Cell Reports, Volume *19*

Supplemental Information

Androgen Receptor Deregulation Drives Bromodomain-Mediated Chromatin Alterations in Prostate Cancer

Alfonso Urbanucci, Stefan J. Barfeld, Ville Kytölä, Harri M. Itkonen, Ilsa M. Coleman, Daniel Vodák, Liisa Sjöblom, Xia Sheng, Teemu Tolonen, Sarah Minner, Christoph Burdelski, Kati K. Kivinummi, Annika Kohvakka, Steven Kregel, Mandeep Takhar, Mohammed Alshalalfa, Elai Davicioni, Nicholas Erho, Paul Lloyd, R. Jeffrey Karnes, Ashley E. Ross, Edward M. Schaeffer, Donald J. Vander Griend, Stefan Knapp, Eva Corey, Felix Y. Feng, Peter S. Nelson, Fahri Saatcioglu, Karen E. Knudsen, Teuvo L.J. Tammela, Guido Sauter, Thorsten Schlomm, Matti Nykter, Tapio Visakorpi, and Ian G. Mills

Supplementary Figure 1

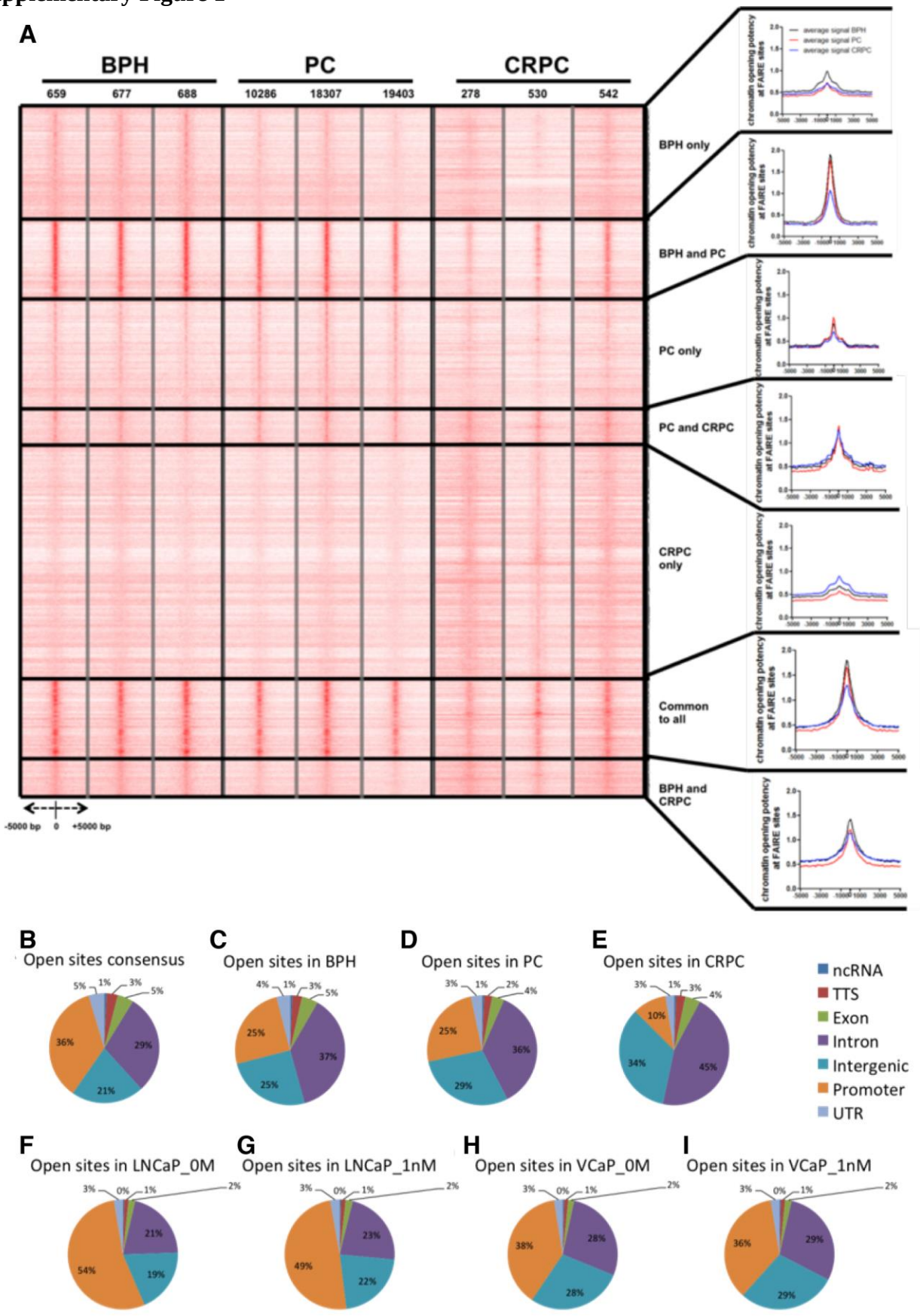


Figure S1. The distribution of FAIRE-seq reads in benign and cancerous prostate tissue samples and their genomic distribution, Related to Figure 1. (A) Formaldehyde–assisted isolation of regulatory elements followed by sequencing (FAIRE-seq) was used to retrieve

accessible genomic regions in tissue specimens from benign prostate hyperplasia (BPH), primary prostate cancer (PC) derived from prostatectomy and locally recurrent castration resistant prostate cancer (CRPC) from transurethral resection of the prostate. Only high confidence open chromatin regions in 3 BPH (23574 sites), 3 primary PC (23341 sites), and 3 CRPC (29160 sites) tissue samples were considered in this analysis. Overlapping and unique peak centres (as indicated on the left) were used to assess distribution of FAIRE-seq reads in the three tissue samples types around these sites (+-5000bp from the centre). Average read distribution was assessed for each category (BPH, PC, and CRPC) of FAIRE-seq peaks and is indicated in the line charts on the right side of the figure. Distribution of common chromatin open sites in all 9 clinical tissue samples (**B**), in all benign prostate hyperplasia (BPH) samples (**C**), in all primary prostate cancer (PC) tissue samples (**D**), and in all castration resistant prostate cancer tissue samples (**E**) according to FAIRE-seq analysis. Distribution of chromatin open sites in LNCaP (**F&G**) and VCaP (**H&I**) cells treated with vehicle (**F&H**) or 1 nM R1881 for 4 hours (**G&I**) following three days of hormone starvation prior to the FAIRE-seq assay.

Supplementary Figure 2

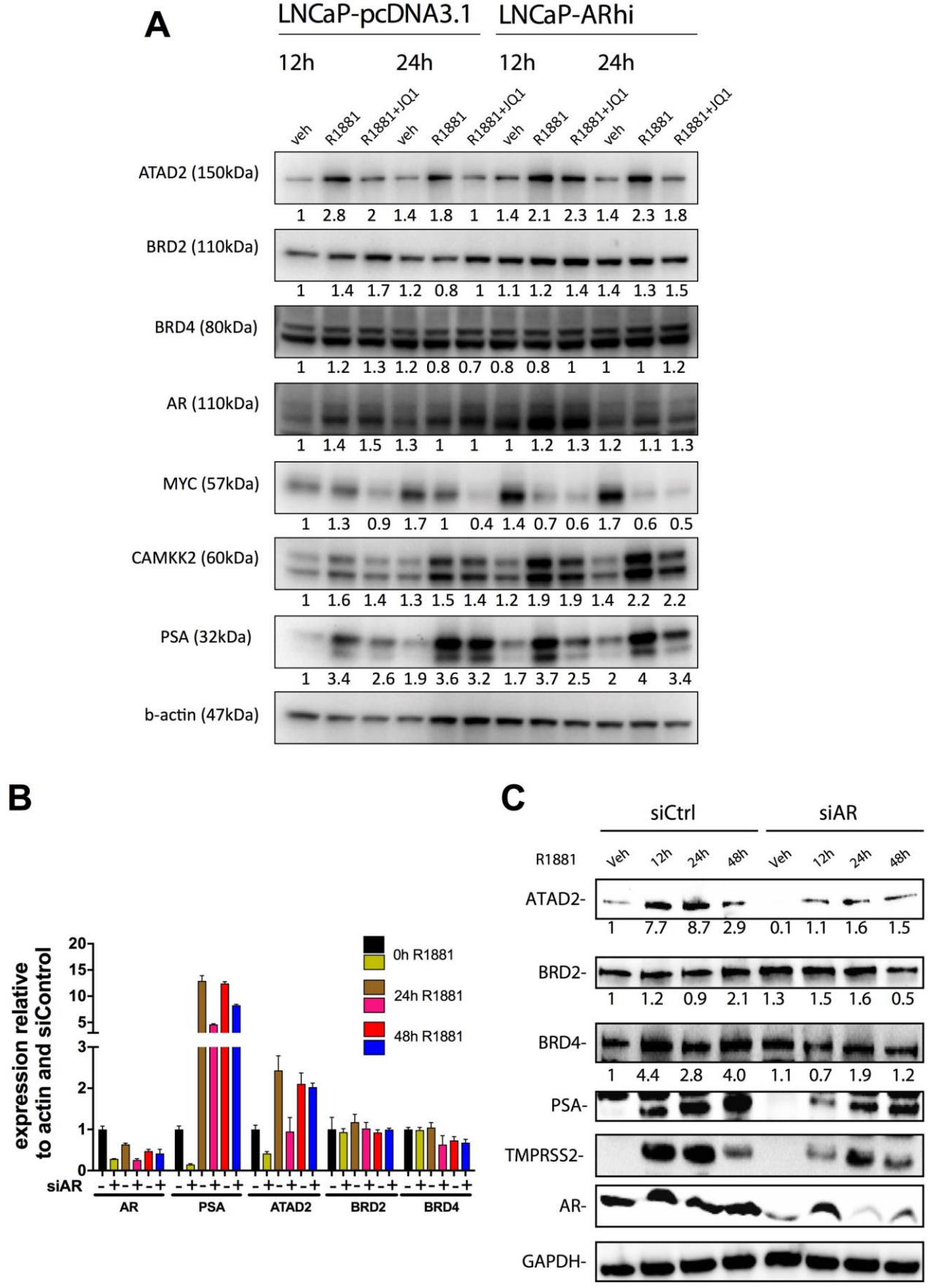
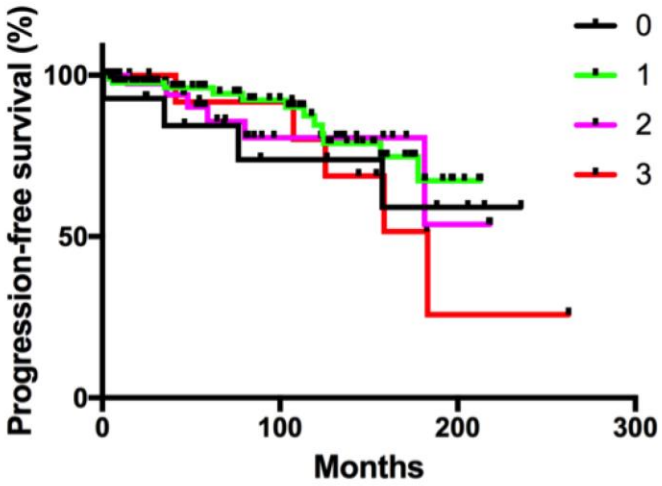


Figure S2. Bromodomain containing protein regulation by androgens, Related to Figure 3. (**A**) Western blot analysis of ATAD2, BRD2, BRD4, AR, MYC, CAMKK2, and PSA in LNCaP-pcDN3.1 and LNCaP-ARhi (LNCaP ARhi) cells 12 and 24 hours after treatment with vehicle (veh

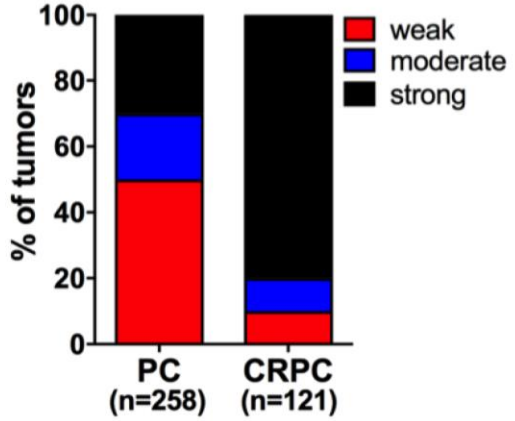
= ethanol), 1nM synthetic androgen R1881 or 1 nM R1881 and 125 nM JQ1. Quantity One software was used to measure the intensity of bands as indicated. β-actin and LNCaP-pcDNA3.1 cells treated for 12 hours with vehicle was used as the point of normalization. Indicated genes' qRT-PCR (**B**) and Western blot analysis (**C**) in LNCaP reverse transfected with either control or siRNA against AR. The cells were hormone starved for 2 days before the treatment with 1 nM R1881 for the indicated time points. Ratio values relative to siControl and are shown for ATAD2, BRD2, and BRD4 blots. TMPRSS2, PSA, and AR levels are shown as control.

Supplementary Figures 3A-H

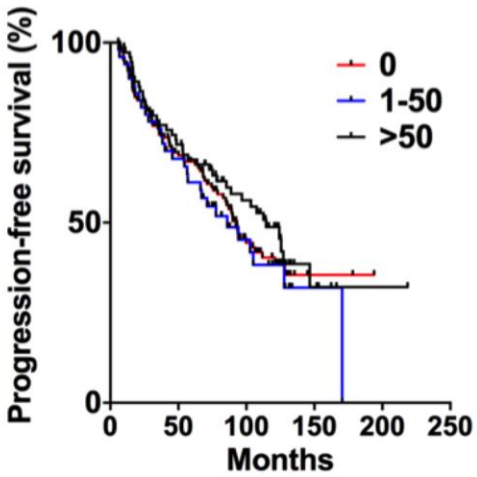
A



B



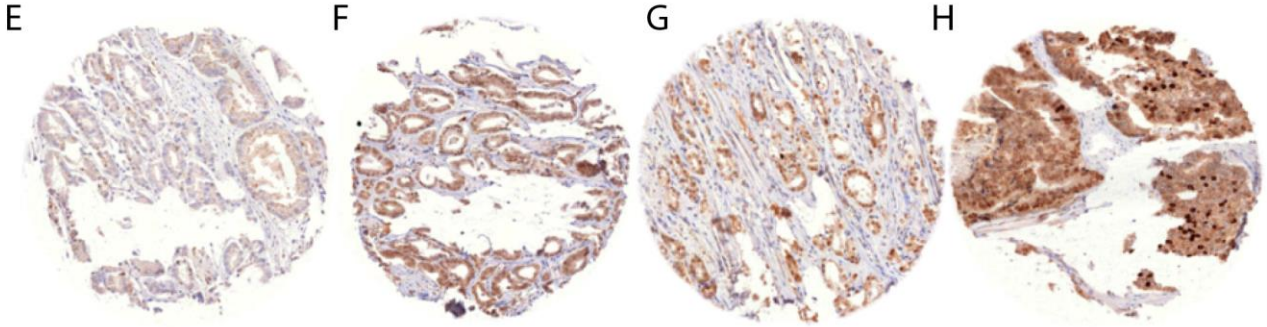
C



D

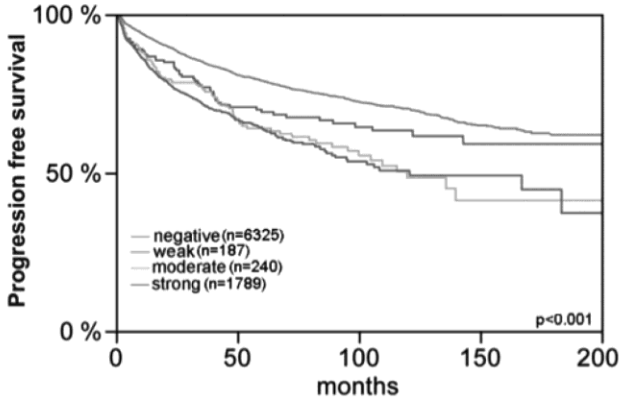
Variable	ATAD2 cytoplasmic histoscore			P	ATAD2 % of positive nuclei			P
	0	1-50	51-300		0	1-5	>5	
Prostatectomy specimens, n (%)	129 (50)	52 (20)	77 (30)		169 (66)	83 (32)	6 (2)	
Locally recurrent CRPCs, n (%)	12 (10)	12 (10)	97 (80)	<0.0001	39(32)	38 (32)	43 (36)	<0.0001
Prostatectomy specimens:								
Gleason score, n (%)								
<7	51 (53)	16 (17)	29 (30)		73 (76)	23 (24)	0 (0)	
7	61 (49)	28 (23)	35 (28)		76 (61)	44 (36)	4 (3)	
>7	17 (49)	7 (20)	11(31)	0.866	20 (57)	13 (37)	2 (6)	0.053
pT-stage, n (%)								
pT2	84 (47)	37 (21)	56 (32)		121 (68)	55 (31)	1 (1)	
pT3	45 (56)	15 (19)	20 (25)	0.407	48 (60)	27 (34)	5 (6)	0.015
PSA ng/ml (mean ± SD)	18.1 ± 25.4	12.8 ± 9.4	12.4 ± 8.4	0.091	16.3 ± 22.1	12.9 ± 9.2	17.0 ± 10.0	0.423
Age (mean ± SD)	63.0 ± 4.9	64.3± 4.8	62.3 ± 5.8	0.125	63.4 ± 4.9	62.2 ± 5.8	63.8 ± 5.8	0.288

E F G H

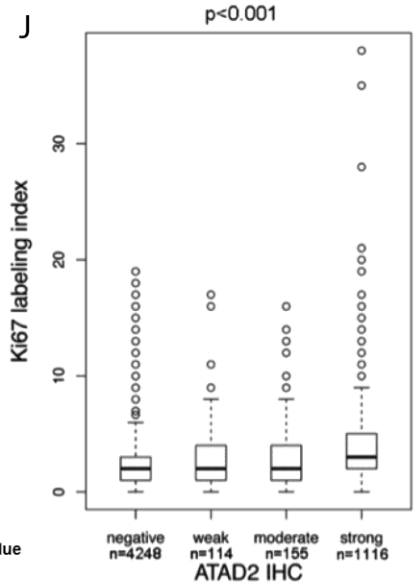


Supplementary Figures 3I-M

I



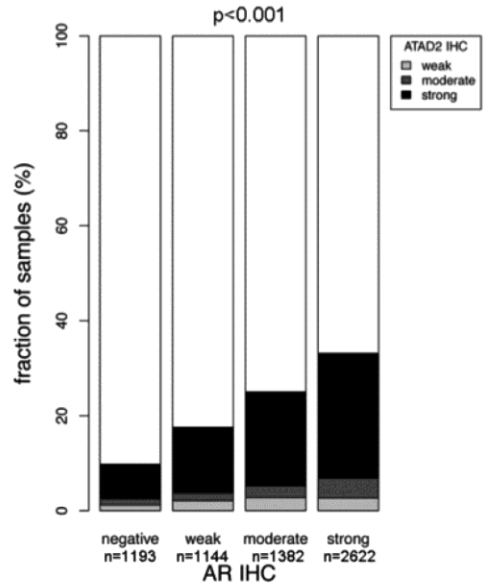
J



K

Parameter	ATAD2				n evaluable	p value
	negative (%)	weak (%)	moderate (%)	strong (%)		
All cancers	73,7	2,2	2,8	21,3	9467	
Tumor stage						<0.0001
pT2	78,1	1,8	2,3	17,7	6085	
pT3a	68,7	2,8	3,2	25,3	2111	
pT3b	60,3	2,9	4,2	32,7	1179	
pT4	55,6	1,9	13,0	29,6	54	
Gleason grade						<0.0001
≤3+3	86,4	2,2	2,1	9,3	2141	
3+4	74,6	1,9	2,4	21,0	5368	
4+3	59,4	2,9	4,3	33,4	1451	
≥4+4	48,5	2,6	5,9	43,0	460	
Lymph node metastasis						<0.0001
N0	71,2	2,2	2,9	23,7	5460	
N+	56,3	2,0	3,3	38,4	549	
Preop. PSA level (ng/ml)						0.0006
<4	75,9	1,6	2,8	19,7	1154	
4-10	74,4	1,9	2,5	21,2	5665	
10-20	71,0	3,1	3,1	22,8	1878	
>20	69,7	3,3	4,2	22,9	669	
Surgical margin						0.0005
negative	74,7	2,2	2,8	20,3	7495	
positive	70,4	2,2	2,6	24,8	1799	

L



M

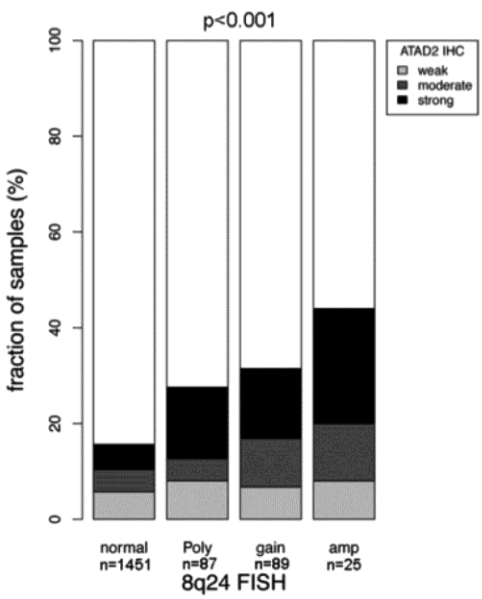


Figure S3. Immunohistochemical analysis of bromodomain containing proteins in two prostate cancer cohorts (Tampere and Hamburg) related to Figure 4. BRD4 (**A** and **B**) and ATAD2 (**C** and **D**) analyses in the Tampere cohort. (**A**) Kaplan–Meier analysis of biochemical progression-free survival in 159 prostatectomy-treated patients according to BRD4 long isoform staining. (**B**) Percentage of tumors according to ATAD2 cytoplasmic histoscore in PC (n=258) and CRPC (n=121) specimens ($p<0.0001$ according to X^2 test). (**C**) Kaplan–Meier analysis of biochemical progression-free survival in prostatectomy-treated patients according to ATAD2 cytoplasmic histoscore ($p=0.5234$ calculated with Mantel–Cox test). (**D**) Association of ATAD2 cytoplasmic histoscore and ATAD2 percentage of positive nuclei with Gleason Score, pT stage, PSA testing, and age. Analysis of ATAD2 in prostate cancer in Hamburg validation cohort;staining intensity of all prostate cancer cases was semi-quantitatively assessed in four categories for which representative images are given: (**E**) negative, (**F**) weak, (**G**) moderate, and (**H**) strong. (**I**). Kaplan–Meier analysis of biochemical progression-free survival in the Hamburg validation cohort of 8541 prostatectomy-treated patients according to ATAD2 staining.Association between ATAD2 immunostaining intensity and Ki67 index label (**J**). (**K**) Association between ATAD2 immunostaining results and prostate cancer phenotype in all cancers. Percentage of tumors that co-stained for androgen receptor (AR) and ATAD2 ($p<0.001$ according to X^2 test) (**L**), and were positive for 8q24 locus (ATAD2 locus) alteration as assessed by fluorescence in situ hybridization (FISH) ($p=0.0009$) (**M**).

Supplementary Figure 4

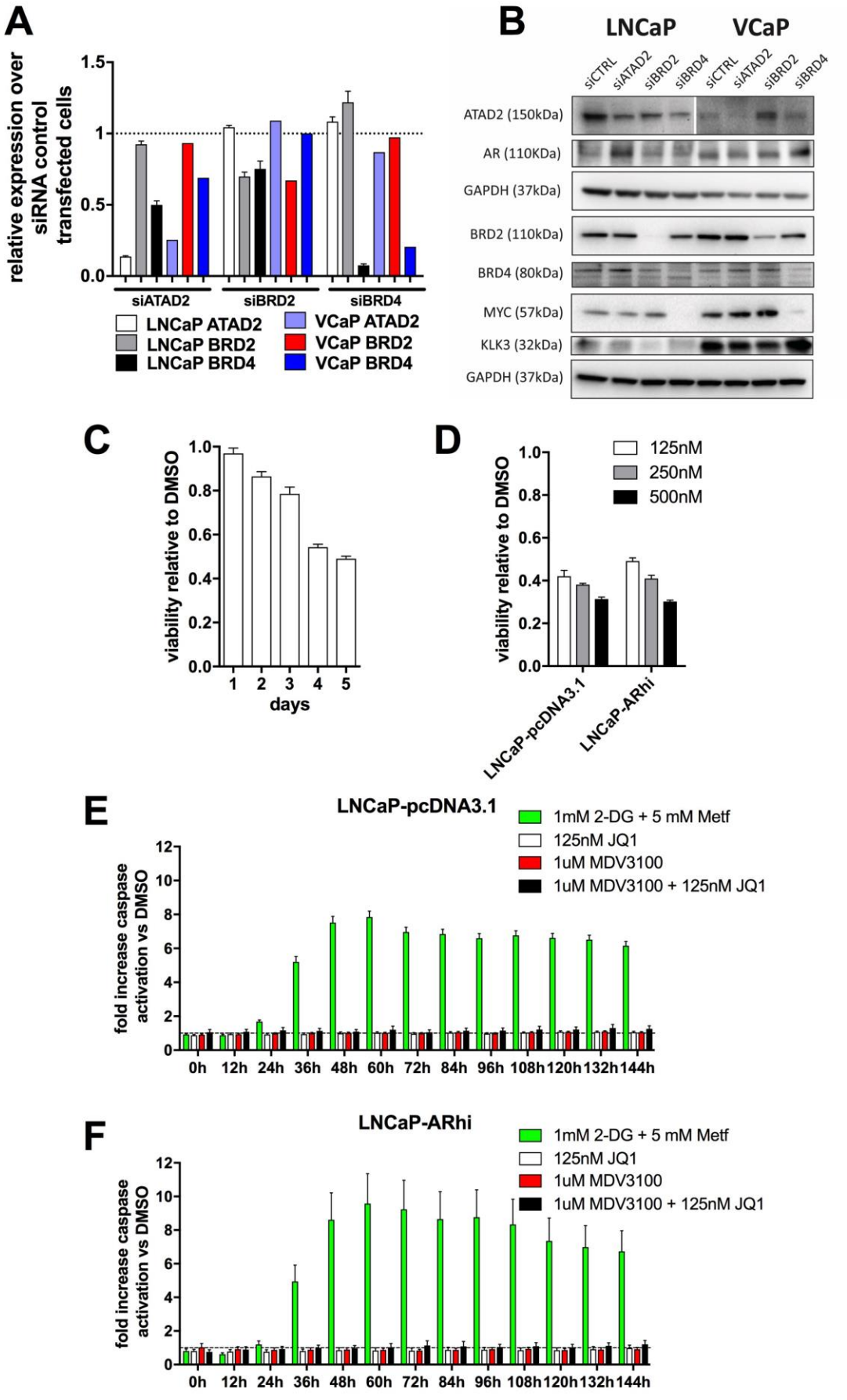
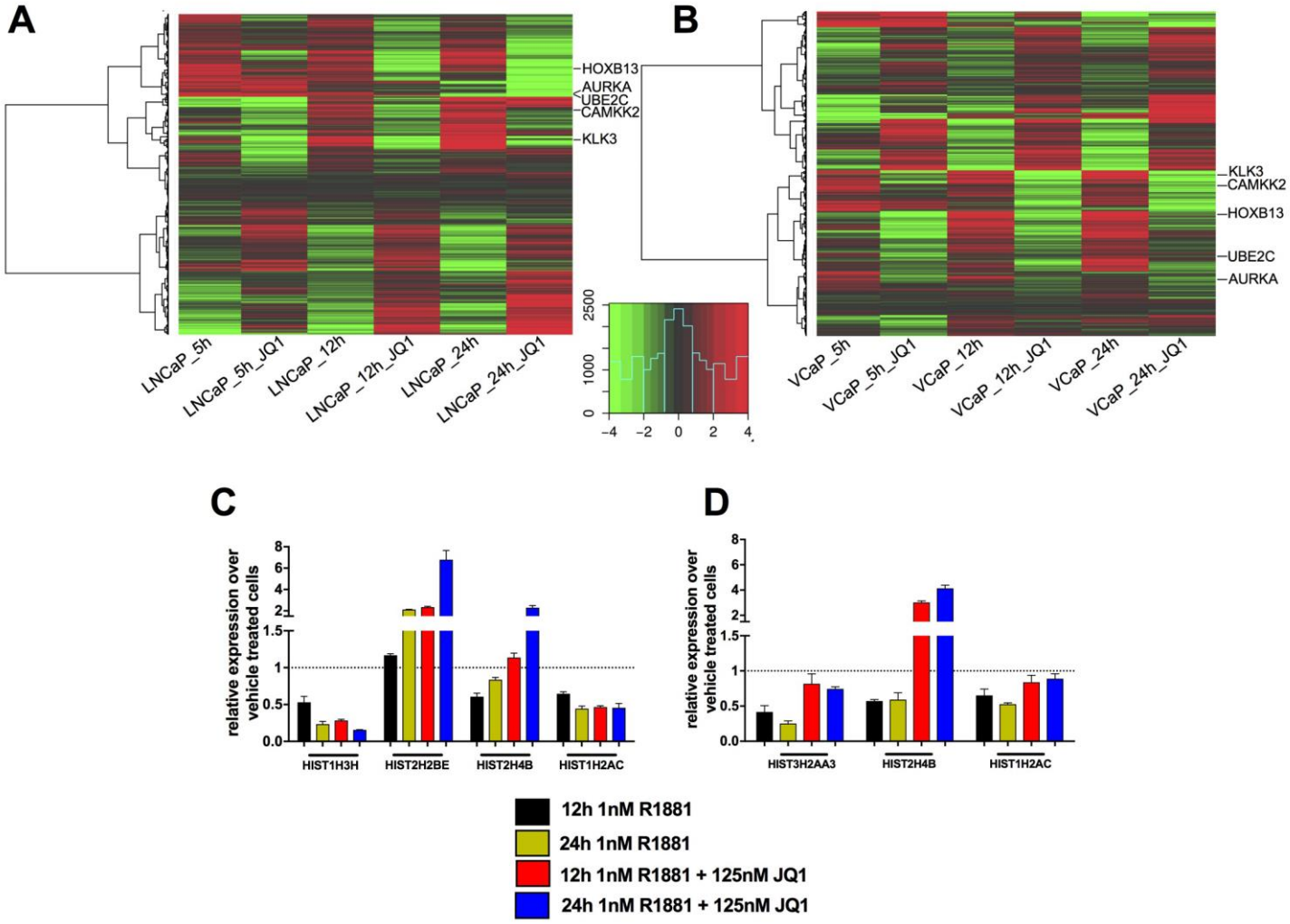


Figure S4. Impact of bromodomain inhibition on androgen receptor positive prostate cancer cells, Related to Figure 5. (**A**) qPCR analysis of ATAD2, BRD2, and BRD4 transcripts in LNCaP

and VCaP cells three days after transfection with siRNA control or siRNA against ATAD2 (siATAD2), siRNA against BRD2 (siBRD2), or siRNA against BRD4 (siBRD4). The relative expression of each gene normalized against house-keeping gene and normalization against the relative siRNA control value are shown. (**B**) Western blot analysis of ATAD2, BRD2, and BRD4 proteins knockdown in LNCaP and VCaP cells following three days transfection with the indicated siRNAs. Protein levels of AR, PSA (KLK3), and MYC are also shown. (**C**) Relative viability of 22RV1 cells cultured in full serum and treated with DMSO or 125 nM bromodomain inhibitor JQ1 for the indicated time. (**D**) Relative viability of LNCaP-pcDNA3.1 cells and AR overexpressing LNCaP-ARhi cells cultured in full serum and treated for 96 hours with the indicated concentrations of JQ1. Caspase activation assay following a time course treatment of 2-deoxy-glucose (2-DG) supplemented with metformin (Metf) to induce caspase activation as positive control, JQ1, MDV3100 or the combination of the last two (indicated concentrations) in LNCaP-pcDNA3.1 (**E**) and LNCaP-ARhi (**F**) cells

Supplementary Figures 5A-D



Supplementary Figures 5 E-J

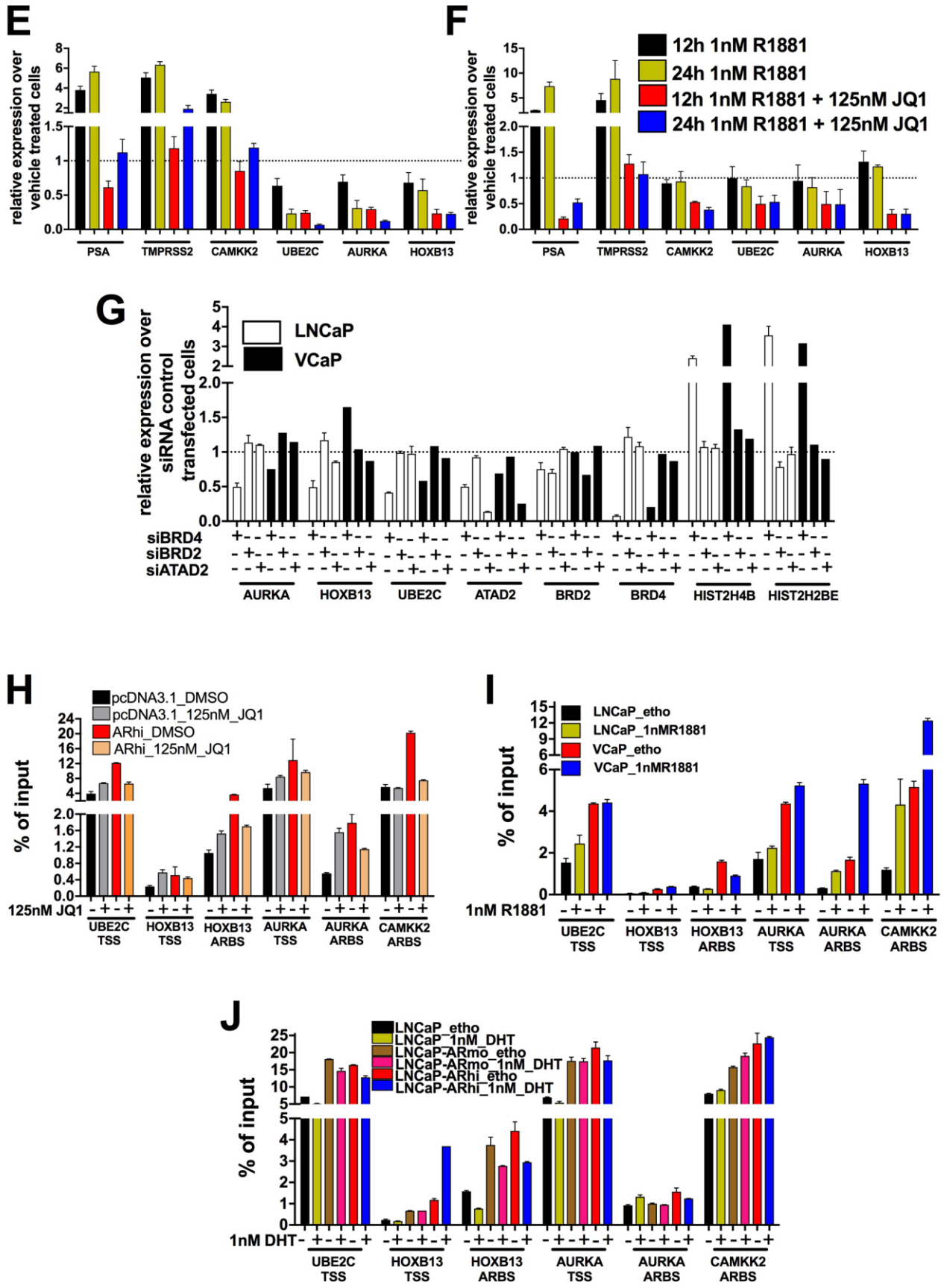
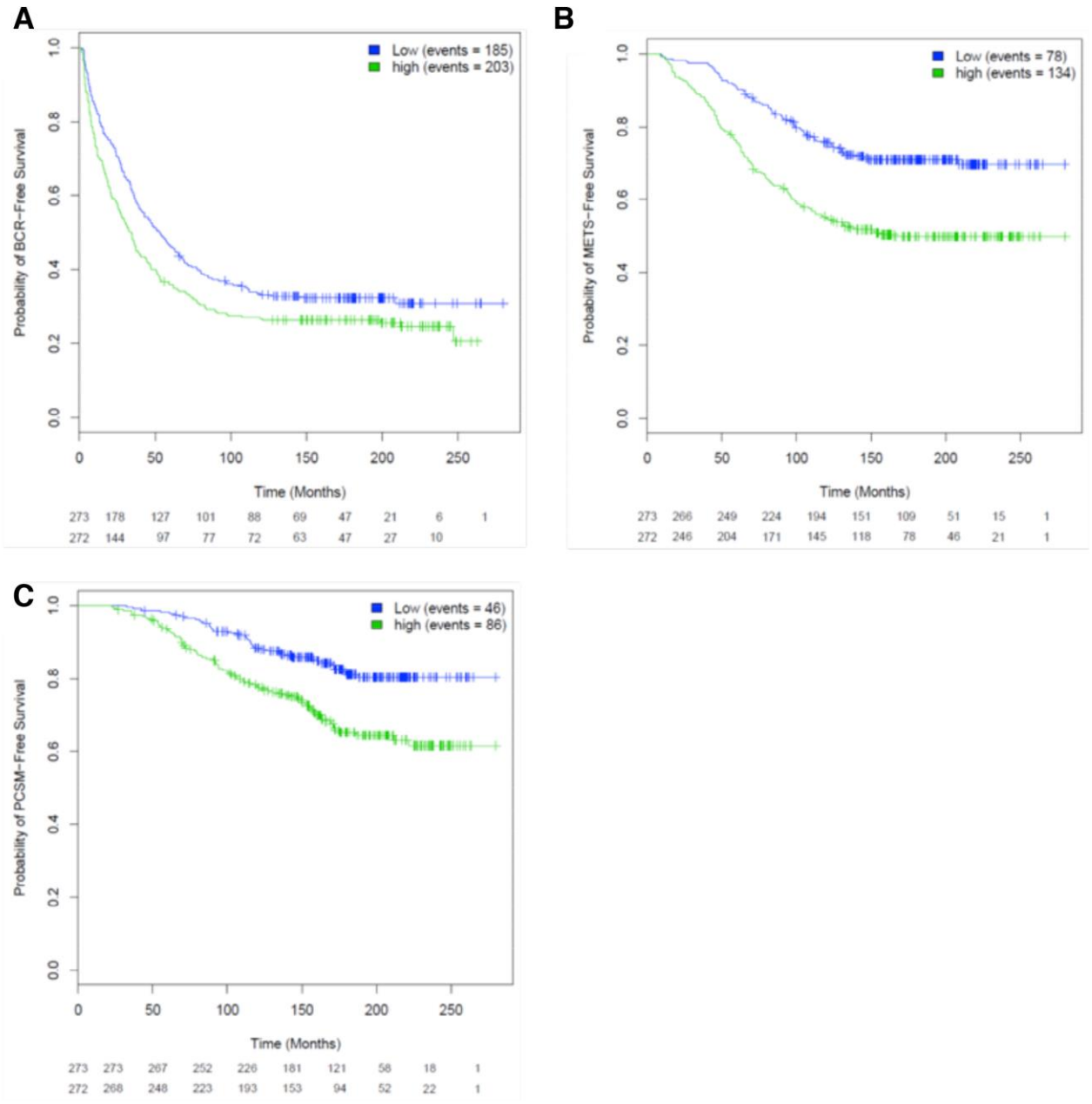


Figure S5. **Impact of JQ1 on the transcriptional program of prostate cancer cells shows downregulation of key androgen receptor target genes and upregulation of histone genes, Related to Figure 6.** Heat maps of genes up- and down-regulated by JQ1 in LNCaP (**A**) and VCaP

(**B**) cells. qRT-PCR validation of micro array data showing upregulation of the indicated histone genes in LNCaP (**C**) and VCaP (**D**) cells upon treatment with JQ1 for the indicated time. qRT-PCR validation of micro array data showing downregulation of the indicated genes in LNCaP (**E**) and VCaP (**F**) cells upon treatment with JQ1 for the indicated time, or upon knockdown of the indicated bromodomains (**G**). Two histone genes also show upregulation upon BRD4 knockdown. Validated FAIRE-seq sites by FAIRE-qPCR analysis (**H-J**). Cells were assayed via FAIRE-qPCR following four days of hormone starvation, in two cases with concomitant treatment with subtoxic concentration of 125 nM small molecule bromodomain inhibitor JQ1 or vehicle (DMSO) (**H**). FAIRE-qPCR validation of local chromatin opening at the indicated loci in LNCaP and VCaP cells treated for 4h with 1 nM R1881 or vehicle (etho) after three days of hormone starvation (**I**) and in a LNCaP-based AR overexpression model (described in Waltering et al.(Waltering et al., 2009)) treated for 4 hours with dihydrotestosterone (DHT) or vehicle (etho) following three days of hormone starvation (**J**).

Supplementary Figures 6A-C



Supplementary Figures 6D-G

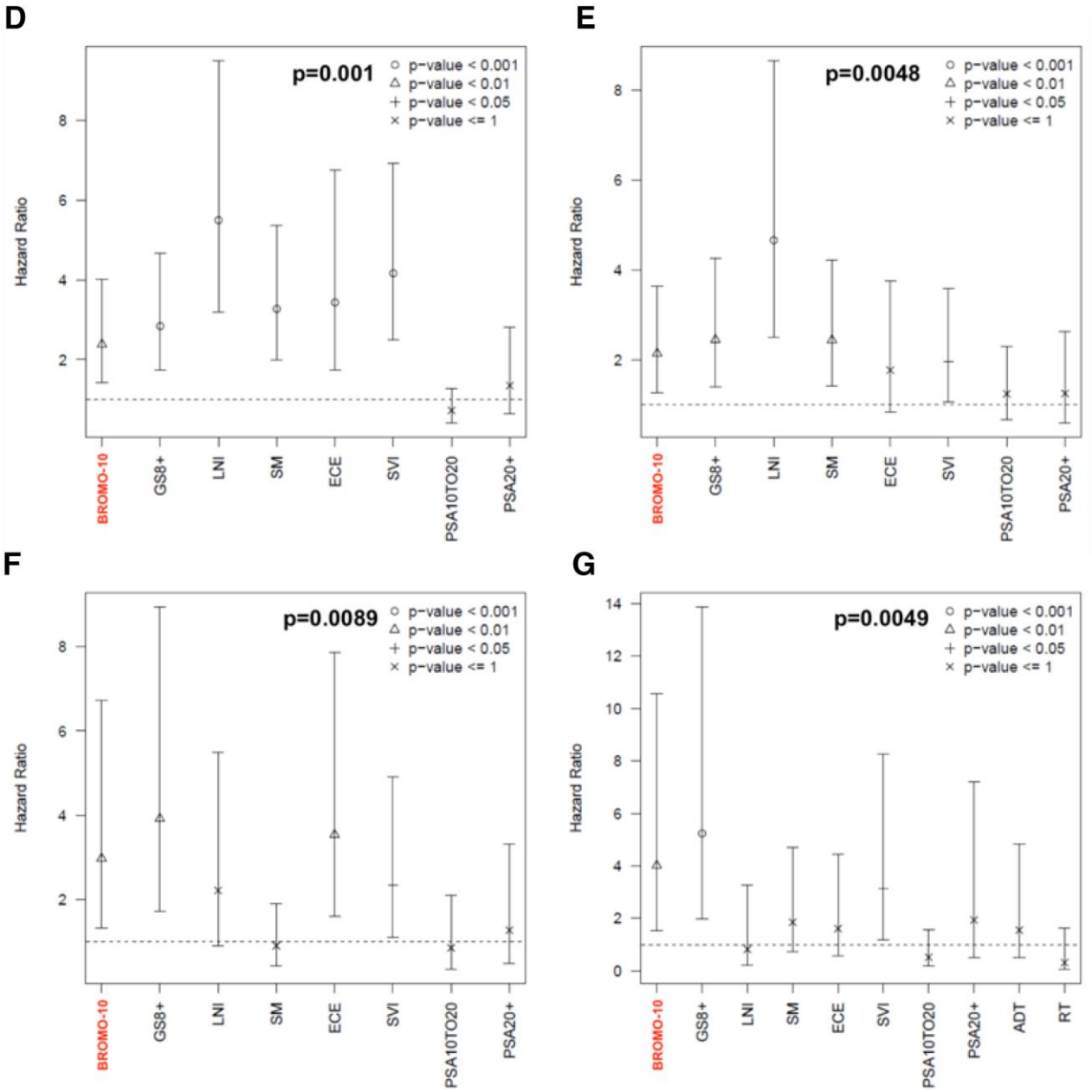


Figure S6. Evaluating the predictive value and independent prognostic contribution of the BROMO-10 signature. Related to Figure 7. Predictive values for the 10 genes in the signature in the MCI cohort of high-risk men, Kaplan-Meier survival curves show (**A**) biochemical recurrence (BCR) -free survival (p=0.00935), (**B**) Metastatic recurrence (METS) -free survival (p=0.0000001), and (**C**) prostate cancer-specific mortality (PCSM) (p=0.00002) for the low and high expression groups of the ten-gene signature (BROMO-10) determined using a median split of the scores in the MCI cohort (n=545)(Erho et al., 2013). The number of patients at risk for each group is show beneath the plot (low and high in the top and bottom raw). The independent prognostic contribution of the BROMO-10 signature over clinicopathological variables as assessed using univariate (**D,F**) and multivariate (**E,G**) analysis and depicted as forest plots applied to the JHMI-RP validation cohort (Ross et al., 2015) for the biochemical recurrence (BCR) endpoint (**D,E**), and in the MCII validation cohort (Karnes et al., 2013) for the prostate cancer-specific mortality (PCSM) endpoint (**F,G**). p-values for each analysis are indicated above the plots. Gleason Score 8 and above with GS 7 as reference (GS8+), lymph node invasion (LNI), surgical margin status (SM) positive, seminal vesicle invasion (SVI), extra capsular extension (ECE), preoperative PSA 10-20 ng/ml

(PSA10TO20), pre-operative PSA 20ng/ml or greater (PSA20+), adjuvant hormone therapy (ADT) and adjuvant radiation therapy (RT) are shown.

Supplemental Experimental Procedures

Clinical material

For Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) followed by sequencing (seq) assays and FAIREqPCR assays 40x10µm microtome slides of freshly frozen tissue samples from 3 benign prostate hyperplasia (BPH), 6 primary prostate cancers (PC) and 3 castration resistant prostate cancers (CRPC) were used.
RNA-sequencing data from transcriptomes of 12 BPH, 28 untreated PCs, and 13 CRPCs was retrieved from Ylipaa et al.,(Ylipaa et al., 2015).
DNA methylation data of 3 BPH, 3 primary PCs and 3 CRPCs samples matching RNAseq and FAIRE-seq data were obtained from Kaukoniemi et al. (manuscript in preparation).
For quantitative PCR analysis of mRNA, freshly frozen tissues from 8 BPH and 27 untreated primary PC samples from prostatectomies, as well as 7 BPH and 15 CRPC specimens from transurethral resection of the prostate (TURP) -treated patients were used. The samples were snap frozen in liquid nitrogen and total RNA was isolated with Trizol™-Reagent (Invitrogen Inc., Carlsbad, California, USA) according to manufacturer's instructions. Tumor samples contained, at least, 70% of cancer cells.
The Tampere patients' cohort of tissue microarrays (TMAs) contained a total of 258 formalin-fixed paraffin embedded prostatectomy and 121 CRPC (TURP) specimens obtained from Tampere University Hospital. Subset of the cohort was used to immunostain for ATAD2, BRD2, and BRD4. For the prostatectomy-treated patients, detectable prostate specific antigen (PSA) values (≥0.5 ng/ml) in two consecutive measurements or the emergence of metastases were considered as signs of progression. The use of TMAs and the above mentioned clinical material has been approved by the ethical committee of Tampere University Hospital and the National Authority for Medicolegal Affairs.
The Hamburg patients' TMA cohort contained 9467 prostatectomy tissue specimens. Radical prostatectomy specimens were available from 12,427 patients, undergoing surgery between 1992 and 2012 at the Department of Urology and the Martini Clinics at the University Medical Center Hamburg-Eppendorf. Follow-up data were available for a total of 12,344 patients with a median follow-up of 36 months (range: 1 to 241 months). PSA values were measured following surgery and PSA recurrence was defined as the time point when postoperative PSA was at least 0.2ng/ml and increasing at subsequent measurements. All prostate specimens were analyzed according to a standard procedure, including a complete embedding of the entire prostate for histological analysis(Schlomm et al., 2008). The TMA manufacturing process was described earlier in detail(Kononen et al., 1998). In short, one 0.6mm core was taken from a representative tissue block from each patient. The tissues were distributed among 27 TMA blocks, each containing 144 to 522 tumor samples. For internal controls, each TMA block also contained various control tissues, including normal prostate tissue. The molecular database attached to this TMA contained results on 8q24 FISH analysis (expanded from El Gammal et al., (El Gammal et al., 2010)), Androgen Receptor (AR) expression (expanded from Minner et al.,(Minner et al., 2011)) and Ki67 labeling index (Ki67LI) data (expanded from Minner at al.,(Minner et al., 2010)). Analysis of patient and corresponding histopathological data for research purposes, as well as construction of tissue microarrays from archived diagnostic tissues, was approved by local laws (HmbKHG, §12,1) and by the local ethics committee (Ethics commission Hamburg, WF-049/09 and PV3652).
All work was carried out in compliance with the Helsinki Declaration.

Cell lines and cell culture procedure

Parental LNCaP, VCaP, and 22RV1 cells were purchased from ATCC and maintained according to the manufacturer instructions in RPMI 1640 or DMEM (Gibco, 21875 and 41966) respectively, supplemented with 10% fetal bovine serum (FBS; Gibco, 10500) at 37 °C and 5% CO2. LNCaP-AR model derivative cells overexpressing AR were described previously(Waltering et al., 2009) and were maintained in geneticin 250µg/ml (Gibco, 10131). The hormone treatments were

performed as previously described (Massie et al., 2011; Urbanucci et al., 2008) with minor modifications: the cells were hormone deprived using phenol free RPMI 1640 or DMEM (Gibco, 11835 and 31053). MDV3100 resistant LNCaP were a gift from Professor Donald J. Vander Griend. The cells were first described in Kregel et al. (2013)(Kregel et al., 2013) and maintained in 10uM MDV3100 prior treatment with JQ1 (a kind gift by Professor Stefan Knapp). LNCaP and VCaP cells were hormone starved for three days and subsequently treated with 1 nM R1881 for 4 hours prior being used for other assays. LNCaP-pcDNA3.1, -ARmo and -ARhi cells were hormone-starved for 4 days and subsequently treated with DHT or with equal volume of ethanol vehicle (0 M).

Formaldehyde-Assisted Isolation of Regulatory Elements

Two replicates were processed for each cell line and condition for subsequent sequencing analysis. 3-5 replicates were processed for qPCR analysis. Four million cells were plated and hormone-deprived for 4 days. Cells were then treated with R1881 or DHT as indicated in the text for 4 hours. Cells were fixed by adding formaldehyde (Merck KGaA, Darmstadt, Germany) in 1% final concentration for 10 minutes at room temperature and lysed in 1% SDS, 10 mM EDTA, 50 mM Tris-HCl containing 2X protease inhibitor (Roche Inc., Mannheim, Germany). To perform tissue FAIRE from the clinical material, 3 ml of PBS containing 2X protease inhibitor (Roche Inc., Mannheim, Germany) were added to 40x20 µm sections of freshly frozen tissue specimens. They were first vigorously mixed 3 times with syringe and 14G needle, then 4 times with 25G needle. The cells were fixed for 10 minutes in room temperature by adding 1/10 volume of fixation solution (11% formaldehyde, 0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM HEPES). Fixation was stopped by adding 1/20 volume of 2.5 M glycine for 5 minutes at room temperature. The cells were pelleted, washed twice in PBS containing 2X protease inhibitor (Roche Inc., Mannheim, Germany) and lysed as above. The chromatin was sonicated to reach a fragment size of 150-300 bp with Bioruptor UCD-200TM-EX instrument (Diagenode Inc., Liège, Belgium).
Three subsequent Phenol:Chloroform:Isoamyl Alcohol (Sigma, P3803) extractions were performed using Phase-Lock heavy tubes (5Prime, 2302830) on the soluble chromatin to isolate protein free DNA in the aqueous phase. The DNA was then further processed as previously described(Urbanucci et al., 2012) prior library preparation with Illumina TrueSeq kit and subsequent sequencing with Illumina HiSeq 2000 or analysed via qPCR.

Chromatin Immunoprecipitation

Chromatin Immunoprecipitation (ChIP) was performed as previously described(Barfeld et al., 2015) using 10 µg of normal rabbit IgG (Santa Cruz Inc., Santa Cruz, California, USA) or 10 µl of N20 AR antibody (Santa Cruz Inc., Santa Cruz, California, USA) previously incubated o/n with magnetic Diagenode beads. The DNA was then purified as previously described(Urbanucci et al., 2012).

RNA isolation and processing for microarrays

Total RNA was isolated using the Qiagen RNeasy kit (Qiagen, 74106) following the manufacturer's recommendations. RNA concentration and purity was measured using a NanoDrop instrument (Thermo Scientific).

Quantitative PCR

For mRNA expression analyses from cell-lines, 500ng to 1µg total RNA were reverse transcribed using the SuperScript VILO kit (Applied Biosystems, 11754) following the manufacturer's recommendations. qRT-PCR was performed using SYBR green master mix (Applied Biosystems, 4385612). Amplification was performed in duplicate series using the ABI 7900HT. The relative expression of each gene against the average value of *TBP* or *β-actin* reference genes was measured using SYBR green master mix (Applied Biosystems, 4385612) and then normalized to vehicle

condition. qRT-PCR of clinical material was performed as previously described(Urbanucci et al., 2012).
For the ChIP-qPCR or FAIRE-qPCR analysis, the enrichment relative to input chromatin was calculated according to the delta Ct method with the percentages being calculated using the formula $2^{-\Delta Ct}$, where ΔCt is Ct(ChIP-template)-Ct(Input), essentially as previously described(Urbanucci et al., 2012). The primers used are listed below (NB. Marked regions refer to **Data Supplemental File 1**):-

FAIREqPCR	
PSA enhancer fw	TGGGACAACTTGCAAACCTG
PSA enhancer rev	CCAGAGTAGGTCTGTTTTCAATCCA
PSA mid region fw	CAGTGGCCATGAGTTTTGTTTG
PSA mid region rev	AACCAATCCAACTGCATTATACACA
PSA promoter fw	CCTAGATGAAGTCTCCATGAGCTACA
PSA promoter rev	GGGAGGGAGAGCTAGCACTTG
H3K27me3-marked region 1 fw	AGAAGCTAAATTAGATACAA
H3K27me3-marked region 1 rev	AGTAAATTTTTCATTCATAC
H3K27me3-marked region 2 fw	TGTTCACCAAATACTGGAGA
H3K27me3-marked region 2 rev	AGTGGGTTTTTGAAGTCTCT
TMPRSS2 promoter fw	GCTCGAGTTTGGGTTAAGGAA
TMPRSS2 promoter rev	TACAGGAGCTCGTGAGGTAGCA
TMPRSS2 enhancer fw	TCCAGGCAGAGGTGTGGC
TMPRSS2 enhancer rev	CGTATGTCTCCCTGCACCACT
UBE2C fw	CACGCGGAGTAAGACGTGTA
UBE2C rev	CGTTGGAAAACGCTAACCAT
HOXB13_TSS fw	AACCATGACCTGCTTTGGTC
HOXB13_TSS rev	TGCCTGGGTAATTCACCATT
HOXB13_ARBS fw	TCCCCTTTCTCAGATGGATG
HOXB13_ARBS rev	TTTCACCACCCTGCTTTCTC
AURKA_TSS fw	CTCGTCCGCCACTGAGATA
AURKA_TSS rev	TTGGAAGACTTGGGTCCTTG
AURKA_ARBS fw	TTTGCAGCCCTAGAGCAAAT
AURKA_ARBS rev	TCAGACAATGACACATTCATGC
CAMKK2_ARBS fw	AGAACACTGTAGCTCACACAGGCA
CAMKK2_ARBS rev	GGGCACTTCCCAACCTTTCTTACT
ChIPqPCR	
BRD4_prox fw	ATCTCCAGCCGTCTGTTTGT
BRD4_prox rev	GAGTGAGCTCAGCCTCCTTG
BRD4_far fw	CTTTTGGCATGGCTCAGAGT
BRD4_far rev	TGTGGCAGGAAATGAACAAG
BRD2 fw	TGTGCTTGTTTGCCTTTCTG
BRD2 rev	TGGCTCAGTTCCAGTTGCTT

ATAD2 fw	TGTTCAGCAACATCATAGTCCA
ATAD2 rev	GGAAACAAATGTTCAGCAAGAA
RT-PCR	
BRD4-L fw	CTGGACCAGCAGAGGGAGT
BRD4-L rev	ACCTAGGTGCGCTCAGAAAA
BRD4-S fw	GACAGCGAAGACTCCGAAAC
BRD4-S rev	TGGGAAGGAATCTGGAACTG
BRD2 fw	TGAAACACTCAAGCCATCCA
BRD2rev	CCTCCTTTGTCTTTCCCACA
ATAD2 fw	CATCGCAAGGACCATGATAA
ATAD2 rev	TCAATTAGGCGGACATGACA
PSA fw	GCAGCATTGAACCAGAGGAG
PSA rev	AGAACTGGGGAGGCTTGAGT
TMPRSS2 fw	CCAGGAGTGTACGGGAATGT
TMPRSS2 rev	CAGCCCCATTGTTTTCTTGT
CAMKK2 fw	TGAAGACCAGGCCCGTTTCTACTT
CAMKK2 rev	TGGAAGGTTTGATGTCACGGTGGA
HOXB13 fw	AGATGTGTTGCCAGGGAGAA
HOXB13 rev	CTTGCTGTACGGAATGCGTT
AURKA fw	AGGCCACTGAATAACACCCA
AURKA rev	TGATGCCAGTTCCTCCTCAG
UBE2C fw	TGGCGATAAAGGGATTTCTGC
UBE2C rev	CGCATTGTAAGGGTAGCCAC
HIST2H4B fw	CTCAGGCAAAGTGGGAGAAG
HIST2H4B rev	AAGCCCAAAACAGTCAATCG
H1H3H fw	CAGAAAGCTGCCTTTTCAGC
H1H3H rev	GGTTGGTGTCCTCAAAGAGC
H2H2BE fw	TCAGAGCCACCCACCTAATC
H2H2BE rev	GTATGCCATTTCCCATGACC
HIST1H2AC fw	CTCCGTAAAGGCAACTACGC
HIST1H2AC rev	TGCGAGTCTTCTTGTTGTCG
HIST2H2AA3 fw	AATAGCGAACCTGGAGCTGA
HIST2H2AA3 rev	GAAGAGCCAAGGCAGTTACG
TBP fw	GGGGAGCTGTGATGTGAAGT
TBP rev	GAGCCATTACGTCGTCTTCC
B-Actin	TGGGACGACATGGAGAAAAT
B-Actin	AGAGGCGTACAGGGATAGCA
G3PDH2 fw	TGAGGAGGGGAGATTCAGTG
G3PDH2 rev	GTCAGTGGTGGACCTGACCT

Transcript expression profiling (mRNA)
For microarray analysis, RNA integrity was confirmed using a 2100 Bioanalyzer (Agilent) and Total RNA Nano Chip (Agilent, 5067-1511). 500ng RNA were reverse transcribed and Biotin-labeled using the TotalPrep-96 RNA Amplification kit (Illumina, 4393543) following the manufacturer's recommendations. Resuspended cRNA samples were hybridized onto Human HT-12 Expression BeadChips (Illumina, BD-103-0204). Missing probes were imputed using Illumina's GenomeStudio Gene Expression Module.
Analysis of microarray data from JQ1/R1881 treated cells was performed as follows: the imputed probe datasets were analyzed using the freely available J-Express 2012 software (http://jexpress.bioinfo.no/site/). The raw data was quantile normalized and log2 transformed prior to analysis. Differential expression analysis was performed using the grouped triplicate experiments and Rank product analysis. Probes with a q-value of <0.05 were considered significantly up- or downregulated. The data are deposited in GEO: GSE73989.

ChIP/DHS-seq re-analysis and FAIRE-seq analysis
ChIP-seq data for AR and ERG were retrieved from (Massie et al., 2011; Sharma et al., 2013; Yu et al., 2010)
The raw reads from described studies were mapped with novoalign (http://www.novocraft.com) and to the human reference genome (build hg19) with default parameters. A maximum of 5 read duplicates per genomic location was allowed by our filtering. Peak detection (i.e., binding site detection) of the chromatin immunoprecipitation (ChIP) and DNA hypersensitive sites (DHS) –seq was performed as previously described(Urbanucci et al., 2012).
Peak detection for FAIRE-seq was performed using MACS(Zhang et al., 2008) with default parameters using inputs of each of the FAIRE samples as controls, and with F-Seq(Boyle et al., 2008). F-seq was used taking into account genomic mappability background and copy number/karyotype correction. The human genome background files were downloaded directly from F-seq website resource (http://fureylab.web.unc.edu/software/fseq/). To correct for copy number/karyotype for our cell lines, we created iff files using iffBuilder (http://fureylab.web.unc.edu/software/fseq/). iffBuilder was fed with chromosome-wise wiggle files produced from the input of each FAIRE sample. To produce wiggle files we used a python based script "create_depth_track_window.py" using a 1000bp sliding window at a step=1. See below:-

Python script used to generate wig files to feed IffBuilder, Related to Experimental Procedure

```
#!usr/bin/env python

import sys, string

if (len(sys.argv) == 1):
  print "Run as:\n"
  print "python create_depth_track_window.py ifolder ifile ofolder win_size [-zig] [-hg18]\n"
  print "- ifolder:  folder containing the input file"
  print "- ifile:    the input file (a coordinate-sorted bed file)"
  print "- ofolder:  folder for output files"
  print "- win_size: size of the sliding window"
  print "- [-zig]  : optional parameter, if \"-zig\" is present as an argument, bases with zero coverage will
be ignored during the computation of mean genomic sequencing depth"
  print "- [-hg18]:  optional parameter, if \"-hg18\" is present as an argument, the script works with
chromosome sizes derived from human genomic build hg18\n"
  sys.exit(0)

# input folder - folder containing the input file
ifolder = sys.argv[1]

if (ifolder[-1:] != "/"):
  ifolder = ifolder + "/"
```

```
# input file - a coordinate-sorted bed file
ifile = open(ifolder + sys.argv[2], "r")

# output folder - folder in which the output files should be stored
ofolder = sys.argv[3]

if (ofolder[-1:] != "/"):
   ofolder = ofolder + "/"

# sliding window size
win_size = int(sys.argv[4])

chrom_sizes =

# hg19 chromosome sizes
chrom_sizes["chr1"] = 249250621
chrom_sizes["chr2"] = 243199373
chrom_sizes["chr3"] = 198022430
chrom_sizes["chr4"] = 191154276
chrom_sizes["chr5"] = 180915260
chrom_sizes["chr6"] = 171115067
chrom_sizes["chr7"] = 159138663
chrom_sizes["chr8"] = 146364022
chrom_sizes["chr9"] = 141213431
chrom_sizes["chr10"] = 135534747
chrom_sizes["chr11"] = 135006516
chrom_sizes["chr12"] = 133851895
chrom_sizes["chr13"] = 115169878
chrom_sizes["chr14"] = 107349540
chrom_sizes["chr15"] = 102531392
chrom_sizes["chr16"] = 90354753
chrom_sizes["chr17"] = 81195210
chrom_sizes["chr18"] = 78077248
chrom_sizes["chr19"] = 59128983
chrom_sizes["chr20"] = 63025520
chrom_sizes["chr21"] = 48129895
chrom_sizes["chr22"] = 51304566
chrom_sizes["chrX"] = 155270560
chrom_sizes["chrY"] = 59373566
chrom_sizes["chrM"] = 16571

if ("-hg18" in sys.argv):
   chrom_sizes["chr1"] = 247249719
   chrom_sizes["chr2"] = 242951149
   chrom_sizes["chr3"] = 199501827
   chrom_sizes["chr4"] = 191273063
   chrom_sizes["chr5"] = 180857866
   chrom_sizes["chr6"] = 170899992
   chrom_sizes["chr7"] = 158821424
   chrom_sizes["chr8"] = 146274826
   chrom_sizes["chr9"] = 140273252
   chrom_sizes["chr10"] = 135374737
   chrom_sizes["chr11"] = 134452384
   chrom_sizes["chr12"] = 132349534
   chrom_sizes["chr13"] = 114142980
   chrom_sizes["chr14"] = 106368585
   chrom_sizes["chr15"] = 100338915
   chrom_sizes["chr16"] = 88827254
   chrom_sizes["chr17"] = 78774742
   chrom_sizes["chr18"] = 76117153
   chrom_sizes["chr19"] = 63811651
```

```
   chrom_sizes["chr20"] = 62435964
   chrom_sizes["chr21"] = 46944323
   chrom_sizes["chr22"] = 49691432
   chrom_sizes["chrX"] = 154913754
   chrom_sizes["chrY"] = 57772954
   chrom_sizes["chrM"] = 16571

   print "Optional parameter \"-hg18\" recognized: the script will be working with chromosome sizes
derived from human genomic build hg18.\n"
else:
   print "Optional parameter \"-hg18\" not recognized: the script will be working with chromosome sizes
derived from human genomic build hg19.\n"
pos_vals = {}   # relevant positions on currently processed chromosome and their seq. depth
final_pos_vals = {}   # values for positions that can no longer be updated (ready for ouput if all values in
given sliding window are available)
out_FPV = 0   # last output position (passed from final_pos_vals to the output)
out_sum_FPV = 0   # sum of sequencing depths for all bases inside the sliding window
elem_count_FPV = 0   # number of bases inside the sliding window
max_FPV = 0   # current maximum position in the final_pos_vals directory

val_distrib = {}   # abundances of sequencing depth values

last_chrom = ""   # chromosome of the last processed read
last_pos = 0   # last genomic position passed from pos_vals to final_pos_vals

# an output buffer and its maximal allowed size
out_buffer = ""
out_buffer_size = 0
out_buffer_max_size = 1024*1024*16

ofile = ""

zero_pos = 0
non_zero_pos = 0
cov_sum = 0

for line in ifile:
   line = string.strip(line)
   line_s = line.split()

   chrom = line_s[0]
   s_pos = int(line_s[1]) + 1
   e_pos = int(line_s[2])

   # upon chromosome change:
   # empty pos_vals and process any trailing zero-depth positions, change the output file
   if not (chrom == last_chrom):

      if not (last_chrom == ""):

         if (len(out_buffer) > 0):
            ofile.write(out_buffer)
            out_buffer = ""
            out_buffer_size = 0

         for i in range(last_pos + 1, chrom_sizes[last_chrom] + 1):

            if (pos_vals.has_key(i)):
               final_pos_vals[i] = pos_vals[i]
               max_FPV = i

               if not (val_distrib.has_key(pos_vals[i])):
```

```
            val_distrib[pos_vals[i]] = 1
          else:
            val_distrib[pos_vals[i]] += 1

          non_zero_pos += 1
          cov_sum += pos_vals[i]
          del(pos_vals[i])

        else:
          max_FPV = i

          if not (val_distrib.has_key(0)):
            val_distrib[0] = 1
          else:
            val_distrib[0] += 1

          zero_pos += 1

      for i in range(out_FPV + 1, chrom_sizes[last_chrom] + 1):

        if (i == 1):
          for j in range(1, win_size/2 + 1):
            elem_count_FPV += 1
            if (final_pos_vals.has_key(j)):
              out_sum_FPV += final_pos_vals[j]

        if (i - win_size/2 > 1):
          elem_count_FPV -= 1

          if (final_pos_vals.has_key(i - win_size/2 - 1)):
            out_sum_FPV -= final_pos_vals[i - win_size/2 - 1]
            del(final_pos_vals[i - win_size/2 - 1])

        if (i + win_size/2 <= chrom_sizes[last_chrom]):
          elem_count_FPV += 1

          if (final_pos_vals.has_key(i + win_size/2)):
            out_sum_FPV += final_pos_vals[i + win_size/2]

        outval = round(float(out_sum_FPV)/elem_count_FPV, 4)
        ofile.write(str(outval) + "\n")

      final_pos_vals = {}
      out_FPV = 0
      out_sum_FPV = 0
      elem_count_FPV = 0
      max_FPV = 0

      pos_vals = {}
      last_pos = 0

    if not (ofile == ""):
      ofile.close()
    ofile = open(ofolder + sys.argv[2] + "." + chrom, "w")
    print("Creating a smooth coverage profile for chromosome " + chrom + " ..\n")
    last_chrom = chrom

  # output values for genommic positions that can't get updated anymore
  if (s_pos > last_pos + 1):
    for i in range(last_pos + 1, s_pos):
      if (pos_vals.has_key(i)):
        final_pos_vals[i] = pos_vals[i]
```

```
        max_FPV = i

        if not (val_distrib.has_key(pos_vals[i])):
          val_distrib[pos_vals[i]] = 1
        else:
          val_distrib[pos_vals[i]] += 1

        non_zero_pos += 1
        cov_sum += pos_vals[i]

        del(pos_vals[i])

      else:
        max_FPV = i

        if not (val_distrib.has_key(0)):
          val_distrib[0] = 1
        else:
          val_distrib[0] += 1

        zero_pos += 1

    last_pos = s_pos - 1

    if (out_FPV + 1 + win_size/2 <= max_FPV):
      for i in range(out_FPV + 1, max_FPV - win_size/2 + 1):

        if (i == 1):
          for j in range(1, win_size/2 + 1):
            elem_count_FPV += 1
            if (final_pos_vals.has_key(j)):
              out_sum_FPV += final_pos_vals[j]

        if (i - win_size/2 > 1):
          elem_count_FPV -= 1

          if (final_pos_vals.has_key(i - win_size/2 - 1)):
            out_sum_FPV -= final_pos_vals[i - win_size/2 - 1]
            del(final_pos_vals[i - win_size/2 - 1])

        if (i + win_size/2 <= chrom_sizes[last_chrom]):
          elem_count_FPV += 1

          if (final_pos_vals.has_key(i + win_size/2)):
            out_sum_FPV += final_pos_vals[i + win_size/2]

        outval = round(float(out_sum_FPV)/elem_count_FPV, 4)
        out_buffer += str(outval) + "\n"
        out_buffer_size += 1

        if (out_buffer_size > out_buffer_max_size):
          ofile.write(out_buffer)
          out_buffer = ""
          out_buffer_size = 0

      out_FPV = max_FPV - win_size/2

  # update the pos_vals dictionary (increment for genomic positions s_pos to e_pos)
  for i in range(s_pos, e_pos + 1):
    if not (pos_vals.has_key(i)):
      pos_vals[i] = 1
    else:
```

```
        pos_vals[i] += 1

ifile.close()

# empty pos_vals and process trailing zero-depth positions for the last chromosome
if (len(out_buffer) > 0):
   ofile.write(out_buffer)
   out_buffer = ""
   out_buffer_size = 0

for i in range(last_pos + 1, chrom_sizes[last_chrom] + 1):
   if (pos_vals.has_key(i)):
      final_pos_vals[i] = pos_vals[i]
      max_FPV = i

      if not (val_distrib.has_key(pos_vals[i])):
         val_distrib[pos_vals[i]] = 1
      else:
         val_distrib[pos_vals[i]] += 1

      non_zero_pos += 1
      cov_sum += pos_vals[i]

      del(pos_vals[i])

   else:
      max_FPV = i

      if not (val_distrib.has_key(0)):
         val_distrib[0] = 1
      else:
         val_distrib[0] += 1

      zero_pos += 1

for i in range(out_FPV + 1, chrom_sizes[last_chrom] + 1):

   if (i == 1):
      for j in range(1, win_size/2 + 1):
         elem_count_FPV += 1
         if (final_pos_vals.has_key(j)):
            out_sum_FPV += final_pos_vals[j]

   if (i - win_size/2 > 1):
      elem_count_FPV -= 1

      if (final_pos_vals.has_key(i - win_size/2 - 1)):
         out_sum_FPV -= final_pos_vals[i - win_size/2 - 1]
         del(final_pos_vals[i - win_size/2 - 1])

   if (i + win_size/2 <= chrom_sizes[last_chrom]):
      elem_count_FPV += 1

      if (final_pos_vals.has_key(i + win_size/2)):
         out_sum_FPV += final_pos_vals[i + win_size/2]

   outval = round(float(out_sum_FPV)/elem_count_FPV, 4)
   ofile.write(str(outval) + "\n")

final_pos_vals = {}
out_FPV = 0
out_sum_FPV = 0
```

```
elem_count_FPV = 0
max_FPV = 0

last_pos = 0

ofile.close()

# output the depth distribution values
dist_values = val_distrib.keys()
dist_values.sort()

genome_length = 0

for chrom in chrom_sizes:
   genome_length += chrom_sizes[chrom]

dfile = open(ofolder + sys.argv[2] + ".depth_value_distribution", "w")
dfile.write("# assembly: hg19\n")
dfile.write("# total genome length: " + str(genome_length) + " bp\n")
dfile.write("# column seq_depth: sequencing depth\n")
dfile.write("# column bp_count: number of genomic locations (1-bp sites) with respective sequencing
depth\n")
dfile.write("# column genome_fraction_%: column bp_count expressed as fraction of the genome
length\n")
dfile.write("# column acc_genome_fraction_%: accumulation of values from column
genome_fraction_%\n")
dfile.write("seq_depth\tbp_count\tgenome_fraction_%\tacc_genome_fraction_%\n")

acc_b_count = 0

for value in dist_values:
   b_count = val_distrib[value]
   acc_b_count += b_count
   g_fraction = round(float(b_count)*100/genome_length, 5)
   acc_g_fraction = round(float(acc_b_count)*100/genome_length, 5)

   dfile.write(str(value) + "\t" + str(b_count) + "\t" + str(g_fraction) + "\t" + str(acc_g_fraction) + "\n")

dfile.close()

print("Number of zero-coverage positions: " + str(zero_pos))
print("Number of non-zero-coverage positions: " + str(non_zero_pos))
print("Total number of genomic positions: " + str(zero_pos + non_zero_pos) + " (check: " +
str(genome_length) + ")")
print("Total coverage sum: " + str(cov_sum))
print("Genomic mean (zero-coverage positions considered): " +
str(round(float(cov_sum)/(zero_pos+non_zero_pos), 4)))
print("Genomic mean (zero-coverage positions not considered): " +
str(round(float(cov_sum)/non_zero_pos, 4)) + "\n")

gen_mean = float(cov_sum)/(zero_pos+non_zero_pos)

if ("-zig" in sys.argv):
   gen_mean = float(cov_sum)/non_zero_pos
   print "Optional parameter \"-zig\" recognized: bases with zero coverage are not considered in the
genomic mean computation.\n"
else:
   print "Optional parameter \"-zig\" not recognized: bases with zero coverage are considered in the
genomic mean computation.\n"

for chrom_name in chrom_sizes:
   print("Creating a wiggle file for chromosome " + chrom_name + " ..\n")
```

```
    cdfile = open(ofolder + sys.argv[2] + "." + chrom_name, "r")
    wigfile = open(ofolder + sys.argv[2] + "." + chrom_name + ".wig", "w")

    out = False
    start_offset = 1

    for inline in cdfile:

      inline = string.strip(inline)

      if ((not out) and (float(inline) == 0.0)):
        start_offset += 1

      elif ((not out) and (float(inline) != 0.0)):
        out = True
        wigfile.write("fixedStep chrom=" + chrom_name + " start=" + str(start_offset) + " step=1\n")
        wigfile.write(str(round(float(inline)/gen_mean, 4)) + "\n")

      else:
        wigfile.write(str(round(float(inline)/gen_mean, 4)) + "\n")

    cdfile.close()
    wigfile.close()

print "All done.\n"
```

FAIRE-seq analysis (continued)

F-Seq was run twice, overlapping peaks from both iterations were then overlapped with MACS peaks. As we had run two biological replicates of the FAIRE samples, we further overlapped the resulting intervals of the corresponding replicates.

For FAIRE-seq from the clinical samples FSeq iteration was repeated twice and resulting bed files were intersected as indicated in the text (category-wise or disease-stage wise) in order to produce consensus open chromatin regions.

The data are deposited in GEO: GSE73989.

Chromatin shape was assessed via FAIRE-seq normalized read counts within the FAIRE-seq peaks indicating chromatin open regions. This is a measure of the potency of the opening events within 1 Kb upstream annotated genes.

Gene-wise correlation between gene expression and local chromatin accessibility or local DNA methylation in prostate benign and cancer tissues samples (Figure 1)

Gaussian distribution deviation from a random distribution of correlative events between raw sequencing reads included within FAIRE-seq peaks or DNA methylated regions and transcript levels (measured as average reads counts) of gene transcription start sites (TSSs) were calculate across a number of intervals. This analysis was supported by the Kolmogorov-Smirnov statistic test shown in **Figure 1E-F**, in the main text. The test verified the statistical significance of the shift of the correlative events (in blue) from the random distribution (in red). A small shift of the Gaussian distributed curve toward the right indicated a correlation of local (as indicated by the intervals) open chromatin and increased gene expression. On the contrary, a small shift to the left indicated a correlation of local (as indicated by the intervals) DNA methylation with low gene expression. When correlation between gene expression and open chromatin upstream the TSS was considered, the Kolmogorov-Smirnov statistic test showed a marked deviation from the random correlation events. Up to a distance of 250kb upstream the TSS, this deviation grows and indicated an increased probability of encountering regulatory regions bound by transcription factors, before decreasing again 500 kb or 1Mb upstream the TSSs (**Figure 1F**). When, on the contrary, correlation between gene expression and local open chromatin around the TSS is considered, the Kolmogorov-Smirnov statistic test showed an indiscriminate deviation from the random correlation events (**Figure 1F**),

indicating the presence of confounding factors such as gene bodies or downstream genes that are actively transcribed. For genomic intervals bigger than 250kb, the shift returned, which once again, indicated the presence of confounding factors such as the presence of other genes (**Figure 1G**).

Reads distribution around peaks, motif enrichment, peaks distribution analysis, gene set enrichment analysis and gene ontology

Read distribution analysis around centre feature peaks was performed using a script(Hurtado et al., 2011) which generate a matrix of " normalized differences between coverage integrals in treated (e.g. FAIRE) versus control (e.g. Input) samples aligned reads" around a certain base pairs window (eg 5 kb). The normalization depends only on the dataset sizes computed as 10M/dataset_size. The script was modified to inspect reads distribution around windows larger than 5 kb.

To assess the presence of motifs of transcription factors (TFs) in the FAIRE-seq dataset, we looked for overrepresented TF motifs in open regions. Prediction of TF binding was performed using "findMotifsGenome.pl", and peaks distribution analysis with annotatePeaks.pl, both parts of the HOMER package(Heinz et al., 2010). Gene set enrichment analysis (GSEA) was performed by applying GSEA v2.07 (Broad Institute(Subramanian et al., 2005)) on a publicly available dataset of 150 AR target genes in castration resistant prostate cancer derived from Sharma et al.,(Sharma et al., 2013). This dataset was tested for enrichment in gene expression data from LNCaP and VcaP cells cultured in presence of JQ1. Gene ontology (GO) of up and down regulated genes upon treatment of the same cells was performed using GOrilla(Eden et al., 2009).

Evaluation of the 15-gene signature prognostic value

Microarray data from the Decipher GRID™ were extracted for three radical prostatectomy cohorts from previously described(Erho et al., 2013; Karnes et al., 2013; Ross et al., 2015) validation studies. Specimen selection, RNA extraction and microarray hybridization was performed for these samples in a Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory facility (GenomeDx Biosciences, San Diego, CA, USA) as previously described(Erho et al., 2013; Ross et al., 2015). Briefly, total RNA was extracted and purified using the Ovation WTA FFPE system (NuGen, San Carlos, CA). RNA was amplified and labeled using the Ovation WTA FFPE system (NuGen, San Carlos, CA) and hybridized to Human Exon 1.0 ST GeneChips (Affymetrix, Santa Clara, CA). After microarray quality control using the Affymetrix Power Toosl packages, probeset normalization was performed using the Single Channel Array Normalization (SCAN) algorithm(Piccolo et al., 2012). Affymetrix Core level summaries for annotated genes were used to summarize gene expression.

A classifier to distinguish between metastatic vs. non-metastatic cancers was developed. Metastatic progression after prostatectomy was defined as a positive CT scan or bone scan. The prognostic value of the signature was further assessed for biochemical recurrence and death from prostate cancer. The classifier was constructed using a generalized linear model with elastic net regularization as previously described(Erho et al., 2013).

Statistical analyses were performed in R, version 3.2.2, all statistical tests were two-sided using a 5% significance level, and the model was constructed using the glmenet package (glmnet_2.0-2). The value of lambda was determined using a 10-fold cross-validation in glmnet. The model was generated using the MCI cohort (GSE46691) as training data. In the final model 10 of the 15 genes contributed to the model score with 6 positively associated and 4 negatively associated genes as determined by the regularized coefficients. The final model output a continuous variable score ranging between 0 and 1 with higher scores indicating a higher probability of metastasis. Scores were then generated for samples from the Mayo Clinic II (MCII) and JHMI-RP validation cohorts and performance in each cohort was assessed using survival analysis. Kaplan-Meier curve p-values were generated with a weighted Cox regression model (R package survival_2.38-3).

Univariableand multivarible Cox proportional hazards analyses were performed using the cch method implemented in the survival package

Immunohistochemistry
In the Tampere TMA cohort, mouse anti-ATAD2 (HPA029424), anti-BRD2 (Bethyl Laboratories, Inc., Montgomery, TX, USA), and anti-BRD4 (Bethyl Laboratories, Inc., Montgomery, TX, USA) were used with Power Vision+ Poly-HRP IHC kit (ImmunoVision Technologies Co., Burlingame, California, USA) according to the manufacturer's instructions. The protocol has previously been described(Leinonen et al., 2010).
BRD2 and BRD4 were scored like in the human protein atlas project: the intensity of staining was scored semiquantitatively (0-3) plus the area percentage (0%, <25%, 25-75%, >75%). Both cytoplasmic and nuclear staining for ATAD2 were scored for intensity (0-3) and percentage (0-100). No digital imagae analysis was used.
The anti-ATAD2 antibody was also used for the Hamburg TMA cohort.
Freshly cut TMA sections were immunostained on one day and in one experiment. Slides were deparaffinized and exposed to heat-induced antigen retrieval for 5 minutes in an autoclave at 121°C in pH 7.8 Tris-EDTA-Citrate buffer. Primary antibody specific for ATAD2 (rabbit polyclonal antibody, Sigma, St. Louis, MO, USA; HPA029424; dilution 1:450) was applied at 37°C for 60 minutes. Bound antibody was then visualized using the EnVision Kit (Dako, Glostrup, Denmark) according to the manufacturer´s directions. Only nuclear ATAD2 staining was evaluated.

siRNA transfections
Silencer® selected siRNAs from Ambion (Applied biosystems/Ambion, Austin, Texas, USA) were used. Cells were reverse-transfected with lipofectamin RNAiMAX transfection reagent (Invitrogen, 13788) according to the manufacturer's protocol. Briefly, the cells were seeded and transfected with 20 nM of siBRD2 (s12070), 20 nM of siBRD4 (s23903), 20 nM of siATAD2 (s26393) or equal concentration of *Silencer*® negative control siRNA #1. Expression levels of *BRD2*, *BRD4* and *ATAD2* relative to *TBP* were measured by qRT-PCR (2.5 days after transfection) and protein levels by Western blot analysis (3 days after transfection).

Viability assays
Viability upon treatment of drugs or gene silencing was assessed using CellTiter-Glo reagent (G3581, Promega, Stockholm, Sweden) MTS assay according to manufacturer's instructions. Colorimetric changes were assessed using a PerkinElmer EnVision® Multilabel Plate Reader.

Fluorescent-based caspase cleavage assay
Appropriate amounts of LNCaP or VCaP cells were seeded in 384-well plates and allowed to attach for 48h at which point they received drug treatment (12 wells per condition). Induction of apoptosis was monitored using the CellPlayer 96-well Caspase-3/7 reagent (Essen Bioscience, 4440) at a final concentration of 1:5,000 on the Incucyte FLR instrument (Essen Bioscience). Phase contrast and fluorescence pictures were taken every two hours for a total of 96h. Analysis was performed using the inbuilt object counting algorithm.

Western Blot analysis
Cells were trypsinized and washed with cold PBS prior to resuspension in RIPA lysis buffer (30 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.5 % NP40, 0.1 % Na-Deoxycholate, 0.1 % SDS, pH 7.4) supplemented with protease inhibitors (Roche, 11873580001), rotated at 4 °C for 10 min and sonicated in a Bioruptor NextGen (Diagenode) at maximum power for ten cycles of 30 s ON, 30 s OFF to break nuclei and other cellular structures. Lysates were centrifuged for 10 min at 18,000 g and 4 °C and the supernatant transferred to a new tube. Protein concentration was determined using a BCA assay (Pierce, 23227) and equalized with RIPA buffer. Extracts were mixed with LDS NuPAGE buffer (Life technologies, NP0008) and Sample Reducing Agent (Life technologies, NP0009) and denatured for 10 min at 70 °C. Equal amounts were loaded onto 4-12 % gradient Bis-Tris NuPAGE gels (Life technologies, NP0323). Separated proteins were wet-blotted (25 mM Tris-Base, 192 mM glycine, 20 % methanol, 0.01 % SDS, pH 9.2) to methanol-activated 0.45 um PVDF

membranes (Millipore, IPVH00010) for 60 min at 30 V. Membranes were blocked in 5 % BSA (Sigma, A2153) in TBS with 0.1 % Tween-20 (Sigma, P5927) for 1 h prior to overnight incubation with appropriate concentrations of primary antibodies. The next day, membranes were washed with TBS with 0.1 % Tween-20 and incubated with appropriate secondary antibody for 1 h at room temperature. After washing with TBS with 0.1 % Tween-20, membranes were developed using the Novex ECL Reagent kit (Life technologies, WP20005) or a super-sensitive HRP substrate (Rockland, FEMTOMAX-110). Primary antibodies used were ATAD2 (Sigma, HPA029424), BRD2 (abcam, ab139690), MYC (Abcam, ab32072), BRD4 (Sigma, AV39076), KLK3 (Dako, D0487), AR (N-20 sc-816), b-actin-HRP (Cell Signaling, 5125) and GAPDH (Cell Signaling, 2118). Secondary HRP-conjugated anti-rabbit and anti-mouse were purchased from Dako (P0448 and P0447, respectively).

Supplemental References

Barfeld, S.J., Fazli, L., Persson, M., Marjavaara, L., Urbanucci, A., Kaukoniemi, K.M., Rennie, P.S., Ceder, Y., Chabes, A., Visakorpi, T., *et al.* (2015). Myc-dependent purine biosynthesis affects nucleolar stress and therapy response in prostate cancer. Oncotarget *6*, 12587-12602.

Boyle, A.P., Guinney, J., Crawford, G.E., and Furey, T.S. (2008). F-Seq: a feature density estimator for high-throughput sequence tags. Bioinformatics *24*, 2537-2538.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC bioinformatics *10*, 48.

El Gammal, A.T., Bruchmann, M., Zustin, J., Isbarn, H., Hellwinkel, O.J., Kollermann, J., Sauter, G., Simon, R., Wilczak, W., Schwarz, J., *et al.* (2010). Chromosome 8p deletions and 8q gains are associated with tumor progression and poor prognosis in prostate cancer. Clin Cancer Res *16*, 56-64.

Erho, N., Crisan, A., Vergara, I.A., Mitra, A.P., Ghadessi, M., Buerki, C., Bergstralh, E.J., Kollmeyer, T., Fink, S., Haddad, Z., *et al.* (2013). Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. PloS one *8*, e66855.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular cell *38*, 576-589.

Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D., and Carroll, J.S. (2011). FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat Genet *43*, 27-33.

Karnes, R.J., Bergstralh, E.J., Davicioni, E., Ghadessi, M., Buerki, C., Mitra, A.P., Crisan, A., Erho, N., Vergara, I.A., Lam, L.L., *et al.* (2013). Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk patient population. The Journal of urology *190*, 2047-2053.

Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M.J., Sauter, G., and Kallioniemi, O.P. (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nature medicine *4*, 844-847.

Kregel, S., Kiriluk, K.J., Rosen, A.M., Cai, Y., Reyes, E.E., Otto, K.B., Tom, W., Paner, G.P., Szmulewitz, R.Z., and Vander Griend, D.J. (2013). Sox2 Is an Androgen Receptor-Repressed Gene That Promotes Castration-Resistant Prostate Cancer. PloS one *8*, e53701.

Leinonen, K.A., Tolonen, T.T., Bracken, H., Stenman, U.H., Tammela, T.L., Saramaki, O.R., and Visakorpi, T. (2010). Association of SPINK1 expression and TMPRSS2:ERG fusion with prognosis in endocrine-treated prostate cancer. Clin Cancer Res *16*, 2845-2851.

Massie, C.E., Lynch, A., Ramos-Montoya, A., Boren, J., Stark, R., Fazli, L., Warren, A., Scott, H., Madhu, B., Sharma, N., *et al.* (2011). The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. EMBO J *30*, 2719-2733.

Minner, S., Enodien, M., Sirma, H., Luebke, A.M., Krohn, A., Mayer, P.S., Simon, R., Tennstedt, P., Muller, J., Scholz, L., *et al.* (2011). ERG status is unrelated to PSA recurrence in radically operated prostate cancer in the absence of antihormonal therapy. Clin Cancer Res *17*, 5878-5888.

Minner, S., Jessen, B., Stiedenroth, L., Burandt, E., Kollermann, J., Mirlacher, M., Erbersdobler, A., Eichelberg, C., Fisch, M., Brummendorf, T.H., *et al.* (2010). Low level HER2 overexpression is associated with rapid tumor cell proliferation and poor prognosis in prostate cancer. Clin Cancer Res *16*, 1553-1560.

Piccolo, S.R., Sun, Y., Campbell, J.D., Lenburg, M.E., Bild, A.H., and Johnson, W.E. (2012). A single-sample microarray normalization method to facilitate personalized-medicine workflows. Genomics *100*, 337-344.

Ross, A.E., Johnson, M.H., Yousefi, K., Davicioni, E., Netto, G.J., Marchionni, L., Fedor, H.L., Glavaris, S., Choeurng, V., Buerki, C., *et al.* (2015). Tissue-based Genomics Augments Post-

prostatectomy Risk Stratification in a Natural History Cohort of Intermediate- and High-Risk Men. European urology.
Schlomm, T., Sultmann, H., and Kollermann, J. (2008). [Identification and validation of clinically relevant molecular alterations in prostate cancer]. Der Urologe Ausg A *47*, 1193-1198.
Sharma, N.L., Massie, C.E., Ramos-Montoya, A., Zecchini, V., Scott, H.E., Lamb, A.D., Macarthur, S., Stark, R., Warren, A.Y., Mills, I.G., *et al.* (2013). The androgen receptor induces a distinct transcriptional program in castration-resistant prostate cancer in man. Cancer Cell *23*, 35-47.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America *102*, 15545-15550.
Urbanucci, A., Sahu, B., Seppala, J., Larjo, A., Latonen, L.M., Waltering, K.K., Tammela, T.L., Vessella, R.L., Lahdesmaki, H., Janne, O.A., *et al.* (2012). Overexpression of androgen receptor enhances the binding of the receptor to the chromatin in prostate cancer. Oncogene *31*, 2153-2163.
Urbanucci, A., Waltering, K.K., Suikki, H.E., Helenius, M.A., and Visakorpi, T. (2008). Androgen regulation of the androgen receptor coregulators. BMC Cancer *8*, 219.
Waltering, K.K., Helenius, M.A., Sahu, B., Manni, V., Linja, M.J., Janne, O.A., and Visakorpi, T. (2009). Increased expression of androgen receptor sensitizes prostate cancer cells to low levels of androgens. Cancer research *69*, 8141-8149.
Ylipaa, A., Kivinummi, K., Kohvakka, A., Annala, M., Latonen, L., Scaravilli, M., Kartasalo, K., Leppanen, S.P., Karakurt, S., Seppala, J., *et al.* (2015). Transcriptome sequencing reveals PCAT5 as a novel ERG-regulated long non-coding RNA in prostate cancer. Cancer Res.
Yu, J., Yu, J., Mani, R.-S., Cao, Q., Brenner, C.J., Cao, X., Wang, X., Wu, L., Li, J., Hu, M., *et al.* (2010). An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. Cancer Cell *17*, 443-454.
Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., Bernstein, B., Nusbaum, C., Myers, R., Brown, M., Li, W., *et al.* (2008). Model-based Analysis of ChIP-Seq (MACS). Genome biology *9*, R137.