the Kruskal-Wallis H test. Boxplots show the median, the 1st and 3rd quartile. Whiskers extend to 1.5 the interquartile range and data beyond that is shown as individual points. **F.** ARfl-ARv7 ChIP-re-ChIP for select binding sites (as indicated) in LNCaP95 cells in the absence of DHT. Cells were first ChIPed with an ARfl-specific antibody (AR C19) and then re-ChIPed with either IgG (control) or an ARv7 antibody. Results are normalized to the negative control locus (Neg) and expressed as a fold change (FC) of the ARv7 re-ChIP over the IgG re-ChIP. P values were calculated relative to control (Neg) using standard t-test; *: $p \leq 0.05$ **G.** Heatmap of ARfl and ARv7 ChIP-seq in indicated 22Rv1 cell lines treated for 4 hr with 10 nM DHT (DHT) or vehicle (ETOH). The heatmap shows the signal intensities of the 10000 most variable peaks of the ETOH-treated ChIP-seq experiments, which were then clustered on the ARfl signals using Pearson correlation. **H.** ChIP-seq signal profiles for the peaks in cluster I from G. P values were calculated using standard t-test; **: $p \leq 0.01$; ****: $p \leq 0.0001$. **I.** Comparison of DNA binding motifs of peaks defined in G. Peaks were randomly subsampled to 1000 each. The height of the bars represents the Zscore from SeqPos, which is proportional to the p value of the motif as well as its position within the peaks.

Table S3 (related to Figure 3) Gene Ontology Hallmark gene sets associated with ARfl only and ARfl/ARv7 shared peaks.

Examined gene set	Gene Set Name	# Genes in Gene Set (K)	Description	# Genes in Overlap (k)	k/K	p value	FDR q value
ARfl only (4737 peaks)	HALLMARK ANDROGEN RESPONSE	101	Genes defining response to androgens.	32	0.3168	3.39E-18	1.69E-16
ARfl only (4737 peaks)	HALLMARK XENOBIOTIC METABOLISM	200	Genes encoding proteins involved in processing of drugs and other xenobiotics.	41	0.205	2.10E-15	5.25E-14
ARfl only (4737 peaks)	HALLMARK IL2 STAT5 SIGNALING	200	Genes up-regulated by STAT5 in response to IL2 stimulation.	40	0.2	1.10E-14	1.83E-13
ARfl only (4737 peaks)	HALLMARK TNFA SIGNALING VIA NFKB	200	Genes regulated by NF-kB in response to TNF [GeneID=7124].	37	0.185	1.31E-12	1.63E-11
ARfl only (4737 peaks)	HALLMARK G2M CHECKPOINT	200	Genes involved in the G2/M checkpoint, as in progression through the cell division cycle.	35	0.175	2.68E-11	2.68E-10
ARfl only (4737 peaks)	HALLMARK FATTY ACID METABOLISM	158	Genes encoding proteins involved in metabolism of fatty acids.	30	0.1899	8.61E-11	7.18E-10
ARfl only (4737 peaks)	HALLMARK ESTROGEN RESPONSE EARLY	200	Genes defining early response to estrogen.	33	0.165	4.81E-10	2.73E-09
ARfl only (4737 peaks)	HALLMARK MTORC1 SIGNALING	200	Genes up-regulated through activation of mTORC1 complex.	33	0.165	4.81E-10	2.73E-09
ARfl only (4737 peaks)	HALLMARK BILE ACID METABOLISM	112	Genes involve in metabolism of bile acids and salts.	24	0.2143	4.91E-10	2.73E-09
ARfl only (4737 peaks)	HALLMARK UV RESPONSE DN	144	Genes down-regulated in response to ultraviolet (UV) radiation.	27	0.1875	9.91E-10	4.95E-09
ARfl only (4737 peaks)	HALLMARK COAGULATION	138	Genes encoding components of blood coagulation system; also up-regulated in platelets.	26	0.1884	1.80E-09	7.42E-09
ARfl only (4737 peaks)	HALLMARK ESTROGEN RESPONSE LATE	200	Genes defining late response to estrogen.	32	0.16	1.93E-09	7.42E-09
ARfl only (4737 peaks)	HALLMARK HYPOXIA	200	Genes up-regulated in response to low oxygen levels (hypoxia).	32	0.16	1.93E-09	7.42E-09
ARfl only (4737 peaks)	HALLMARK P53 PATHWAY	200	Genes involved in p53 pathways and networks.	31	0.155	7.47E-09	2.67E-08
ARfl only (4737 peaks)	HALLMARK PEROXISOME	104	Genes encoding components of peroxisome.	21	0.2019	1.79E-08	5.96E-08
ARfl only (4737 peaks)	HALLMARK GLYCOLYSIS	200	Genes encoding proteins involved in glycolysis and gluconeogenesis.	30	0.15	2.79E-08	8.20E-08
ARfl only (4737 peaks)	HALLMARK INFLAMMATORY RESPONSE	200	Genes defining inflammatory response.	30	0.15	2.79E-08	8.20E-08
ARfl only (4737 peaks)	HALLMARK ADIPOGENESIS	200	Genes up-regulated during adipocyte differentiation (adipogenesis).	28	0.14	3.45E-07	9.60E-07
ARfl only (4737 peaks)	HALLMARK COMPLEMENT	200	Genes encoding components of the complement system, which is part of the innate immune system.	27	0.135	1.15E-06	2.87E-06
ARfl only (4737 peaks)	HALLMARK KRAS SIGNALING UP	200	Genes up-regulated by KRAS activation.	27	0.135	1.15E-06	2.87E-06
ARfl and ARv7 (2629 peaks)	HALLMARK ANDROGEN RESPONSE	101	Genes defining response to androgens.	25	0.2475	2.11E-17	1.06E-15
ARfl and ARv7 (2629 peaks)	HALLMARK TNFA SIGNALING VIA NFKB	200	Genes regulated by NF-kB in response to TNF [GeneID=7124].	28	0.14	1.23E-12	3.07E-11
ARfl and ARv7 (2629 peaks)	HALLMARK G2M CHECKPOINT	200	Genes involved in the G2/M checkpoint, as in progression through the cell division cycle.	23	0.115	6.16E-09	1.03E-07

ARfl and ARv7 (2629 peaks)	HALLMARK MITOTIC SPINDLE	200	Genes important for mitotic spindle assembly.	22	0.11	2.95E-08	2.95E-07
ARfl and ARv7 (2629 peaks)	HALLMARK XENOBIOTIC METABOLISM	200	Genes encoding proteins involved in processing of drugs and other xenobiotics.	22	0.11	2.95E-08	2.95E-07
ARfl and ARv7 (2629 peaks)	HALLMARK ADIPOGENESIS	200	Genes up-regulated during adipocyte differentiation (adipogenesis).	21	0.105	1.34E-07	9.60E-07
ARfl and ARv7 (2629 peaks)	HALLMARK KRAS SIGNALING UP	200	Genes up-regulated by KRAS activation.	21	0.105	1.34E-07	9.60E-07
ARfl and ARv7 (2629 peaks)	HALLMARK IL2 STAT5 SIGNALING	200	Genes up-regulated by STAT5 in response to IL2 stimulation.	20	0.1	5.82E-07	3.64E-06
ARfl and ARv7 (2629 peaks)	HALLMARK HEME METABOLISM	200	Genes involved in metabolism of heme (a cofactor consisting of iron and porphyrin) and erythroblast differentiation.	19	0.095	2.39E-06	1.33E-05
ARfl and ARv7 (2629 peaks)	HALLMARK SPERMATOGENESIS	135	Genes up-regulated during production of male gametes (sperm), as in spermatogenesis.	15	0.1111	4.00E-06	2.00E-05
ARfl and ARv7 (2629 peaks)	HALLMARK INFLAMMATORY RESPONSE	200	Genes defining inflammatory response.	18	0.09	9.26E-06	4.21E-05
ARfl and ARv7 (2629 peaks)	HALLMARK CHOLESTEROL HOMEOSTASIS	74	Genes involved in cholesterol homeostasis.	10	0.1351	2.94E-05	1.21E-04
ARfl and ARv7 (2629 peaks)	HALLMARK HYPOXIA	200	Genes up-regulated in response to low oxygen levels (hypoxia).	17	0.085	3.39E-05	1.21E-04
ARfl and ARv7 (2629 peaks)	HALLMARK OXIDATIVE PHOSPHORYLATION	200	Genes encoding proteins involved in oxidative phosphorylation.	17	0.085	3.39E-05	1.21E-04
ARfl and ARv7 (2629 peaks)	HALLMARK APICAL JUNCTION	200	Genes encoding components of apical junction complex.	16	0.08	1.17E-04	3.24E-04
ARfl and ARv7 (2629 peaks)	HALLMARK E2F TARGETS	200	Genes encoding cell cycle related targets of E2F transcription factors.	16	0.08	1.17E-04	3.24E-04
ARfl and ARv7 (2629 peaks)	HALLMARK GLYCOLYSIS	200	Genes encoding proteins involved in glycolysis and gluconeogenesis.	16	0.08	1.17E-04	3.24E-04
ARfl and ARv7 (2629 peaks)	HALLMARK MTORC1 SIGNALING	200	Genes up-regulated through activation of mTORC1 complex.	16	0.08	1.17E-04	3.24E-04
ARfl and ARv7 (2629 peaks)	HALLMARK FATTY ACID METABOLISM	158	Genes encoding proteins involved in metabolism of fatty acids.	13	0.0823	3.77E-04	8.99E-04
ARfl and ARv7 (2629 peaks)	HALLMARK EPITHELIAL MESENCHYMAL TRANSITION	200	Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis.	15	0.075	3.77E-04	8.99E-04

Pathways highlighted in blue are with the top 20 enriched pathways for both the genes associated with the ARfl only and the ARfl and ARv7 peaks.

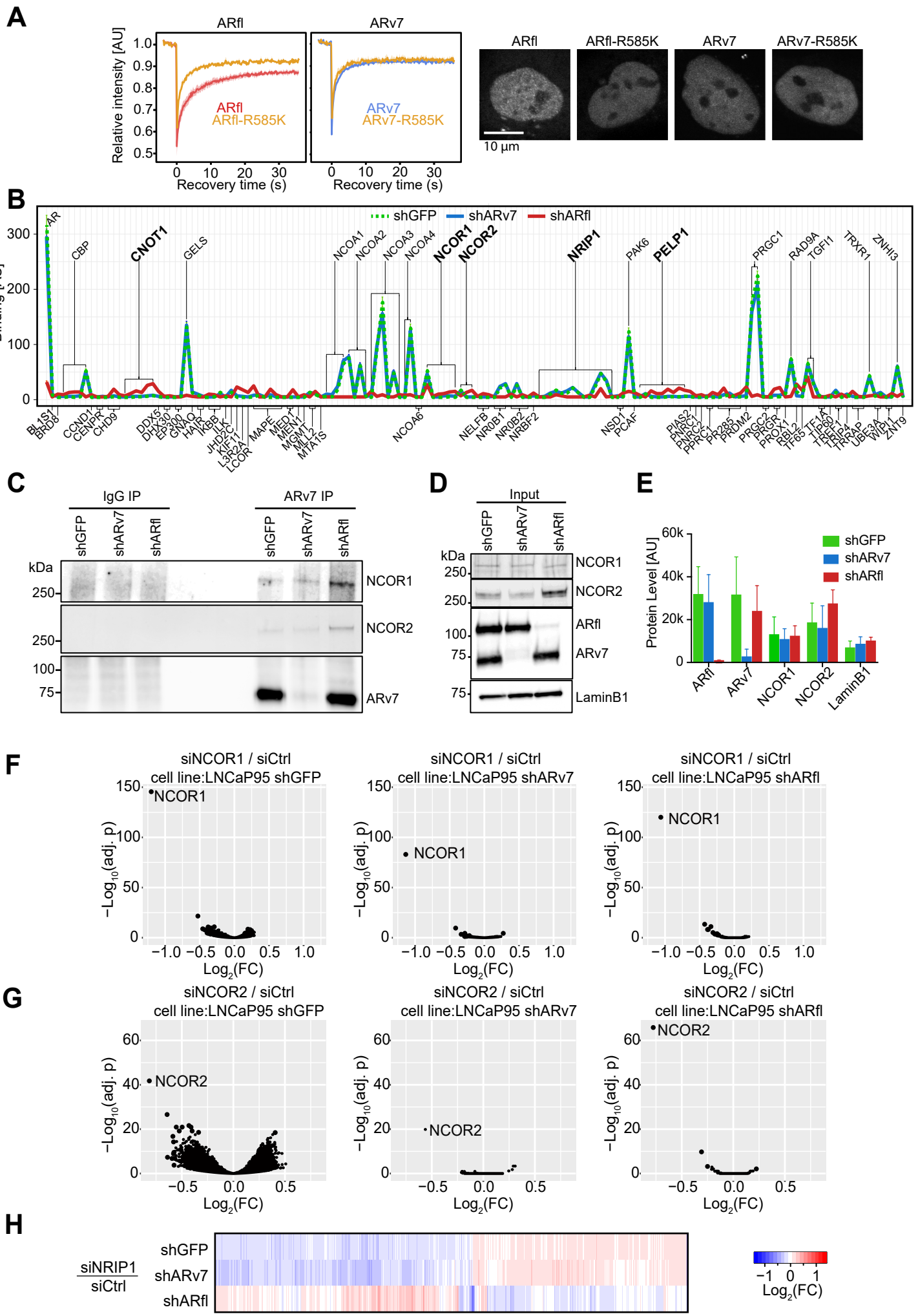
A
ARfl
ARv7
Relative intensity [AU]
ARfl
ARfl-R585K
ARv7
ARv7-R585K
Recovery time (s)
ARfl
ARfl-R585K
ARv7
ARv7-R585K
10 µm
B
shGFP
shARv7
shARfl
Binding [AU]
AR
CBP
CNOT1
GELS
NCOA1
NCOA2
NCOA3
NCOA4
NCOR1
NCOR2
NRIP1
PAK6
PELP1
PRGC1
RAD9A
TGFI1
TRXR1
ZNHI3
C
IgG IP
ARv7 IP
shGFP
shARv7
shARfl
kDa
NCOR1
NCOR2
ARv7
D
Input
shGFP
shARv7
shARfl
kDa
NCOR1
NCOR2
ARfl
ARv7
LaminB1
E
Protein Level [AU]
shGFP
shARv7
shARfl
ARfl
ARv7
NCOR1
NCOR2
LaminB1
F
siNCOR1 / siCtrl
cell line:LNCaP95 shGFP
cell line:LNCaP95 shARv7
cell line:LNCaP95 shARfl
NCOR1
−Log10(adj. p)
Log2(FC)
G
siNCOR2 / siCtrl
cell line:LNCaP95 shGFP
cell line:LNCaP95 shARv7
cell line:LNCaP95 shARfl
NCOR2
−Log10(adj. p)
Log2(FC)
H
siNRIP1
siCtrl
shGFP
shARv7
shARfl
Log2(FC)

Figure S4 (related to Figure 4). ARv7 and ARfl coregulator interaction assays

A. Line plots (*left*) and representative images (*right*) of FRAP microscopy of wild-type ARfl or ARv7 and their respective DNA-binding mutants (R585K). Results are the mean of 33-47 cells ± SEM and are shown in arbitrary units (AU). Image scale bar is 10 µm. **B.** MARCoNI peptide binding array of cell lysates from shGFP (dotted green), shARv7 (solid blue) and shARfl cells (solid red) grown in the presence of dox for 3 days and induced with 10 nM DHT for 30 min. Peptide binding for AR was assessed with a pan-AR antibody, and mean binding (± SEM), displayed as arbitrary binding units (AU) of triplicate experiments, is shown. Statistically significant interactions with classical corepressors are highlighted in bold. **C.** Representative example of the co-immunoprecipitation (co-IP) of ARv7 with selected co-repressors (NCOR1 and NCOR2) using nuclear lysates of shGFP, shARv7 and shARfl LNCaP95 cells grown in the absence of hormone (used in Figure 4B). **D.** Representative example of Input samples (1/25th of the IP samples) for co-IPs from B. before IP with IgG or ARv7. **E.** Quantification of all Input samples in arbitrary units (AU). **F, G.** Volcano plots for the experiments depicted in Figure 4C. Expression fold changes (FC) and adjusted p values (adj. p) after siRNA-mediated knock-down of NCOR1 (siNCOR1, F) and siNCOR2 (siNCOR2, G) relative to a si control (siCtrl) are shown in LNCaP95 cells with additional depletion of either ARv7 (shARv7), ARfl (shARfl) or a control (shGFP). **H.** Heatmap of genes significantly (adj.p < 0.05) dysregulated upon siRNA-mediated silencing of NRIP1 (siNRIP1) in LNCaP95 cell lines with an additional knock down of either ARv7 (shARv7), ARfl (shARfl) or a control (shGFP). The analysis was performed together with the one in Figure 4C. The heatmap shows the $\log_2$ fold changes of the dysregulated genes for the indicated siRNA-treatments relative to a si control in each row.

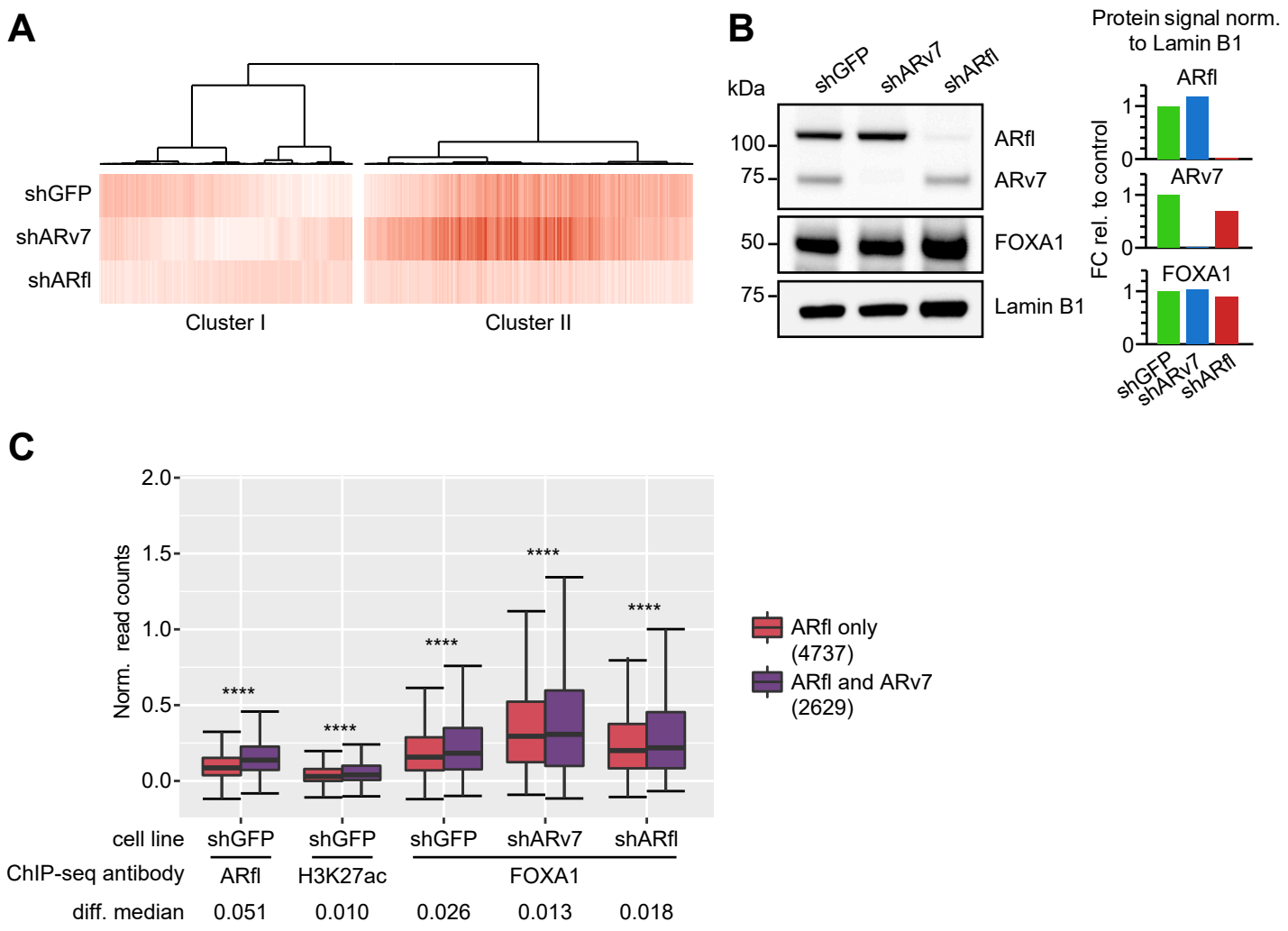

Figure S5 (related to Figure 5). Clustering of ChIP-seq peaks and FOXA1 protein levels in ARfl and ARv7 knockdown cells

A. Hierarchical clustering of the normalized read-counts from H3K27ac ChIP-seq, for the ARv7 and ARfl peaks defined in Figure 3B. The clustering is used in Figures 5A-5F. **B.** *Left*: Western blot of nuclear lysates of dox-inducible shGFP, shARv7 and shARfl cells, grown with dox for 3 days. Proteins were detected with indicated antibodies. Lamin B1 serves as a loading control. *Right*: Quantification of Western Blot signals. Signals of the ARfl-, ARv7- and FOXA1 antibodies were normalized to their Lamin B1 values and expressed as a fold change (FC) relative to shGFP **C.** ARfl, H3K27ac and FOXA1 ChIP-seq signals in indicated LNCaP95 cell lines (shGFP, shARfl and shARv7) at ARfl only (n=4737) and at ARfl/ARv7 shared peaks (n=2629), as defined in Figure 3B. The difference in the medians between ARfl/ARv7 shared and ARfl only peaks is indicated. Boxplots show the median, the 1st and 3rd quartile. Whiskers extend to 1.5 the interquartile range. P values were calculated using standard t-test; ****: $p \leq 0.0001$.

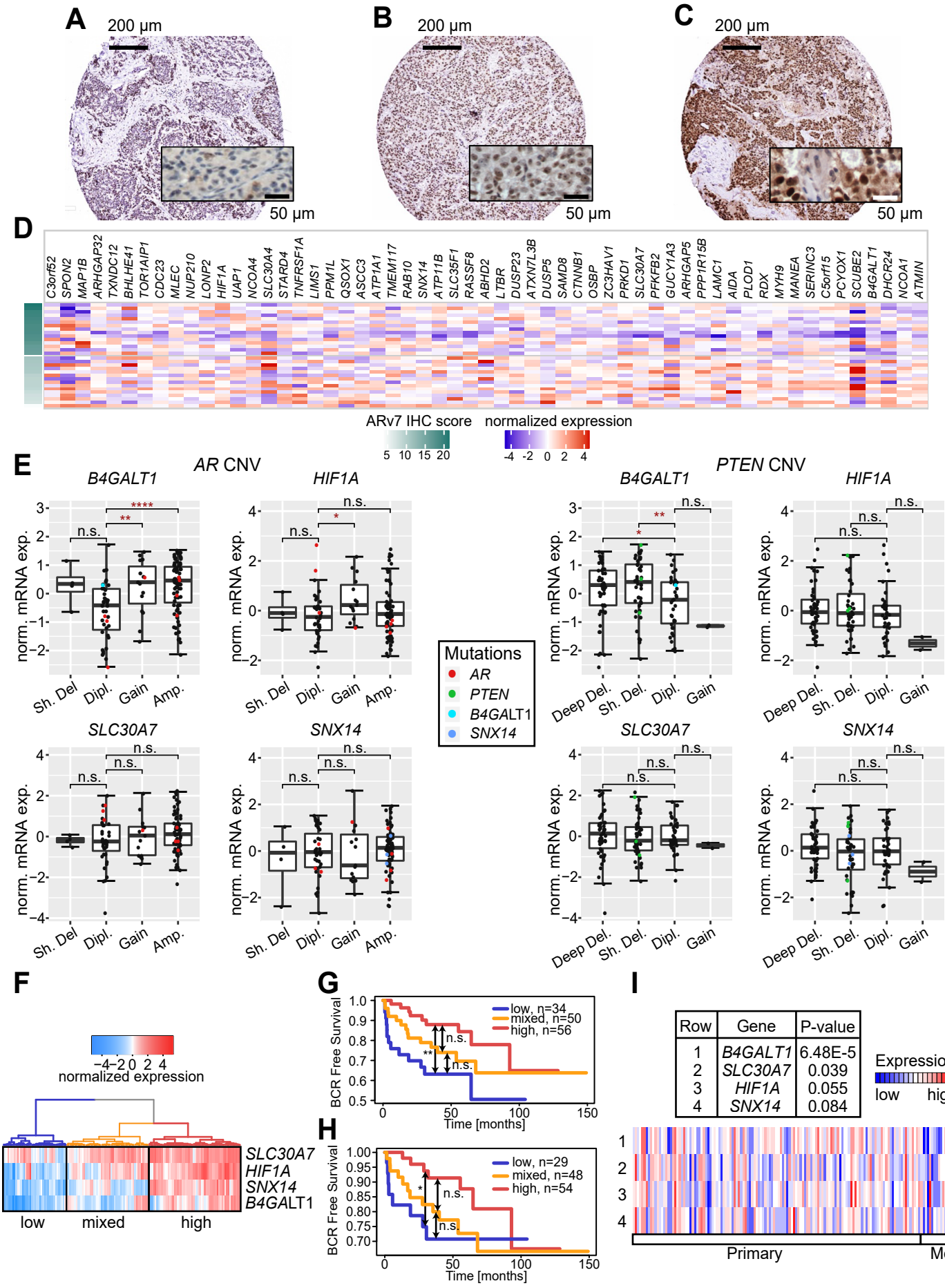
A
200 µm
50 µm
B
200 µm
50 µm
C
200 µm
50 µm
D
C3orf52
SPON2
MAP1B
ARHGAP32
TXNDC12
BHLHE41
TOR1AIP1
CDC23
MLEC
NUP210
LONP2
HIF1A
UAP1
NCOA4
SLC30A4
STARD4
TNFRSF1A
LIMS1
PPM1L
QSOX1
ASCC3
ATP1A1
TMEM117
RAB10
SNX14
ATP11B
SLC35F1
RASSF8
ABHD2
LTBR
DUSP23
ATXN7L3B
DUSP5
SAMD8
CTNNB1
OSBP
ZC3HAV1
PRKD1
SLC30A7
PFKFB2
GUCY1A3
ARHGAP5
PPP1R15B
LAMC1
AIDA
PLOD1
RDX
MYH9
MANEA
SERINC3
C5orf15
PCYOX1
SCUBE2
B4GALT1
DHCR24
NCOA1
ATMIN
ARv7 IHC score
5 10 15 20
normalized expression
-4 -2 0 2 4
E
AR CNV
PTEN CNV
B4GALT1
HIF1A
SLC30A7
SNX14
norm. mRNA exp.
Sh. Del
Dipl.
Gain
Amp.
Deep Del.
Mutations
AR
PTEN
B4GALT1
SNX14
n.s.
F
normalized expression
SLC30A7
HIF1A
SNX14
B4GALT1
low
mixed
high
G
BCR Free Survival
Time [months]
low, n=34
mixed, n=50
high, n=56
H
low, n=29
mixed, n=48
high, n=54
I
Row
Gene
P-value
1
B4GALT1
6.48E-5
2
SLC30A7
0.039
3
HIF1A
0.055
4
SNX14
0.084
Expression
low
high
Primary
Met

Figure S6 (related to Figure 6). IHC controls for the ARv7-specific antibody and expression levels of ARv7-repressed target genes that inhibit CRPC growth

A-C. Representative staining and scores (average nuclear optical density multiplied with the fraction of positive nuclei) from PCa spots on TMA 55, stained with an ARv7-specific antibody. A. low ARv7, score: 4; B. medium ARv7, score: 14; C. high ARv7, score: 24. **D.** Heatmap of leading edge gene expression levels in patient samples, as determined by GSEA (see Figure 6B). Patients were selected according to the lowest (low ARv7 expression) and highest quartiles (high ARv7 expression) of ARv7 IHC scores. **E.** Expression of *B4GALT1*, *HIF1A*, *SLC30A7* and *SNX14* in patients with known *AR* (left) and *PTEN* (right) copy number variations (CNV). Patients were grouped by cbioportal (Cerami et al., 2012) into deep deletions potentially missing both copies (Deep Del), shallow deletion potentially missing one copy (Sh. Del.), diploid having two copies (Dipl.) gain indicating a low level amplification (Gain) and amplification indicating a high level amplification (Amp.). Samples with *AR* (red), *PTEN* (green), *B4GALT1* (azure) or *SNX14* mutations (blue) are color-coded. No mutations were detected for *SLC30A7* or *HIF1A*. Boxplots show the median, the 1st and 3rd quartile. Whiskers extend to 1.5 the interquartile range. P values were calculated using Wilcoxon test; n.s.; not significant $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ****: $p \leq 0.0001$. **F.** Hierarchical clustering of the 4 ARv7-regulated, positively-selected target genes from Figure 6D, and gene expression data from patient samples (Taylor et al., 2010). Three clusters, low (blue), mixed (orange) and high (red), based on the average expression of the 4 genes, are shown. **G, H.** Kaplan-Meier graphs of PSA recurrence-free survival (BCR Free Survival) for the three patient clusters defined in F. Data for either primary or metastatic samples (G), or primary samples only are shown (H). P values were determined using the log rank test; ns: not significant; *: $p < 0.05$; **: $p < 0.01$. **I.** Heatmap of gene expression levels of the ARv7-repressed target genes B4GALT1, SLC30A7, HIF1A, and SNX14 in primary compared to metastatic PCa in the Taylor dataset (Taylor et al., 2010).