SUPPLEMENTAL INFORMATION

Dual Role of FoxA1 in Androgen Receptor Binding to Chromatin, Androgen Signaling and Prostate Cancer

Biswajyoti Sahu[1], Marko Laakso[1,2], Kristian Ovaska[1,2], Tuomas Mirtti[3,4], Johan Lundin[4], Antti Rannikko[5], Anna Sankila[3], Juha-Pekka Turunen[3], Mikael Lundin[4], Juho Konsti[4], Tiina Vesterinen[4], Stig Nordling[3], Olli Kallioniemi[4], Sampsa Hautaniemi[1,2], and Olli A. Jänne[1,6]

[1]Institute of Biomedicine and [2]Research Programs Unit, Genome-Scale Biology, Biomedicum Helsinki, University of Helsinki, FI-00014 Helsinki, Finland, [3]Department of Pathology, Haartman Institute, University of Helsinki and HUSLAB, Helsinki University Central Hospital, FI-00014 Helsinki, Finland, [4]Institute for Molecular Medicine, University of Helsinki, FI-00014 Helsinki, Finland, [5]Department of Urology, Helsinki University Central Hospital, FI-00290 Helsinki, Finland, and [6]Department of Clinical Chemistry, Helsinki University Central Hospital, FI-00290 Helsinki, Finland

Supplementary Figure Legends

Figure S1. Loading of AR to *PSA* and *TMPRSS2* enhancers in parental and FoxA1-depleted LNCaP-1F5 cells in response to a 2-h exposure to different concentrations of 5α-dihydrotestosterone (DHT). Each point represents the mean value of two independent experiments.

Figure S2. Distribution of the distance between AR- and FoxA1-binding peaks in parental cells (A) and in the subgroup of ARBs pioneered by FoxA1 (B). The median distances were 43 and 41 nt (red line) in parental cell and FoxA1-pioneered ARBs, respectively.

Figure S3. Examples of the co-occupancy between AR- and FoxA1-binding sites around AR target genes. The number of reads for AR binding (left, in blue) and FoxA1 binding (right, in purple) is shown on the y-axes. Below each plot, there is a schematic description of the transcription unit in question, with the transcription start site and the direction of transcription being denoted by arrows, exons by black solid bars and introns by horizontal lines, respectively.

Figure S4. (A) LNCaP-1F5 cells were treated for 72 h with siRNA specific for FoxA1 mRNA (siFoxA1) or control siRNA (parental), after which total RNA samples were subjected to qPCR using primers specific for FoxA1 mRNA. 5α-Dihydrotestosterone (DHT) exposure time was 2 h. GAPDH mRNA was used for data normalization. (B) Total AR levels in parental and FoxA1-depleted LNCaP-1F5 and VCaP cells after a 2-h exposure to DHT. Immunoblotting was performed with anti-AR (Kang *et al*, 2004) and GAPDH (sc-47724, Santa Cruz) antibodies.

Figure S5. Directed ChIP validation of ARBs in the three different categories defined by FoxA1. The ChIP-qPCR assays were performed in parental (white bars) and siFoxA1 (black bars) cells after a 2-h exposure to 100 nM DHT (+) or vehicle (–). Mean + S.E.M. values are shown for duplicate samples. Primer pairs used for the selected sites are given in Supplementary Table S13.

Figure S6. Directed ChIP validation of FoxA1-independent and FoxA1-pioneered ARBs in parental (white bars) and FoxA1-depleted (black bars) VCaP cells. The cells were exposed for 2 h to 100 nM DHT (+) or vehicle (–). Mean + S.E.M. values are shown for duplicate samples. Primer pairs used for the selected sites are given in Supplementary Table S13.

Figure S7. Temporal and DHT dose-dependent accumulation of androgen-regulated mRNAs in the three classes defined by FoxA1. For each group, two mRNAs were selected as examples, and their accumulation was followed after exposure to 100 nM DHT for 8, 16, and 24 h in parental (yellow bars) or siFoxA1 (red bars) cells (upper panels). Accumulation of the same mRNAs in response to 0–100 nM DHT was determined at 24 h (lower panel). (**A**) FoxA1-independent genes; (**B**) FoxA1-pioneered genes, and (**C**) new genes in siFoxA1 cells. Mean + S.E.M. values are shown for three biological replicate experiments. Transcript levels were determined by qPCR using the primers listed Supplemental information, Table S11.

Figure S8. Examples of AR target gene expression in parental and siFoxA1 cells after a 24-h exposure to 100 nM DHT for the three classes of genes defined by FoxA1. Transcript

levels were determined by qPCR using the primers listed Supplemental information, Table S11. The results are shown as relative mRNA levels (left panels) and as fold changes (right panels). (A) Androgen regulation of genes that were independent of FoxA1. (B) Androgen regulation of genes that required the presence of FoxA1. (C) Examples of genes that became androgen-responsive in siFoxA1 cells. Mean ± SEM values are shown for three biological replicates.

Figure S9. AR- and FoxA1-binding sites together with H3K4me2 marks shown as raw tag counts from ChIP-seq in both parental and siFoxA1 cells for the *kallikrein* cluster located on human chromosome 19.

Figure S10. (A) Disease-specific survival of prostate cancer patients according to the intensity of AR staining in the primary tumor. $\chi 2 = 4.78$, $P = 0.03$ (log-rank test). (B) Expression of FoxA1 mRNA in normal prostate tissue and in prostate adenocarcinoma specimens. The data are adapted from Kilpinen *et al,* 2008.

Supplementary Tables S1, S6–S11

Table S1. List of the top *cis*-elements enriched in the AR cistrome of the parental LNCaP-1F5 cells.

Motif	Enrichment Ratio	p-value
AR	9.2338	1.80E-72
FoxA1	2.6169	5.06E-244
FoxA2	2.4796	2.53E-237
FoxF2	2.4091	3.40E-62
E2F1	2.1628	4.60E-05
Tal1/GATA1	2.0978	8.90E-10
STAT1	2.0541	6.73E-07
STAT3	1.7763	3.72E-24
GABPA	1.7693	4.08E-11
FoxD1	1.715	2.81E-203

Table S6. Pathway enrichment analysis in LNCaP-1F5 cells

Pathway analysis was performed for enriched gene sets in parental and FoxA1-depleted cells using WebGestalt (http://bioinfo.vanderbilt.edu/webgestalt/). For the enriched gene sets, ratio of enrichment (R) and p-value from hypergeometric test was calculated and p-value adjusted by multiple test adjustment (adjP). Enriched gene sets with an adjP <0.002 are included in the table.

Pathway	Genes	Enrichment Ratio	adjP	Regulation*
FoxA1-independent				
Metabolic pathways	13	5.18	2.94e-05	Up
Biosynthesis of unsaturated fatty acids	3	60.03	0.0004	Up
Prostate cancer	4	14.84	0.005	Up
Synaptic transmission	2	34.88	0.009	Down
FoxA1-pioneered				
Metabolic pathways	20	5.66	1.53e-08	Up
Insulin signaling	7	13.68	1.56-05	Up
Adipocytokine signaling pathway	3	14	0.001	Up
Metabolism of lipids and lipoproteins	7	8.86	0.001	Up
Prostaglandin synthesis and regulation	3	30.26	0.001	Up
Peroxisome lipid metabolism	3	39.09	0.004	Up
Wnt signaling pathway	4	19.17	0.0006	Down
Starch and sucrose metabolism	3	30.42	0.0016	Down
Myometrial relaxation and contraction	4	13.52	0.002	Down
Unique to siFoxA1 cells				
IFN-gamma pathway	14	10.49	1.65e-08	Up
Pathways in cancer	12	9.31	5.85e-07	Up
Metabolic pathways	19	4.41	5.51e-06	Up
Focal adhesion	9	11.47	7.77e-06	Up
Regulation of SMAD2/3 signaling	11	10.63	1.54e-06	Up
TGFBR signaling	11	10.63	1.54e-06	Up
Glypican pathway	13	7.64	3.48e-06	Up
IL-6 mediated signaling	7	34.15	3.91e-06	Up
Adherens junction	6	19.96	4.29e-05	Up
Neutrophin signaling	6	12.20	0.0008	Up
Wnt signaling pathway	5	21.35	0.0002	Up
MAPK signaling pathway	7	11.07	0.0002	Up
Erbb signaling pathway	5	14.72	0.001	Up
IGF1 pathway	4	34.15	0.001	Up
Metabolic pathways	14	3.27	0.003	Down
Synthesis of bile acids and bile salts	4	59.45	0.0024	Down
AR signaling	4	8.96	0.002	Down
IFN-gamma pathway	8	6.03	0.009	Down

*Direction of regulation refers to the up- or down-regulated gene sets used for pathway analysis and not to the overall direction in pathway regulation.

.

Table S7. Pathway enrichment analysis for LNCaP-1F5 and VCaP cells

Pathway analysis was performed for enriched gene sets for ChIP-seq mapped genes in parental and FoxA1-depleted cells common to LNCaP-1F5 and VCaP cells.

Pathway	Genes	Enrichment Ratio	P-value
New sites common to LNCaP-1F5 and VCaP cells under FoxA1 depletion conditions (siFoxA1)			
Insulin signaling pathway	8	7.53	6.02e-05
Neurotrophin signaling pathway	5	5.12	0.014
Chronic myeloid leukemia	4	6.78	0.019
Pathways in cancer	7	2.71	0.036
Type II diabetes mellitus	3	8.11	0.049
FoxA1-independent in LNCaP-1F5 and VCaP cells			
Biosynthesis of unsaturated fatty acids	4	17.77	0.001
Aldosterone-regulated sodium reabsorption	5	6.51	0.006
Apoptosis	5	5.62	0.01
Jak-STAT signaling pathway	6	3.78	0.01

Table S8. Gene ontology enrichment analysis for LNCaP-1F5 and VCaP cells

Gene ontology (GO) analysis was performed for enriched gene sets for ChIP-seq mapped genes in parental and FoxA1-depleted cells common to LNCaP-1F5 and VCaP cells.

GO Term	Genes	Enrichment Ratio	P-value
New sites common to LNCaP-1F5 and VCaP cells under FoxA1 depletion conditions (siFoxA1)			
GO:0009725~response to hormone stimulus	12	9.02	4.63e-08
GO:0032870~cellular response to hormone/peptide hormone stimulus	7	14.53	1.26e-07
GO:0008286~insulin receptor signaling pathway	4	29.84	2.99e-04
GO:0007242~intracellular signaling cascade	14	3.07	3.13e-04
GO:0042127~regulation of cell proliferation	10	3.50	0.0015
GO:0007167~enzyme linked receptor protein signaling pathway	6	4.84	0.007
GO:0051094~positive regulation of developmental process	5	4.96	0.016
GO:0030334~regulation of cell migration	4	6.53	0.021
FoxA1-independent in LNCaP-1F5 and VCaP cells			
GO:0030521~androgen receptor signaling pathway	3	21.67	0.008
GO:0030522~intracellular receptor-mediated signaling pathway	3	10.40	0.032
GO:0008610~lipid biosynthetic process	5	4.02	0.033
GO:0008593~regulation of Notch signaling pathway	2	47.30	0.040
GO:0006508~proteolysis	9	2.22	0.042
GO:0006916~anti-apoptosis	4	5.05	0.042

Table S9. List of *cis*-elements enriched in the vicinity of the three ARB categories shown in Figs. 1 and 2. In each case, the ten most enriched *cis*-elements are shown.

(A) AR-binding sites independent of FoxA1.

Motif	**Enrichment Ratio**	**p-value**
AR	19.8956	78.53E-176
TLX1::NFIC	4.0920	35.35E-12
ESR2	2.8879	3.19E-08
FOXA1	2.2878	1.19E-272
TAL1::GATA1	2.1914	8.15E-16
GABPA	2.0527	3.29E-26
FOXF2	2.00	6.45E-52
STAT3	1.9964	4.18E-46
FOXA2	1.8807	1.459E-208
ELK4	1.8534	7.178E-06

(B) AR-binding sites pioneered by FoxA1.

Motif	**Enrichment Ratio**	**p-value**
AR	4.0404	5.03E-16
FOXA1::AR*	3.8718	1.63E-114
FOXA1	3.7707	0.00E00
FOXA2	3.7237	0.00E00
FOXF2	3.3948	9.57E-124
FOXD1	2.2243	0.00E00
FOXQ1	2.0888	36.32E-162
FOXI1	1.7994	69.35E-224
TLX1::NFIC	1.7994	1.656E-02
FOXO3	1.4102	40.99E-270

(C) New AR-binding sites in siFoxA1 cells.

Motif	**Enrichment Ratio**	**p-value**
AR	13.9930	0.00E00
TLX1::NFIC	4.8639	4.709E-76
ESR1	4.0552	1.053E-04
CTCF	2.7758	1.315E-22
ESR2	2.7498	57.25E-58
GABPA	2.4615	14.59E-222
ELK4	2.1605	26.88E-54
STAT3	2.1004	3.21E-314
TCFCP2I1	2.0065	7.07E-114
PPARG::RXRA	1.9990	2.844E-34
TAL1::GATA1	1.9884	17.43E-70

*The unique *cis*-element shown in Figs. 3D and 3G.

Table S10. Demographics and clinicopathological variables of the prostate cancer patients.

Variable	**Total Sample (n=350)**	
	n	Median (range)
Age at diagnosis, years	350	63 (44–83)
Follow-up time, years	350	13.3 (11.33–25.0)
Preoperative PSA, μg/l	278	10.1 (0–120)
Tumor area*	342	15% (0.1–100)
Gleason grade		
4	1 (0.3)	
5	9 (2.6)	
6	76 (21.7)	
7	207 (59.1)	
8	44 (12.6)	
9	13 (3.7)	
pT		
2a	31 (9.5)	
2b	2 (0.6)	
2c	160 (49.1)	
3a	77 (23.6)	
3b	55 (16.9)	
3c	1 (0.3)	
NA	24	
Positive surgical margins		
Yes	63 (22.1)	
No	222 (77.9)	
NA	65	
Lymph node involvement		
Yes	10 (2.9)	
No	334 (97.1)	
NA	6	
FoxA1 nuclear stained area	350	95% (0–100)
FoxA1 nuclear staining intensity– Maximum score		
negative	1 (0.3)	
weak	23 (6.6)	
moderate	71 (20.3)	
strong	255 (72.9)	

*Tumor area is determined as percentage of tumor in all histological sections of each patient; NA = not available.

Table S11. Univariate and multivariate Cox proportional hazards regression analysis of tumor characteristics in a series of patients (n=350) with prostate cancer.

		Univariate		Multivariate	
Characteristic	No. (%)	HR (CI 95%)	*P*	HR (CI 95%)	*P*
FoxA1 ratio positive nuclei					
<=90 %	111 (32)	1			
> 90 %	239 (68)	1.80 (0.79–4.14)	0.16		
FoxA1 staining intensity					
Neg-Mod	95 (27)	1		1	
Strong	255 (73)	2.89 (1.02–8.21)	0.04	1.72 (0.59–5.04)	0.32
Gleason grade					
4-7	293 (84)	1		1	
8-9	57 (16)	3.04 (1.50–6.14)	0.005	2.38 (1.07–5.30)	0.03
Margins positive					
No	222 (78)	1			
Yes	63 (22)	0.64 (0.14–2.89)	0.56		
pT					
pT2	193 (59)	1		1	
pT3	133 (41)	20.3 (4.79–85.8)	<0.0001	16.0 (3.72–69.3)	<0.0001
Lymph nodes					
Neg	344 (97)	1		1	
Pos	10 (3)	4.24 (1.45–12.4)	0.008	2.88 (0.81–10.2)	0.10

Table S12. Pair-wise associations between FoxA1 expression and tumor characteristics (Pearson's product moment correlation) in a series of patients (n=350) with prostate cancer.

Characteristic	FoxA1 % positive nuclei	FoxA1 staining intensity	Gleason grade	pT	Nodal status	Margins positive	Preoperative PSA
FoxA1 % positive nuclei	1						
FoxA1 staining intensity	0.67***	1					
Gleason grade	0.19**	0.20***	1				
pT	0.15	0.12	0.40***				
Nodal status	0.07	0.07	0.10	0.18*	1		
Margins positive	0.00	0.00	0.23**	0.26**	0.10	1	
Preoperative PSA	0.00	-0.03	0.16	0.29***	0.03	0.37***	1

*** $p<0.0001$
** $p<0.001$
* $p<0.01$

Table S13. Nuclear FOXO1 and FOXO3 staining intensity in prostate cancer tissue specimens containing low-moderate or high FoxA1 staining intensity. AR staining intensity was high in all samples*.

	Nuclear FOXO1 staining	
	Low-moderate FoxA1/High AR	High FoxA1/High AR
Strong	6	4
Moderate	3	12
Weak	1	14

	Nuclear FOXO3 staining	
	Low-moderate FoxA1/High AR	High FoxA1/High AR
Strong	7	10
Moderate	3	12
Weak	0	8

*Three TMA slides containing 10 low-moderate FoxA1/high AR spots and 30 high FoxA1/high AR spots were evaluated for staining with FOXO1 and FOXO3 antibodies. The distribution of staining intensities between the low-moderate FoxA1 and high FoxA1 samples is significantly different ($p<0.05$).

Table S14. Sequences of the siRNAs used to deplete FoxA1.

siRNA sequences	**Sequences (5'→3')**
siFoxA1 (ON-Targetplus SMARTpool sequences from Dharmacon)	1) GCACUGCAAUACUCGCCUU 2) CCUCGGAGCAGCAGCAUAA 3) GAACAGCUACUACGCAGAC 4) CCUAAACACUUCCUAGCUC
Non-targeting pool	1) UGGUUUACAUGUCGACUAA 2) UGGUUUACAUGUUGUGUGA 3) UGGUUUACAUGUUUUCUGA 4) UGGUUUACAUGUUUUCCUA

Table S15. ChIP-sequencing aligned read statistics.

Sample	Yield (kbases)	% PF clusters	% Align
AR (rep 1)	626712	81.8	73.7
AR (rep 2)	628314	80.08	73.9
FoxA1 (rep 1)	701234	78.36	74.89
FoxA1 (rep 2)	708246	77.97	75.19
AR-siFoxA1 (rep 1)	758533	74.56	73.89
AR-siFoxA1 (rep 2)	753454	74.52	74.39
FoxA1-siFoxA1 (rep 1)	542030	74.5	77.02
FoxA1-siFoxA1 (rep 2)	552898	72.84	76.38
H3K4me2	894045	83.01	78.57
H3K4me2-siFoxA1	847641	85.38	78.53
DHS	829000	84.29	63.35
DHS-siFoxA1	907000	82.53	60.85
rIgG	743283	80.59	75.27
GR	804000	82.05	76.63
GR-siFoxA1	987000	77.30	76.26
VCaP-AR	1392000	84.04	79.46
VCaP-AR-siFoxA1	1436000	79.46	73.10
VCaP-rIgG	1433000	80.28	75.51

Table S16. Sequences of the primers used for ChIP assays and qRT-PCR.

Primer name	Primer sequence (5'→3')
PSA enhancer+ (Wang et al. 2005)	TGGGACAACTTGCAAACCTG
PSA enhancer- (Wang et al. 2005)	CCAGAGTAGGTCTGTTTTCAATCCA
FoxA1- Independent	
AR5695+	CCATCTTTCATTCCCGATTTT
AR5695-	ACAGGGCTGTTCCGTCAA
AR3752+	TGTCCCATGAGTGACCTCAA
AR3752-	ACTGCTCCAGGAAAACATGG
AR2823+	TGGATTTGTCCATCTGTCCA
AR2823-	GGCAATGACCCAGGATTTC
AR3174+	ACCAGCGTGCCTTTTATCAC
AR3174-	CTCGATTTCCTCCAATCCAA
AR5138+	AGGGAGGATGTGGTGAGAACT
AR5138-	AAGCGATCCGTCTACCAATG
AR5327+	CAGCCAGCATTTTGTGAAAGT
AR5327-	GGCAGAGAGCCAGAATCAAC
AR5328+	CCACAGGGAGAGAGTGGGTA
AR5328-	CCATCGGTTCATAGTGTGCTT
AR4310+	TTCCCGCTGAATACACACAC
AR4310-	GCATGAACAGAGAGTGAAGCTG
VCaP-I2+	AGGGAGGATGTGGTGAGAACT
VCaP-I2-	AAGCGATCCGTCTACCAATG
VCaP-I3+	CCACAGGGAGAGAGTGGGTA
VCaP-I3-	CCATCGGTTCATAGTGTGCTT
VCaP-I4+	GGCTGCTGCAGGATAAAGG
VCaP-I4-	ATGGAGTGCAGGTTGGAGAG
VCaP-I5+	GCAGCAGCCTGGAGTTTTAC
VCaP-I5-	AAATGAGAGAGTGGGGAAAGC
VCaP-I6+	TCATGCACAGAAGGAAGTGG
VCaP-I6-	CGTCCAGTAGCAGCCTTTTC
VCaP-I7+	CATCAGGGGTAGAGGGTTCA
VCaP-I7-	GGCCAAATCAATGTTTCCAC
VCaP-I8+	AAGGCAGGAGATGGGACAG
VCaP-I8-	ACCCTGCTGCTGGCTAGTT

VCaP-I9+	ACGTTGGGTGTCCAAAACAT
VCaP-I9-	AGGATTGTGCCACACCTGAT
FoxA1-Pioneered	
AR2827+	GGGAGACAGATTTTCCCTGAG
AR2827-	CGGGCATACCTCTTGTGAA
AR5569+	TGCCATCAGAAGAATCCACA
AR5569-	ACCAACATTCTGGGACACGTA
AR5144+	CACACTATGATGGCCTGTGG
AR5144-	TGTGTGGCTGTGTTCCAATAA
AR3757+	TAAAGGAAGGGGGAGAAAGG
AR3757-	TGGCCCTGAAGATGAATAGG
AR1177+	TGTTGATGGGATGCTTGTAGTC
AR1177-	GAAAAGCCCTCAGTGTGTCTCT
AR2830+	TTCTCCTCTTGCAGTGTTTTCA
AR2830-	CCCCCATCAGAAATGCTCT
AR10+	AAATTTGCCAGGCTGGTTC
AR10-	GGTGTCCGAAATCCACTGTAA
AR5141+	TTTCCCCGACAGATTTTTAGAG
AR5141-	GGGTAAACAAATTGGGACACA
VCaP-U1+	AGTTGCTATGGGGCCAGTT
VCaP-U1-	TACTGCCAGCACGTCTTCTG
VCaP-U2+	CCACACAAACTTCTGGCTCA
VCaP-U2-	GTATGCTGCCCACTTGGTTT
VCaP-U3+	GGACGTGACAGATTCCTCAAA
VCaP-U3-	CCCTTGCTATTGTTTGCTTTTC
VCaP-U4+	CGTATGCCTTCCCACAATG
VCaP-U4-	TGCCAGAGGTACGTGTGTTC
VCaP-U5+	GGAGGACATCAGGGATACCA
VCaP-U5-	GGGAGCGTAAGAACCACCTT
VCaP-U6+	AACATCCAAAGAAATGCTTCGT
VCaP-U6-	TGTCCTGTGCAAGCAACAAT
VCaP-U7+	TCCTGACTTCACCCCTGACT
VCaP-U7-	TATGCAGTGGCAAGGAAACA
VCaP-U8+	CAGCTCACCTGCACATAGTCA
VCaP-U8-	AGAAAGCAAGGGAAGAAATGC

VCaP-U9+	CGGCAAATATTGATTTGAGGA
VCaP-U9-	TGGCACCTATTCAGCACAGA
New sites in siFoxA1	
AR11836+	CCTGACCTCAGCTCTCCTTTT
AR11836-	TTCAGGGAGCAGGAAGTTTG
AR2002+	ATTCAGACCTGGCATGAACC
AR2002-	GAGGGCTCTTCTTGGACAGA
AR8720+	CTGGGGCCTGAAACCTTAG
AR8720-	GTGGGGACGTCATGTTCTG
AR10310+	TGTGTGACCACAGTACCCTGA
AR10310-	ATCTGCCTGGCAAAGGTCT
AR3233+	TGTCTGACAGTGAGGCAAGG
AR3233-	GCTTTCCCTGTTCTGATCCA
AR15327+	ATCTGCCACTCCCCACTTC
AR15327-	TAGAGTCTCCCGGGAACAAA
AR10313+	GCCCTGTAAAGCACAGTTCAG
AR10313-	GTTCCCAGTGAAGCCACAGT
AR5+	GCTAAACGACCAGGTCAAGC
AR5-	CTTCATGCCAGAGTGTTCCA
VCaP-N1+	AATGCCCAGATGGAAAACAG
VCaP-N1-	GCTGTTCCTAGGGTGCTGAG
VCaP-N2+	TCGTTCTCTCCCTCCCACTA
VCaP-N2-	TCTCCTCAGCTCTTTGGTCAG
VCaP-N3+	GGCATTCCAGCAGAGTGTG
VCaP-N3-	TTTGCACCACAGAAAGCAAG
VCaP-N4+	GCAGGGAGGAGCATGAAA
VCaP-N4-	TGTACTGGCGAGGAGAGAGAG
VCaP-N5+	GACAGCAAGGGACAAAAGGA
VCaP-N5-	CTTCAGCTGTGGTGCAAGAA
VCaP-N6+	CTAATTGGTCACAGCCAGCA
VCaP-N6-	CTCCGCCATTTTGCTTTTTA
VCaP-N7+	CTAGGGTCCTAGCGATGTGG
VCaP-N7-	GAATGCAAGACCTGACAGCA
VCaP-N8+	TCTTTCCTGCCTTCCATCAC
VCaP-N8-	AGTCACTGGCTGTCCTCTCAG

VCaP-N9+	CAGCTGTGAGGACAGAGGGTA
VCaP-N9-	GTCGTACGGTTGGGACAGTAG
mRNA qRT-PCR primer name	
PSA mRNA+ (Wang et al. 2007)	TGTGTGCTGGACGCTGGA
PSA mRNA- (Wang et al. 2007)	CACTGCCCCATGACGTGAT
FoxA1 mRNA+	GTGGCTCCAGGATGTTAGGA
FoxA1 mRNA-	GAGTAGGCCTCCTGCGTGT
NFKBIA mRNA+	GGGACTCGTTCCTGCACTT
NFKBIA mRNA-	GTCTGCTGCAGGTTGTTCTG
SPDEF mRNA+	AAGTGCTCAAGGACATCGAGA
SPDEF mRNA-	AGGAGCCACTTCTGCACATT
LPAR3 mRNA+	CTCATGGCCTTCCTCATCAT
LPAR3 mRNA-	TACCACAAACGCCCCTAAGA
LRIG1 mRNA+	CCGAACCTACAGGAAGTGTACC
LRIG1 mRNA-	TGCGAATCTTGTTGTGCTG
EXTL2 mRNA+	CCTGAACTGGAAACCAATGC
EXTL2 mRNA-	TCAGGAAATTGCTGCCAAA
AFF3 mRNA+	GATGCAGAGCCAGAGAGTCC
AFF3 mRNA-	GCCTGCTGTTCATTCTCCTC
ETS2 mRNA+	GGATTCCATTTCTCATGACTCC
ETS2 mRNA-	CAGGGGTTCTTTGGAATGC
EDN2 mRNA+	TGGGTGAACACTCCTGAACA
EDN2 mRNA-	AGGCAGAAGGTGGCACAG
FOXO1 mRNA+	CGGGCTGGAAGAATTCAA
FOXO1 mRNA-	GTTGTTGTCCATGGATGCAG
CITED1 mRNA+	CTCACCTGCGAAGGAGGAT
CITED1 mRNA-	TGGTTCCATTTGAGGCTACC

Extended Experimental Procedures

Chromatin Immunoprecipitation (ChIP)

Five million LNCaP-1F5 cells were cultured on 150-mm dishes in RPMI-1640 medium containing 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 10 mM HEPES supplemented with antibiotics (penicillin and streptomycin) for 4 days and then exposed to 100 nM DHT (or 100 nM dexamethasone) for 2 h. Cells were fixed in 1% formaldehyde (Merck KGaA, Darmstadt, Germany) for 10 min at room temperature and washed twice with ice-cold PBS. Cells were scraped in lysis buffer (5 mM PIPES, pH 8.0, 85 mM KCl, and 0.5% NP-40) containing 2 x protease inhibitors (Roche Inc., Mannheim, Germany).

Antibodies specific for the protein of interest were coupled to Dynal protein-G magnetic beads (Invitrogen) according to the manufacturer's instructions. For each immunoprecipitation (IP), 100 µl of sonicated chromatin was diluted 1:10 with 900 µl of RIPA buffer containing PIC and 10% was stored as input fraction. To the remaining, 100 µl of antibody-coupled magnetic beads were added and incubated on a rotator overnight at 4°C. After incubation, beads were washed 5 times with LiCl wash buffer (100 mM Tris-HCl, pH 7.5, 500 mM LiCl, 1% NP-40, and 1% sodium deoxycholate) and followed by two washes with 10 mM Tris-HCl (pH 8.0) containing 1 mM EDTA. Chromatin-antibody samples were eluted from beads by incubating for 1 h at 65°C in IP elution buffer (1% SDS, 0.1 M $NaHCO_3$), followed by overnight incubation at 65°C to reverse cross-linking. DNA was purified using QIAquick PCR purification kit (Qiagen Inc., California, USA) and eluted in 100 µl of elution buffer.

ChIP-sequencing (ChIP-seq)

ChIP experiments were carried out using the following antibodies: AR (Kang et al. 2004), FoxA1 (ab23738, Abcam), H3K4me2 (07-030, Millipore), CTCF (ab70303, Abcam), GR (Widen et al., 2000), normal rabbit IgG (sc-2027, Santa Cruz), and normal mouse IgG (sc-2025, Santa Cruz). Immunoprecipitated samples together with input samples were processed for library preparation according to Illumina's instructions. In brief, DNA samples were blunt-ended and ligated to sequencing adapters. Adapter-ligated DNA fragments (size range 150–300 bp) were excised from agarose gel and purified with Qiagen gel extraction kit. Isolated DNA was amplified by PCR (15 cycles), purified, and DNA library was sequenced on Illumina Genome Analyzer II.

The ChIP-seq reads were filtered using the default Illumina chastity filter during the base-calling process and used ELAND (Illumina) to align 30-bp reads to human genome (hg19), allowing up to two mismatches per read. All ChIP-seq experiments were performed in biological duplicates, and only peaks present in both samples were used for downstream analysis.

DNaseI-hypersensitivity site mapping by sequencing (DHS-seq)

The DNaseI-hypersensitivity assay and deep sequencing was performed as described by Song and Crawford (2010). In brief, 30 x 10^6 cells were centrifuged at 900 rpm for 5 min and cells pellets washed twice with ice cold PBS. Cells were lysed with RSB buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, and 3 mM $MgCl_2$) with 0.1% NP-40 and nuclei were collected by centrifugation at 500 rpm for 10 min. Intact nuclei were treated with different concentrations (0–12 U) of DNaseI (Roche) for 15 min at 37°C, and reactions were stopped with 0.1 M EDTA. DNaseI-digested high-molecular-weight DNA was embedded in 1.0% InCert (Lonza) low-melt agarose gel in 50 mM EDTA. To remove protein, DNA plugs were washed twice with LIDS buffer (1% lauryl sulfate, 10 mM Tris-HCl, and 100 mM

EDTA) at room temperature, then incubated overnight at 37°C and subsequently washed five times with 50 mM EDTA, 1 h each, by shaking at 60 rpm at room temperature. The DNaseI-digested plugs were run through pulsed field gel electrophoresis system to identify optimal DNaseI digestion. Optimal concentrations of DNaseI generated a smear of high-molecular-weight fragments (>100 kb) when analyzed by pulsed field gel electrophoresis. DNaseI-digested ends were blunt ended in gel with T4 DNA polymerase, melted at 65°C, and purified by phenol extraction and ethanol precipitation followed by library construction for deep sequencing.

The annealed biotinylated linkers containing an *MmeI* restriction site (TCCGAC) at the 3′ end were attached to the blunt-ended DNaseI-digested ends followed by digestion with *MmeI*. *MmeI* cuts 20 bp into the sequence adjacent to the DNaseI hypersensitive site, leaving a 2-bp overhang. After purification of DNaseI-digested ends on streptavidin beads, a second set of phosphorylated annealed linkers containing a two-base degenerate overhang was ligated to the dephosphorylated *MmeI* ends. Inserts, along with the linkers were amplified by PCR, and the inserts were purified by PAGE to minimize contamination by adaptors. The insert + linker band was cut and purified by phenol extraction and ethanol precipitation. The purified DNA was sequenced at the BMGen core facility with custom sequencing primers. The linker and sequencing primer sequences are listed in the original publication by Song and Crawford, Cold Spring Harbor Protocols, 2010.

The sequencing reads for DHS-seq were filtered using the default Illumina chastity filter during the base-calling process and used ELAND (Illumina) to align 20-bp reads to human genome (hg19), allowing up to two mismatches per read. The clustering of sequence tags and the DHS peaks were analyzed using F-seq algorithm (Boyle et al., 2008b). F-seq uses kernel density estimation that provides both discrete and continuous probability landscape across the whole genome. These kernel density estimation based probabilities are

computed at each base and are directly proportional to the probability of seeing a sequence read at that location.

Bioinformatics Analyses

Pre-defined motifs of transcription factors were obtained from JASPAR (Portales-Casamar et al. 2010). Over-representation of a motif in ChIP-seq binding sites was calculated against randomly selected regions from the same chromosomes and of the same length as the actual binding sites. The significance of the over- or under-representation was addressed by comparing the binding site frequencies between the groups using χ^2-test (Wang *et al,* 2007). *De novo* motif analysis was carried out using MEME (Bailey and Elkan, 1994; Bailey, 2002). The details of all the analyses are provided in the Supplementary information.

Evaluation of FoxA1 and AR staining intensity in tissues specimens

The slides were scanned with an automated whole-slide scanner (Mirax Scan, Zeiss, Göttingen, Germany), using a 20 x objective and a DFW-X710 camera (Sony, Tokyo, Japan). The pixel resolution was 0.26 μm per pixel. The virtual slides were compressed to a wavelet file format (Enhanced Compressed Wavelet, ECW, ER Mapper, Erdas Inc, Atlanta, Georgia) with a compression ratio of 1:5. Staining intensity was evaluated using virtual microscopy software (WebMicroscope, HUCH Clin Res Institute Ltd, Helsinki) without prior knowledge of the Gleason grades. Nuclear staining intensity was scored as being negative, weak, moderate or strong. Proportion of positive nuclei was estimated in semiquantitative manner in five percent-unit intervals. In the survival analysis, a cut-off value of 90% of the maximum score was considered as threshold for nuclear positivity. Automated immunohistochemistry scoring was performed as previously described (Konsti

et al, 2011). The separation of different color channels was achieved by a built-in vector for Fast Red in the color deconvolution plugin.

Additional references

Bailey TL (2002) Discovering novel sequence motifs with MEME. *Curr Protoc Bioinformatics.* Chapter *2*, Unit 2 4

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second International Conference on Intelligent Sytems for Molecular Biology*, pp. 28–36

Boyle AP, Guinney J, Crawford GE, Furey TS (2008) F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* **24:** 2537–2538

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M*,* Mpindi JP, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S, Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9:** R139

Konsti J, Lundin M, Joensuu H, Lehtimäki T, Sihto H, Holli K, Turpeenniemi-Hujanen T, Kataja V, Sailas L, Isola J, Lundin J (2011) Development and evaluation of a virtual microscopy application for automated assessment of Ki-67 expression in breast cancer. *BMC Clin Pathol* **11:** 3

Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, Valen E, Yusuf D, Lenhard B, Wasserman WW, Sandelin A (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucl Acids Res* **38:** D105–110

Figure S1

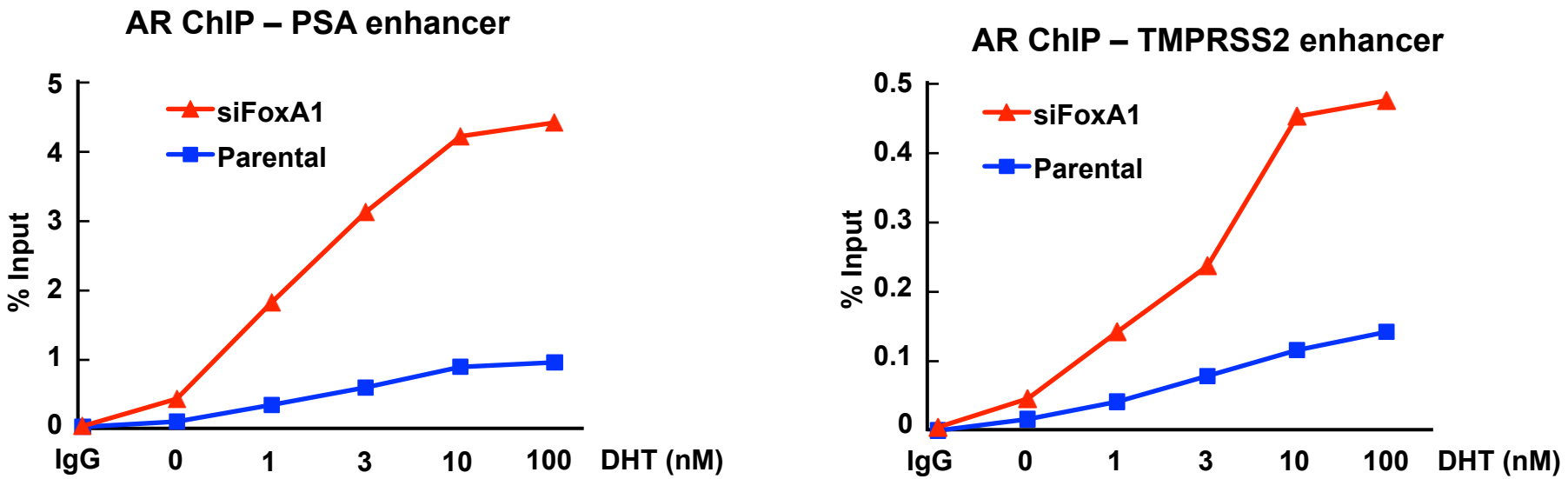

AR ChIP – PSA enhancer
% Input
5
4
3
2
1
0
siFoxA1
Parental
IgG
0
1
3
10
100
DHT (nM)
AR ChIP – TMPRSS2 enhancer
% Input
0.5
0.4
0.3
0.2
0.1
0
siFoxA1
Parental
IgG
0
1
3
10
100
DHT (nM)

Figure S2

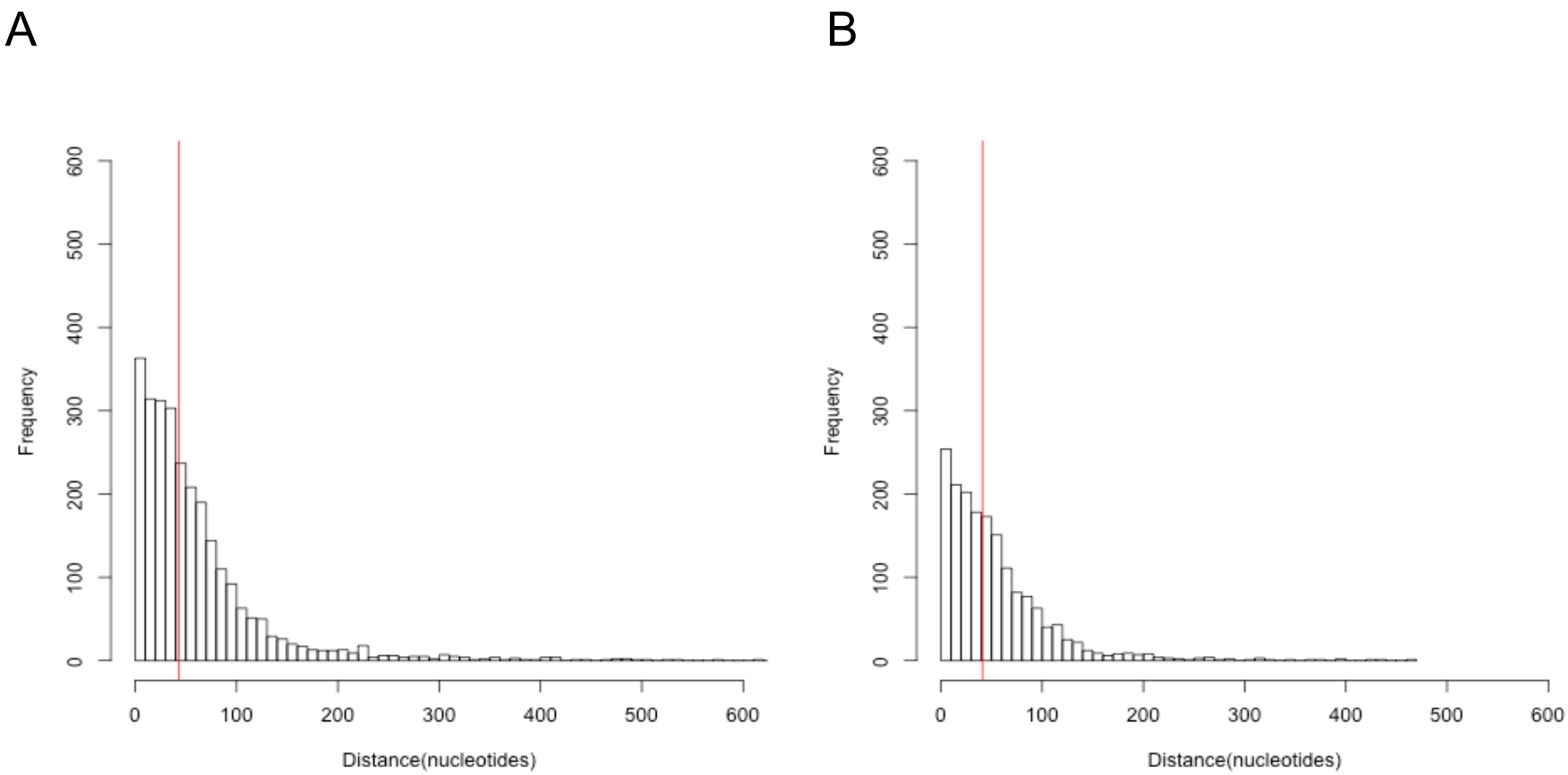


Figure S3

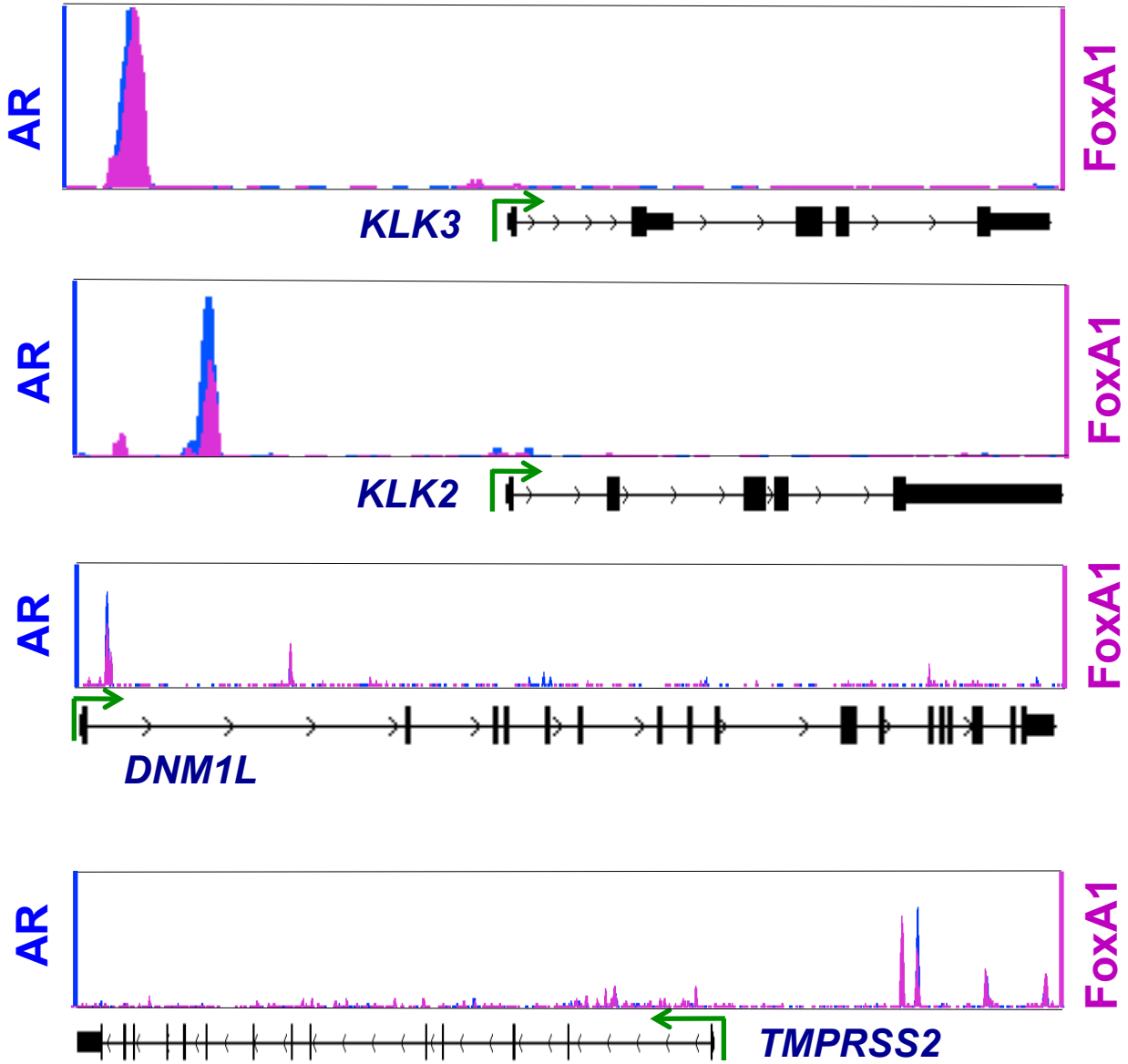


Figure S4

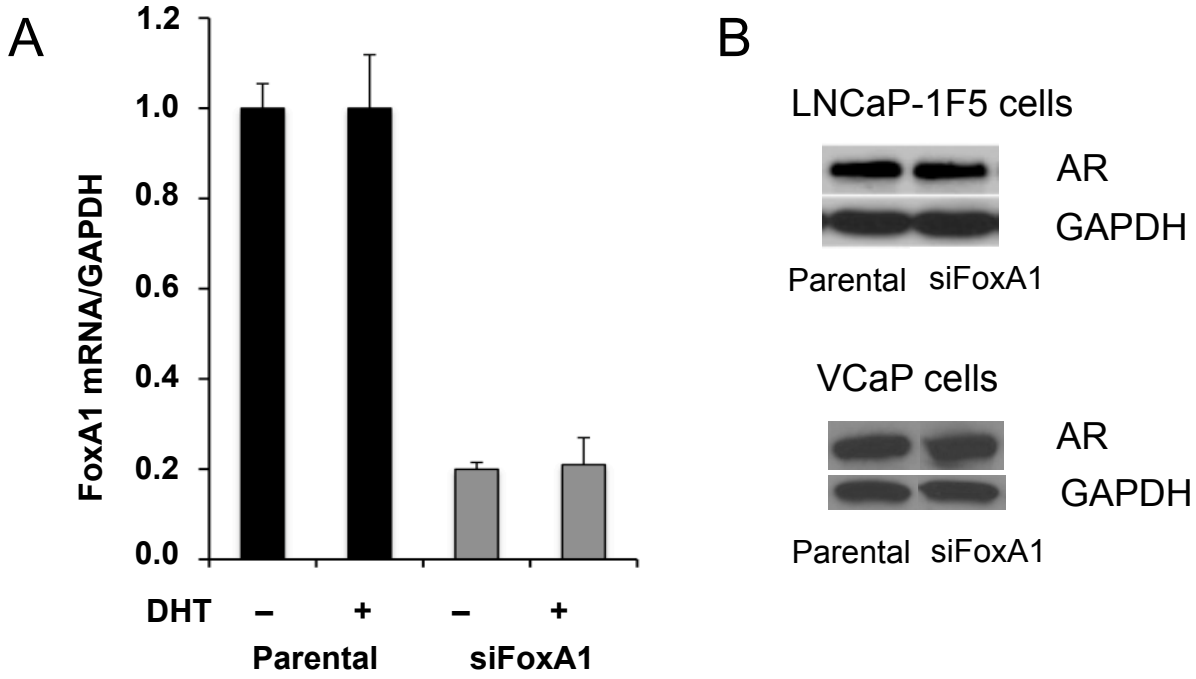

A
1.2
1.0
0.8
0.6
0.4
0.2
0.0
FoxA1 mRNA/GAPDH
DHT – + – +
Parental
siFoxA1
B
LNCaP-1F5 cells
AR
GAPDH
Parental siFoxA1
VCaP cells
AR
GAPDH
Parental siFoxA1

Figure S5

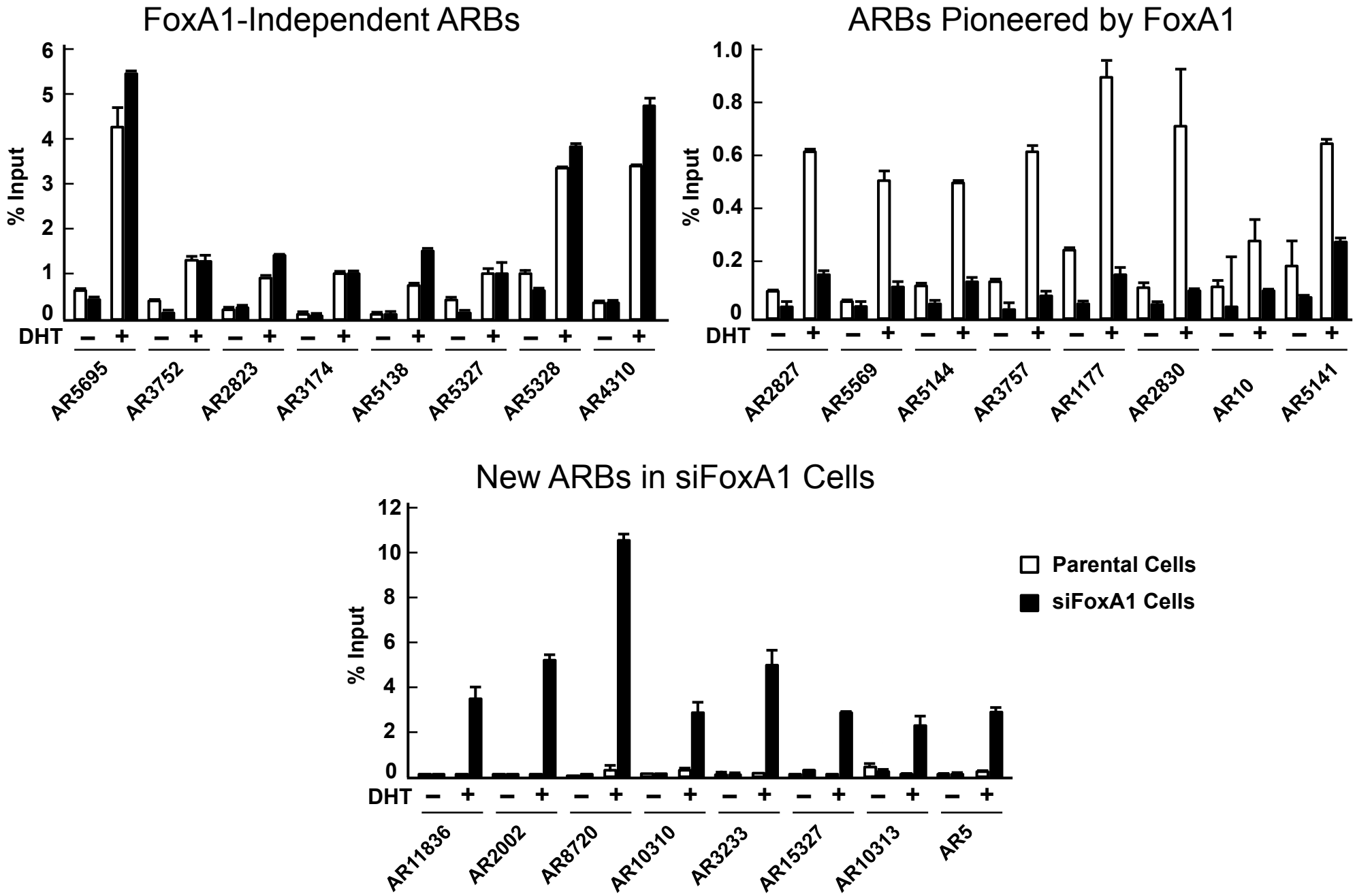

FoxA1-Independent ARBs
% Input
DHT
AR5695
AR3752
AR2823
AR3174
AR5138
AR5327
AR5328
AR4310
ARBs Pioneered by FoxA1
AR2827
AR5569
AR5144
AR3757
AR1177
AR2830
AR10
AR5141
New ARBs in siFoxA1 Cells
AR11836
AR2002
AR8720
AR10310
AR3233
AR15327
AR10313
AR5
Parental Cells
siFoxA1 Cells

Figure S6

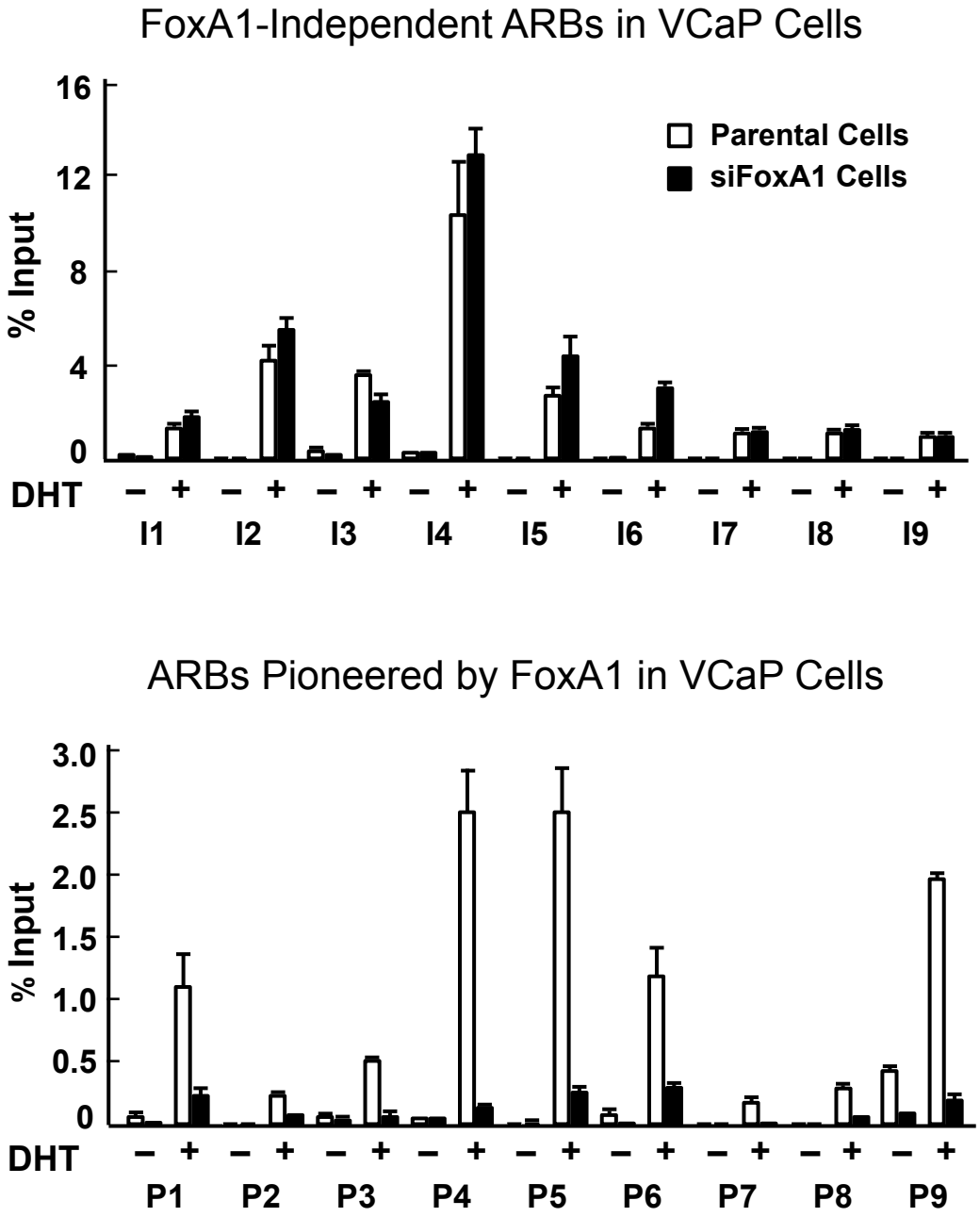

FoxA1-Independent ARBs in VCaP Cells
Parental Cells
siFoxA1 Cells
% Input
16
12
8
4
0
DHT
– + – + – + – + – + – + – + – + – +
I1 I2 I3 I4 I5 I6 I7 I8 I9
ARBs Pioneered by FoxA1 in VCaP Cells
% Input
3.0
2.5
2.0
1.5
1.0
0.5
0
DHT
– + – + – + – + – + – + – + – + – +
P1 P2 P3 P4 P5 P6 P7 P8 P9

Figure S7A

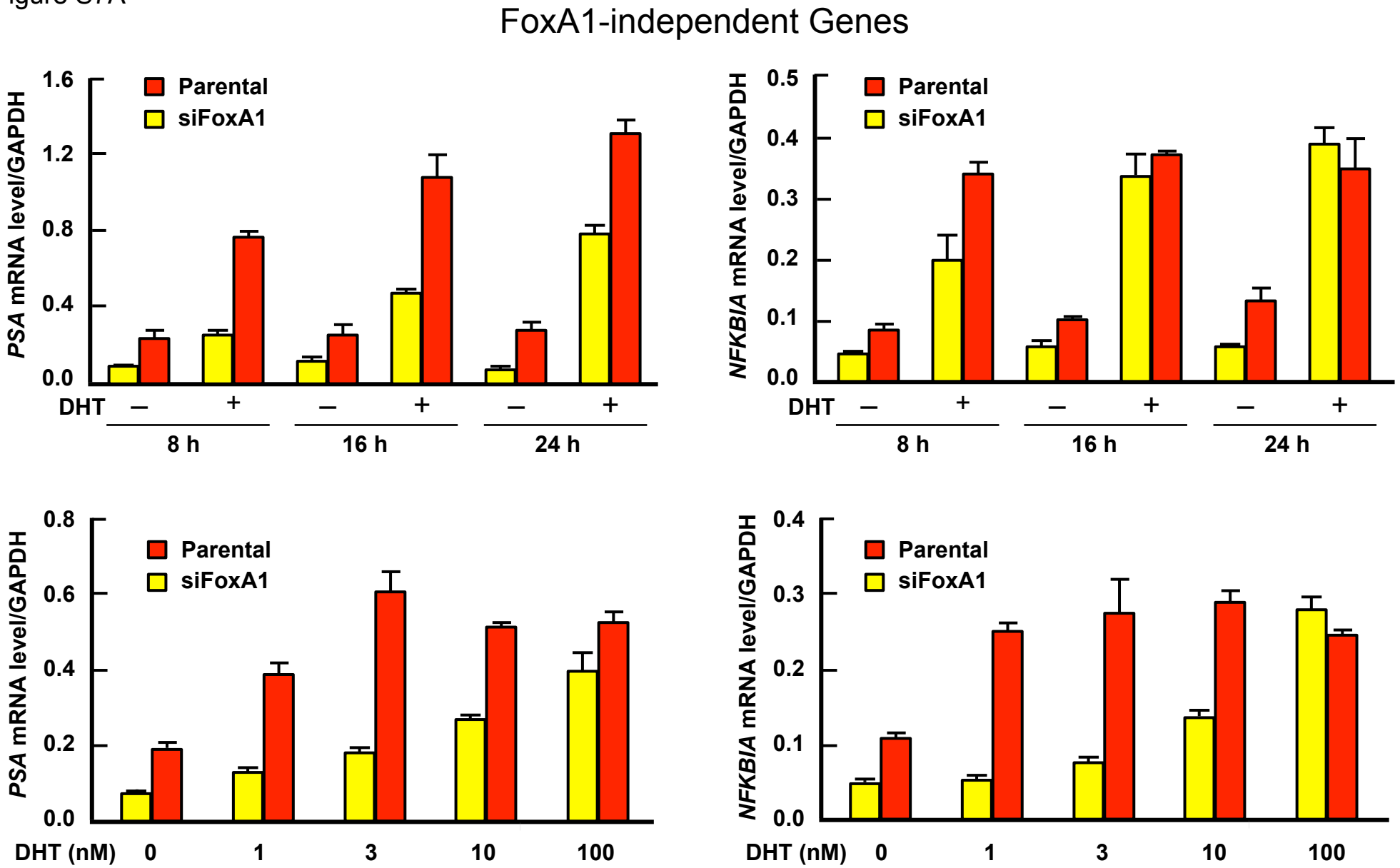

FoxA1-independent Genes
PSA mRNA level/GAPDH
Parental
siFoxA1
1.6
1.2
0.8
0.4
0.0
DHT
–
+
8 h
16 h
24 h
NFKBIA mRNA level/GAPDH
0.5
0.4
0.3
0.2
0.1
0.0
DHT (nM)
0
1
3
10
100

Figure S7B

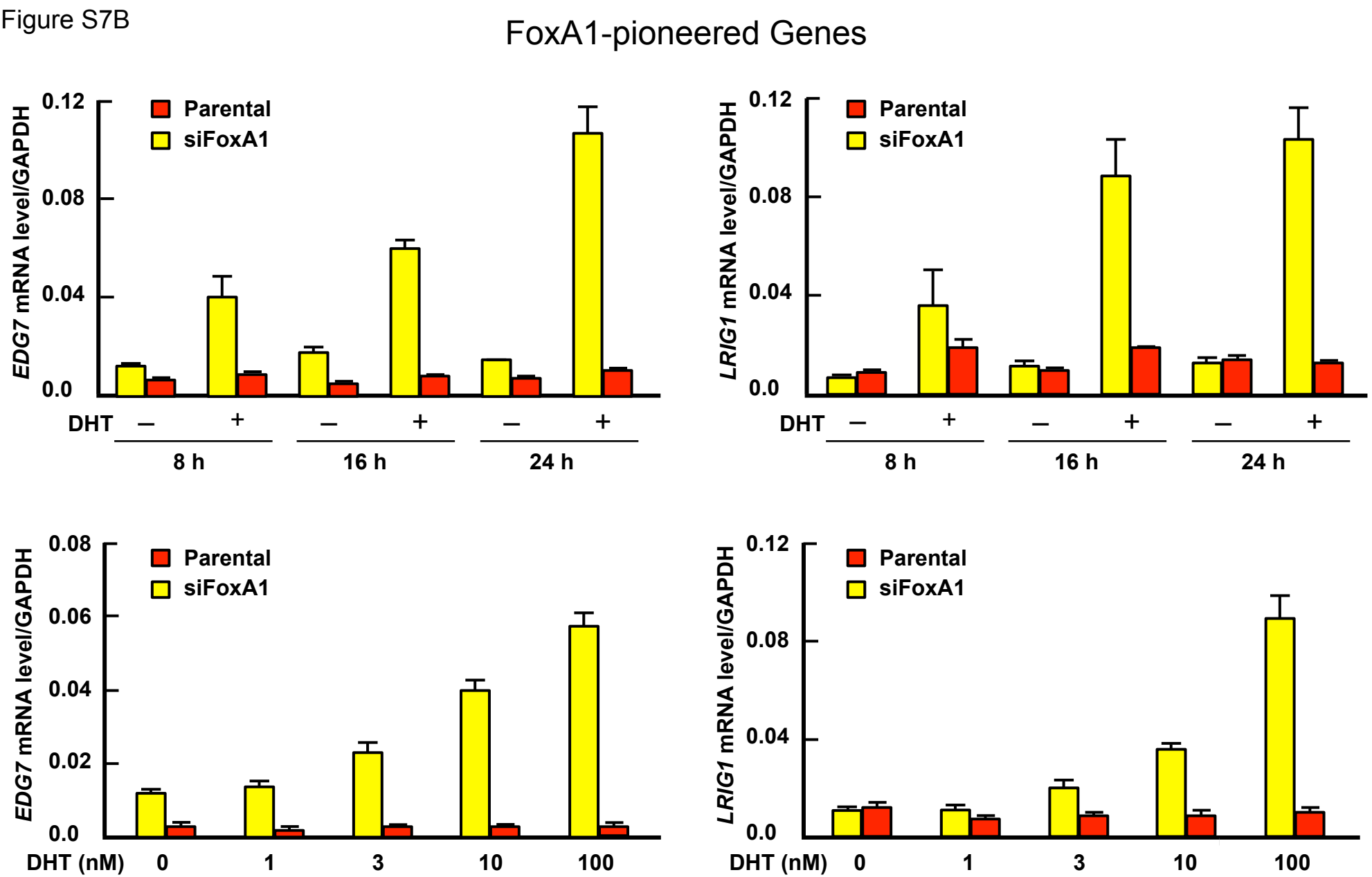

FoxA1-pioneered Genes
EDG7 mRNA level/GAPDH
LRIG1 mRNA level/GAPDH
Parental
siFoxA1
0.12
0.08
0.04
0.0
DHT
–
+
8 h
16 h
24 h
0.06
0.02
DHT (nM)
0
1
3
10
100

Figure S7C

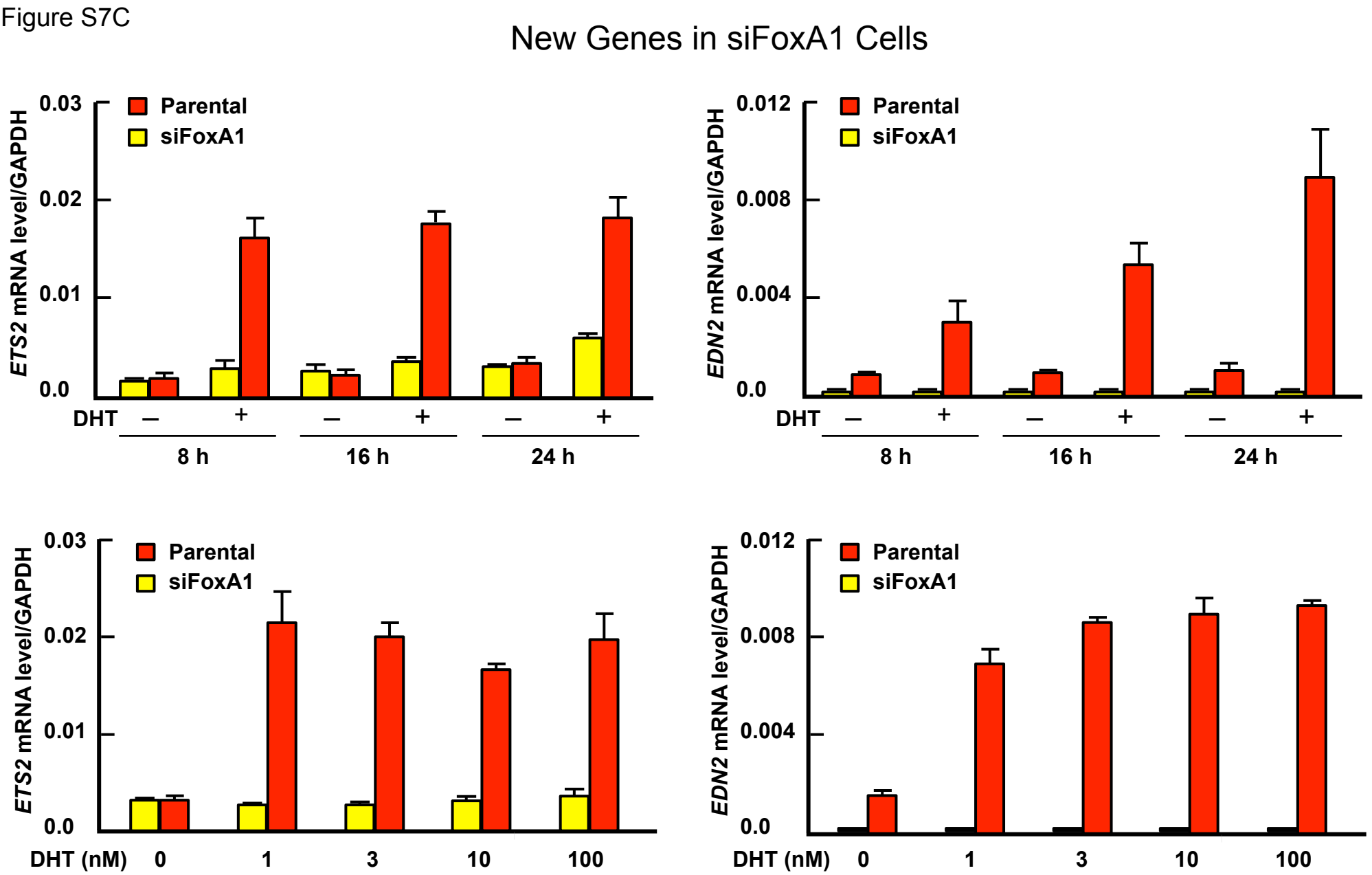

New Genes in siFoxA1 Cells
ETS2 mRNA level/GAPDH
EDN2 mRNA level/GAPDH
Parental
siFoxA1
0.03
0.02
0.01
0.0
0.012
0.008
0.004
DHT
–
+
8 h
16 h
24 h
DHT (nM)
0
1
3
10
100

Figure S8

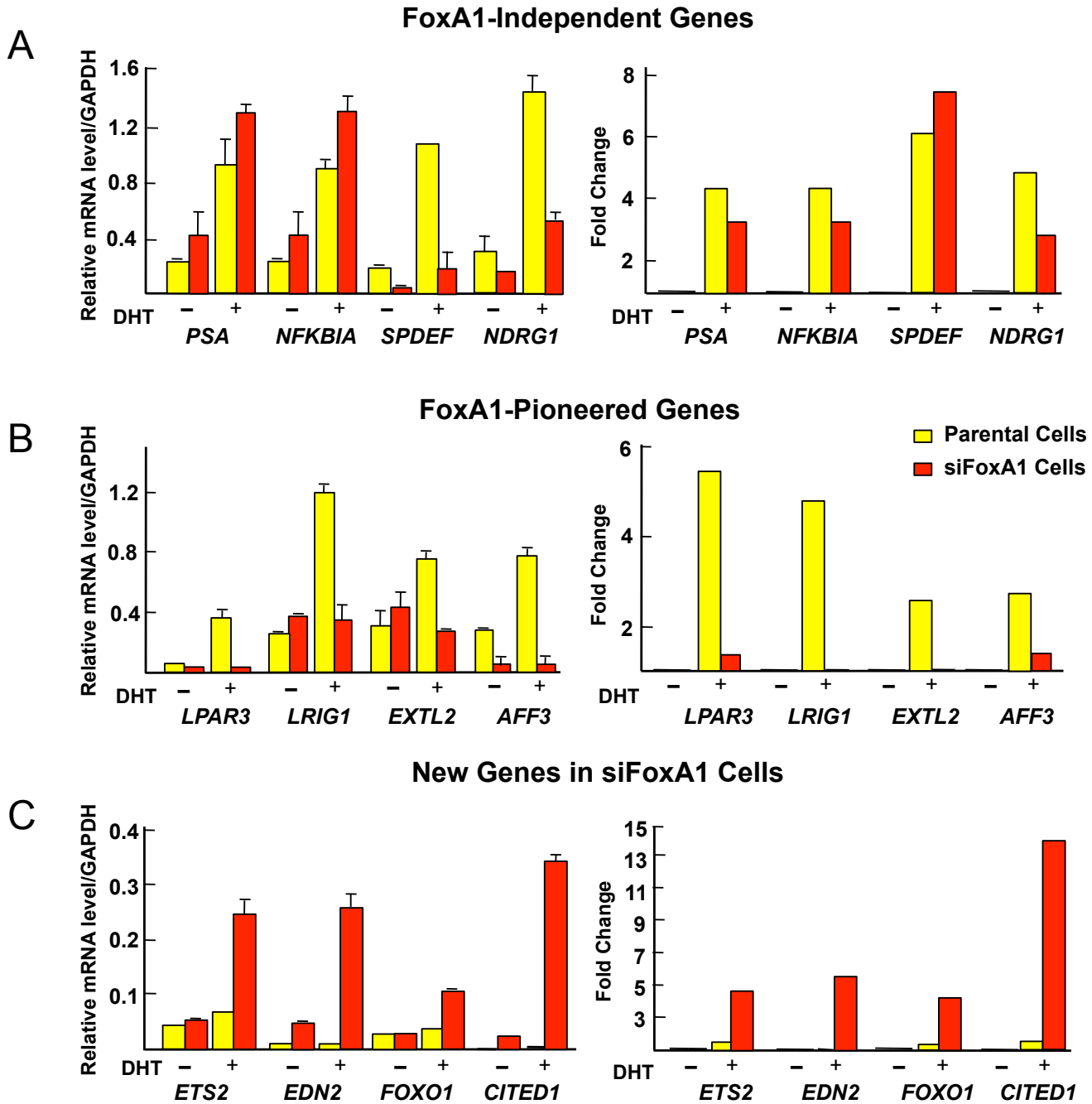

A
FoxA1-Independent Genes
Relative mRNA level/GAPDH
1.6
1.2
0.8
0.4
DHT
PSA
NFKBIA
SPDEF
NDRG1
Fold Change
8
6
4
2
B
FoxA1-Pioneered Genes
Parental Cells
siFoxA1 Cells
LPAR3
LRIG1
EXTL2
AFF3
C
New Genes in siFoxA1 Cells
0.4
0.3
0.2
0.1
15
13
11
9
7
5
3
ETS2
EDN2
FOXO1
CITED1

Figure S9

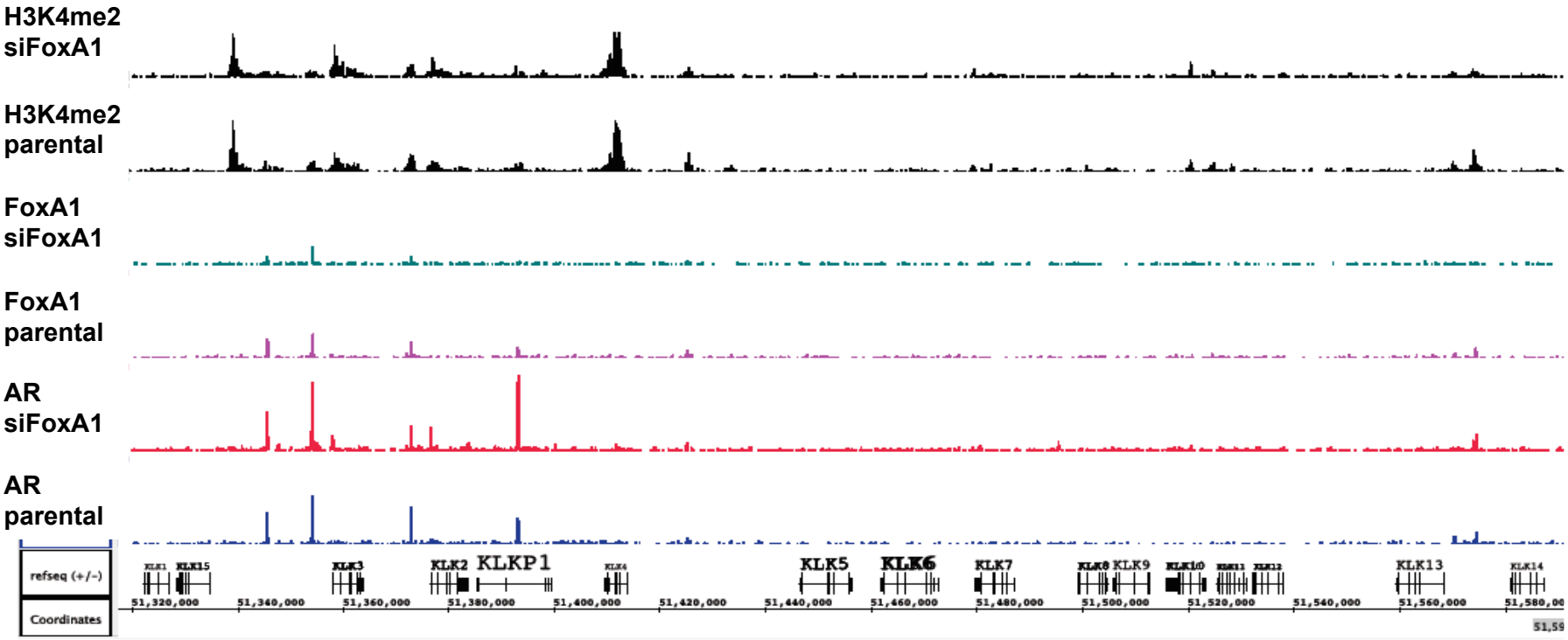
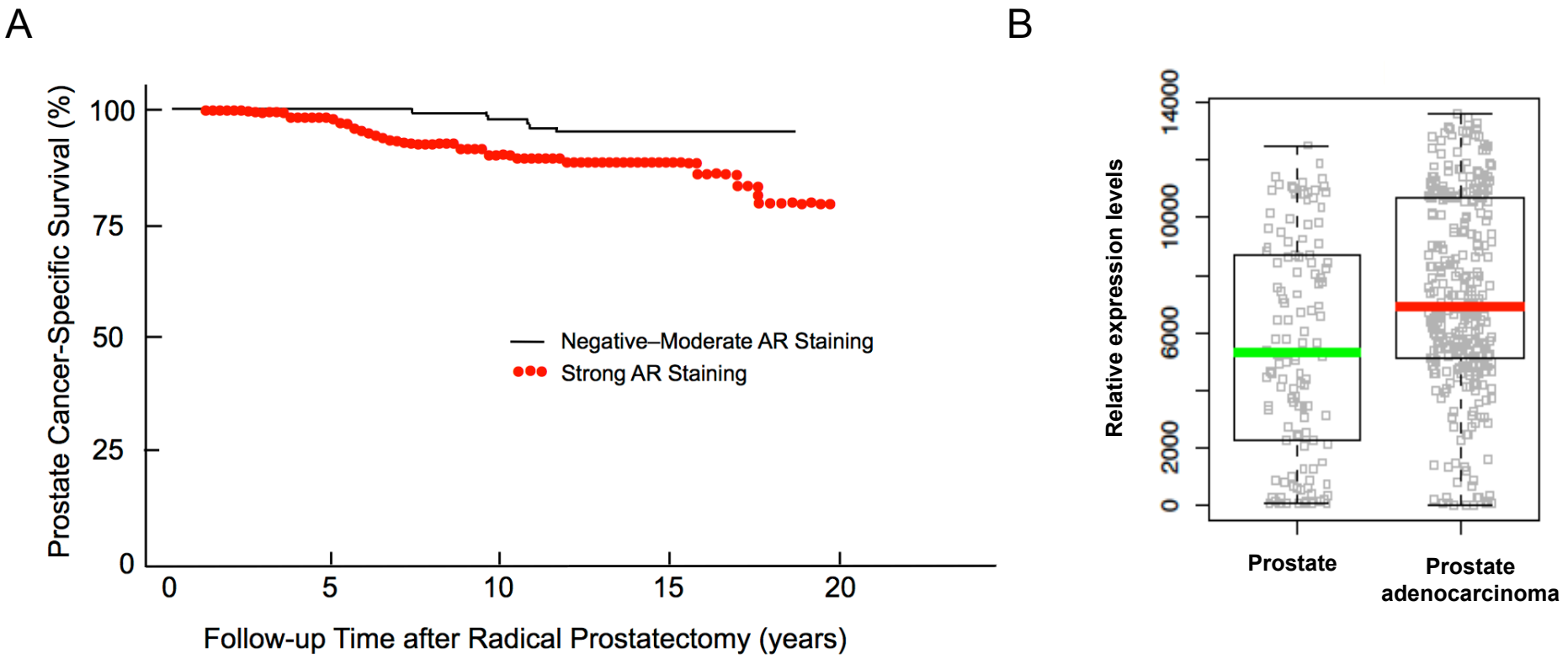

H3K4me2
siFoxA1
H3K4me2
parental
FoxA1
siFoxA1
FoxA1
parental
AR
siFoxA1
AR
parental
refseq (+/-)
Coordinates
KLK1
KLK15
KLK3
KLK2
KLKP1
KLK4
KLK5
KLK6
KLK7
KLK8
KLK9
KLK10
KLK11
KLK12
KLK13
KLK14
51,320,000
51,340,000
51,360,000
51,380,000
51,400,000
51,420,000
51,440,000
51,460,000
51,480,000
51,500,000
51,520,000
51,540,000
51,560,000
51,580,00
51,59

Figure S10


A
Prostate Cancer-Specific Survival (%)
100
75
50
25
0
0
5
10
15
20
Follow-up Time after Radical Prostatectomy (years)
Negative–Moderate AR Staining
Strong AR Staining
B
Relative expression levels
0
2000
6000
10000
14000
Prostate
Prostate adenocarcinoma