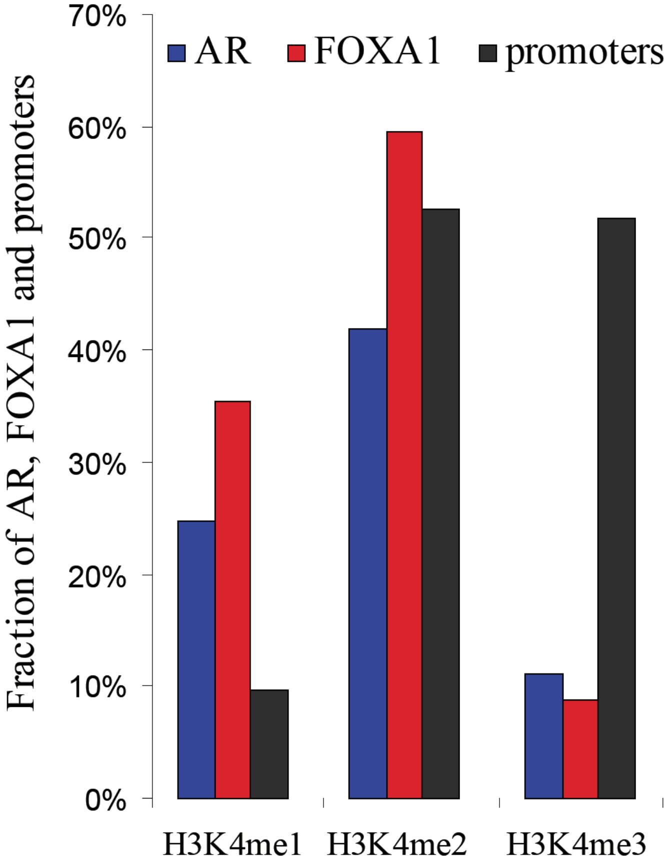
Supporting online material for

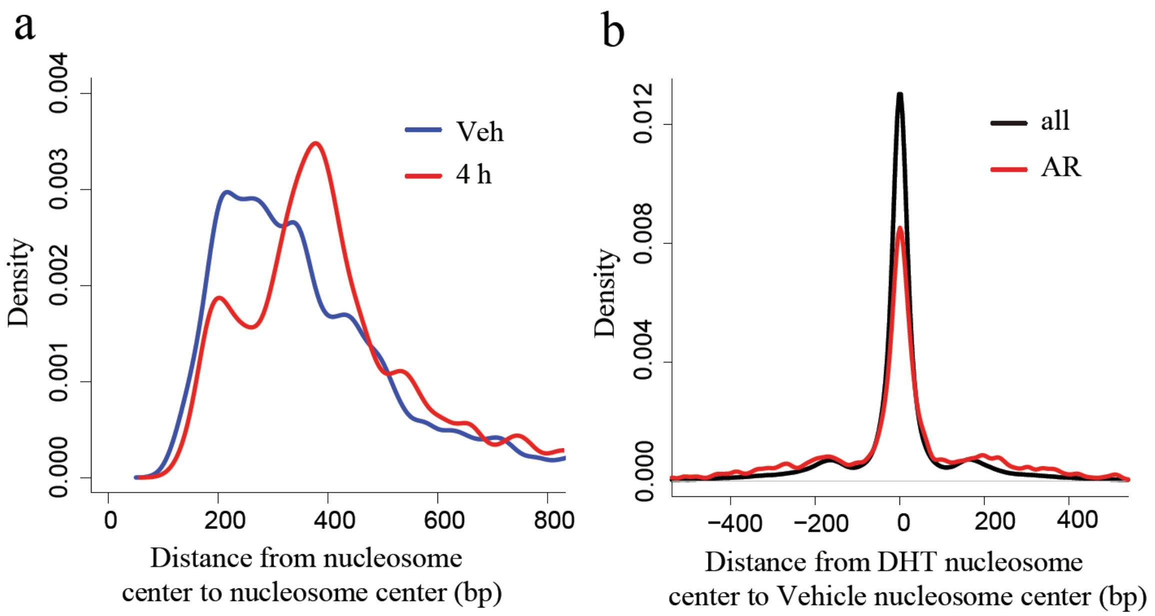
Nucleosome Dynamics Define Transcriptional Enhancers

Housheng Hansen He, Clifford A Meyer, Hyunjin Shin, Shannon T Bailey, Gang Wei, Qianben Wang, Yong Zhang, Kexin Xu, Min Ni, Mathieu Lupien, Piotr Mieczkowski, Jason D Lieb, Keji Zhao, Myles Brown*, Xiaole Shirley Liu*
* To whom correspondence should be addressed. Email: Myles_Brown@dfci.harvard.edu, xsliu@jimmy.harvard.edu

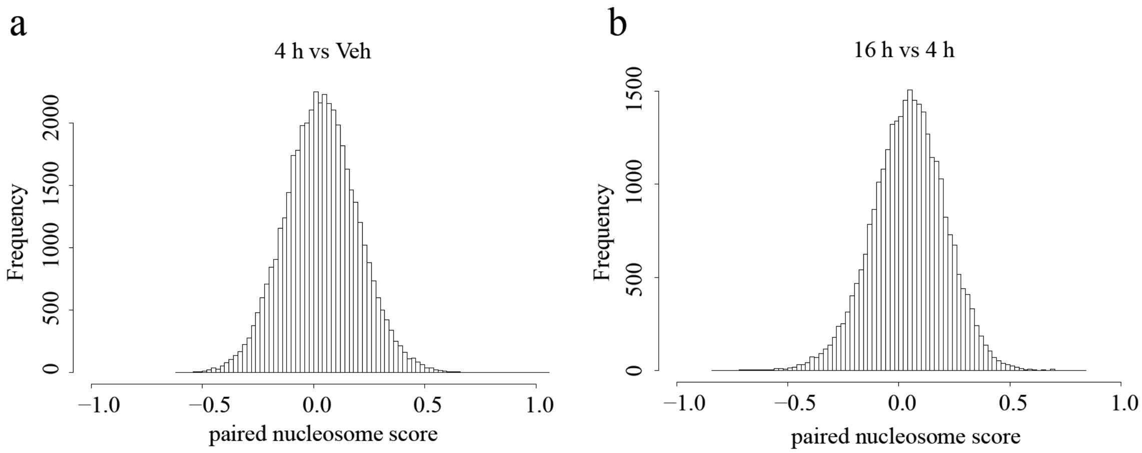
Supplementary Figures



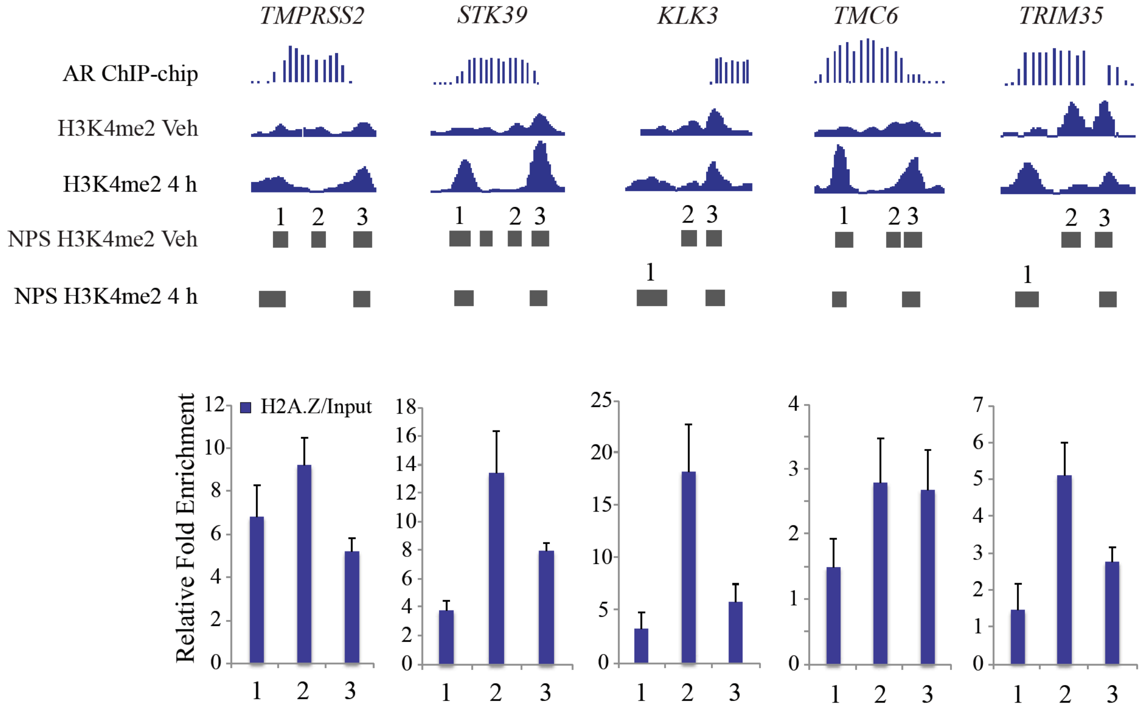
Supplementary Figure 1. Fraction of AR, FOXA1 binding sites and refseq promoter regions overlapping with H3K4me1/2/3 peak regions. 8727 AR and 10006 FOXA1 binding sites were identified by ChIP-chip. Upstream and downstream 1 kb regions of the transcription start sites of the 39289 refseq genes (hg18) were determined as promoter regions. H3K4me1/2/3 peak regions were identified using MACS software.



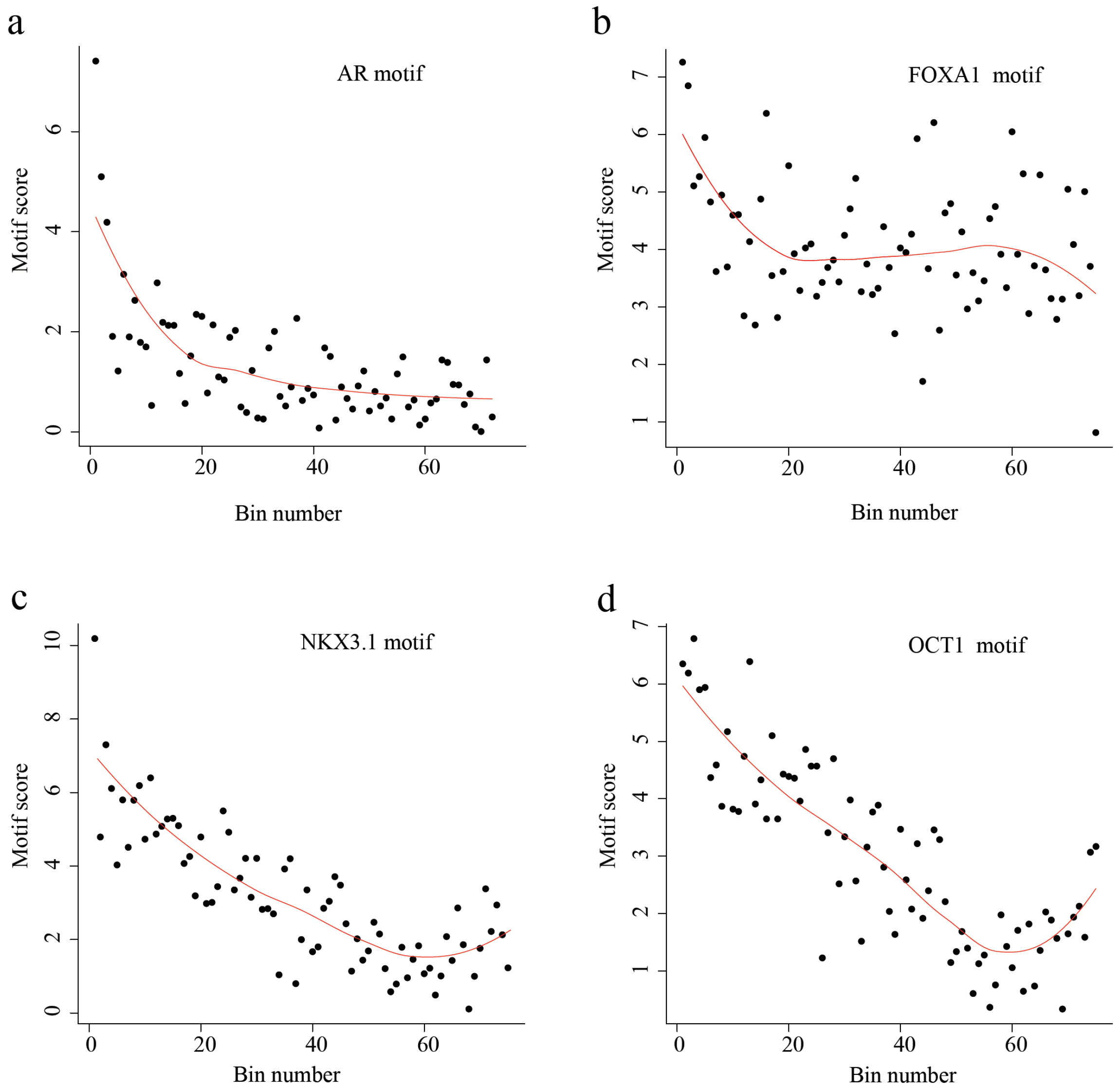
Supplementary Figure 2. Signal distribution and nucleosome position analysis in the AR and FOXA1 binding regions identified by ChIP-chip experiment. (*a*) Nucleosome center-to-center distance distributions between the nearest nucleosomes upstream and downstream of the AR motif in the AR binding sites under vehicle (red) and 4 hours after DHT stimulation (blue). (*b*) Distributions of location shifts between nucleosomes under vehicle and after 4 hours of DHT treatment. Nucleosomes detected within 500 bp of AR binding sites (red) under DHT and further than 500 bp of AR sites (black) are matched with the nearest identified nucleosome detected under vehicle conditions.



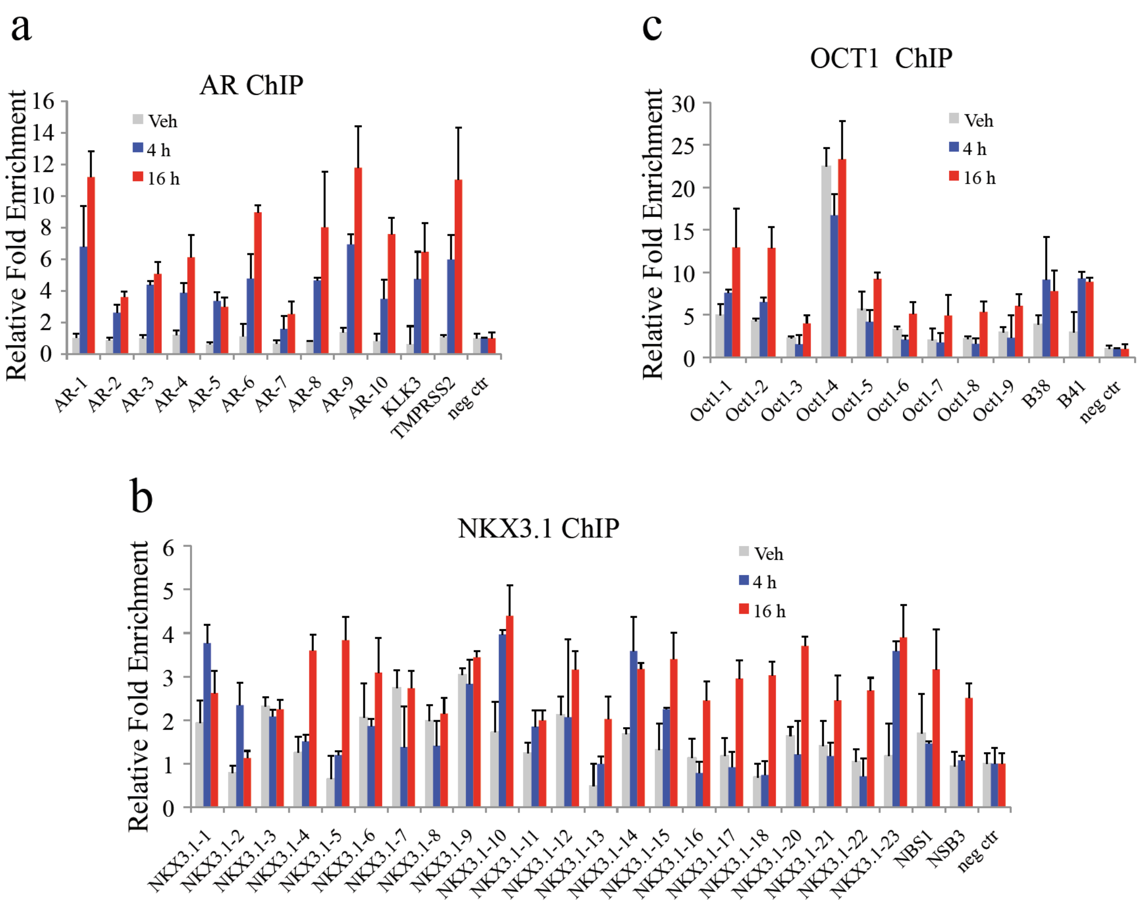
Supplementary Figure 3. Distribution of the scores of paired nucleosome regions under the conditions of DHT 4 hours vs vehicle (*a*) and DHT 16 hours vs 4 hours (*b*). The NSD score for each paired nucleosome region was calculated using the formula in the manuscript. The y-axis represents the number of the paired nucleosome regions within each score bin.



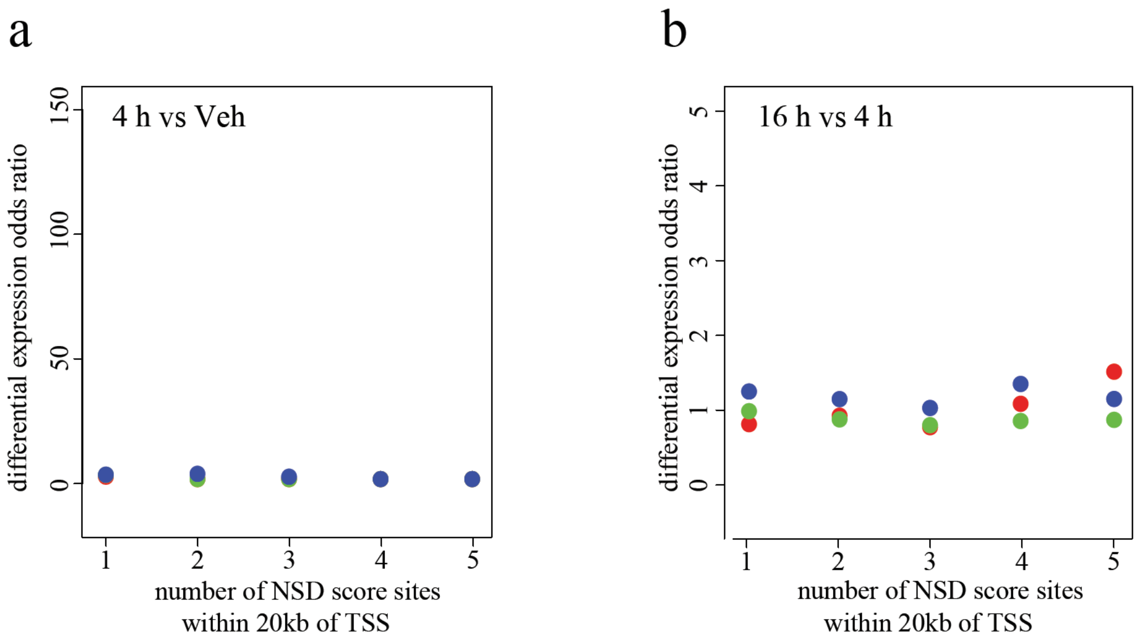
Supplementary Figure 4. H2A.Z ChIP-qPCR at 5 AR binding sites. Five AR binding sites near the genes *TMPRSS2*, *STK39*, *KLK3*, *TMC6* and *TRIM35 are shown*. "AR ChIP-chip" represents the AR ChIP-chip signals; "H3K4me2 Veh" and "H3K4me2 4 h" represent H3K4me2 ChIP-Seq signals under vehicle condition and after 4 hours of DHT treatment. Each horizontal bar represents a NPS peak region. “H2A.Z/Input” represents the qPCR assay of fold change for H2A.Z signal under vehicle condition.



Supplementary Figure 5. Motif scores of AR (*a*), FOXA1 (*b*), NKX3.1 (*c*) and OCT1 (*d*) in paired nucleosome bins. Paired nucleosome regions are ranked by scores representing the differences in H3K4me2 tag counts before and after DHT treatment. These ranked regions are grouped into bins of 500. The centrality bias (as described in Materials and Methods) statistic for each of these bins is represented on the y-axis.



Supplementary Figure 6. ChIP-qPCR validation of predicted (a) AR, (b) NKX3.1 and (c) OCT1 binding sites. Ten sites with an AR motif, but not overlapping with AR ChIP-chip binding sites, are selected from the top 1000 paired nucleosome regions (4 h vs. Veh). Two known AR binding sites in the enhancer region of KLK3 and TMPRSS2 are selected as positive controls. Twenty-two sites with NKX3.1 motif are selected from the top 1000 paired nucleosome regions (16 h vs. 4 h). NBS1 and NBS3 are 2 known NKX3.1 binding regions [1]. Nine sites with OCT1 motif are selected from the top 1000 paired nucleosome regions (16 h vs. 4 h), B38 and B41 are 2 known OCT1 binding sites [2].



Supplementary Figure 7. Correlation of score and number of NSD scoring sites and down-regulated gene expression. (*a*) 4 h vs Veh, (*b*) 16 h vs 4 h. X axis represents the lower bound *n* of the number of sites within 20 kb of the TSS of a gene, Y axis represents the odds ratio calculated by the formula (down-regulated genes with at least *n* sites/non-regulated genes with at least *n* sites)/(all down-regulated genes/all non-regulated genes). Red, green and blue dots represent the top 5000, 10,000 and 20,000 NSD score sites, respectively. This shows that the likelihood of a gene being down-regulated dose not change with the number and score of paired nucleosome sites.

Supplementary Table 1. Summary data for ChIP-seq regions enriched in H3K4me2 and H3K4me3

Antibody	Condition	Tag number	MACs peaks	NPS Positioned nucleosomes
H3K4me3	Veh	6.5 M	18424	74491
H3K4me3	DHT 4 h	6.5 M	20273	104022
H3K4me2	Veh	12.7 M	85830	235549
H3K4me2	DHT 4 h	12.7 M	90053	248486
H3K4me2	DHT 16 h	12.7 M	86728	237595

Supplementary Table 2. Motifs enriched in 1000 paired nucleosome regions with highest NSD scores.

4h vs. Veh		**16h vs. 4h**	
motif	p-value	motif	p-value
M00289.HFH-3.matrix	1.00E-20	M00131.HNF-3beta.matrix	1.00E-20
M00422.FOXJ2.matrix	1.00E-20	M00724.HNF-3alpha.matrix	1.13E-19
M00266.Croc.matrix	1.00E-20	M00130.FOXD3.matrix	1.13E-19
M00724.HNF-3alpha.matrix	1.00E-20	M00422.FOXJ2.matrix	1.13E-19
M00293.Freac-7.matrix	1.13E-19	M00266.Croc.matrix	1.13E-19
M00131.HNF-3beta.matrix	1.13E-19	M00293.Freac-7.matrix	9.48E-18
M00809.FOX.matrix	1.13E-19	M00289.HFH-3.matrix	6.22E-16
M00290.Freac-2.matrix	9.48E-18	M00477.FOXO3.matrix	6.22E-16
M00129.HFH-1.matrix	6.22E-16	M00742.HFH-4.matrix	3.19E-14
M01012.HNF3.matrix	6.22E-16	M00791.HNF-3.matrix	3.19E-14
M00481.AR.matrix	6.22E-16	M00809.FOX.matrix	3.19E-14
M00791.HNF-3.matrix	6.22E-16	M00129.HFH-1.matrix	3.19E-14
M00477.FOXO3.matrix	3.19E-14	M00290.Freac-2.matrix	1.28E-12
M00130.FOXD3.matrix	3.19E-14	M01011.HNF1.matrix	1.28E-12
M00481.AR.matrix	1.28E-12	M00292.Freac-4.matrix	1.28E-12
M00291.Freac-3.matrix	1.28E-12	M01012.HNF3.matrix	4.02E-11
M00956.AR.matrix	1.28E-12	M00093.BR-C.matrix	4.02E-11
M00957.PR.matrix	1.28E-12	M00471.TBP.matrix	9.87E-10
M00953.AR.matrix	1.28E-12	M01010.HMGIY.matrix	1.38E-09
M00955.GR.matrix	4.02E-11	M00476.FOXO4.matrix	4.64E-09
M00292.Freac-4.matrix	4.02E-11	M00160.SRY.matrix	6.59E-09
M00473.FOXO1.matrix	4.02E-11	M00473.FOXO1.matrix	1.22E-08
M00476.FOXO4.matrix	4.02E-11	M00092.BR-C.matrix	1.39E-08
M00472.FOXO4.matrix	4.02E-11	**M00138.OCT1.matrix**	1.34E-08
M00474.FOXO1.matrix	4.02E-11	M00472.FOXO4.matrix	1.85E-08
M00742.HFH-4.matrix	4.02E-11	M00096.Pbx-1.matrix	2.89E-08
M00093.BR-C.matrix	9.87E-10	M00291.Freac-3.matrix	2.78E-08
M00954.PR.matrix	9.87E-10	**M00451.NKX3A.matrix**	1.03E-07
M00921.GR.matrix	1.14E-09	M00980.TBP.matrix	2.30E-07
M00160.SRY.matrix	1.52E-09	M00987.FOXP1.matrix	5.99E-07
M00960.PR.matrix	1.22E-08	M00790.HNF-1.matrix	7.73E-07
M00790.HNF-1.matrix	1.39E-08		
M00980.TBP.matrix	1.42E-08		
M00471.TBP.matrix	4.09E-08		
M01011.HNF1.matrix	6.75E-08		

Supplementary Table 3. Primers used for PCR.

Primers used in Fig. 2			
KLK3_nuc1_F	GCAGTCTAGGTGGATGCTGT	KLK3_nuc1_R	GGTTTGCAAGTTGTCCCAGTA
KLK3_nuc2_F	TTATCTAGGACAGTAAGCAAGCC	KLK3_nuc2_R	GGATGTTTGTAAAGCAGGCAT
KLK3_nuc3_F	TCCCAACACTGTTCAGCCA	KLK3_nuc3_R	GCAGGATGTGACAGAAGAAGTAT
TMPRSS2_nuc1_F	TCTGAGCCCCCTACAATTGCA	TMPRSS2_nuc1_R	TAGGGCAGGGGTGAGGAAGT
TMPRSS2_nuc2_F	GCCACCTGGTGAAGTGCAGA	TMPRSS2_nuc2_R	TGGAGCTAGTGCTGCATGTC
TMPRSS2_nuc3_F	AAAGGGCCTACAAGATAAGATC	TMPRSS2_nuc3_R	TCAACCATGGAAAGCAGGTG
TRIM35_nuc1_F	GGATTTTAGTAACATCTTTGGGG	TRIM35_nuc1_R	CCCAATACATTATAATTTAGAGCAG
TRIM35_nuc2_F	AAAGGAGACAGACTGTTCTAGTACA	TRIM35_nuc2_R	TTGTGTTGTTTGATTAGATAATCTG
TRIM35_nuc3_F	CAACATAACCATTTTCTCATGC	TRIM35_nuc3_R	CAGATTTTTAACAAGAACGTTTGCA
TMC6_nuc1_F	GGGATTCTCCCAGCACCACT	TMC6_nuc1_R	TGGGTGGCATTTCAGAGC
TMC6_nuc2_F	AACTAGCCAGCAGCAGGGTCA	TMC6_nuc2_R	TGGCGTCTGCCAAGGCGA
TMC6_nuc3_F	CTAGCCCTGGTGTGGAGCAT	TMC6_nuc3_R	GCCATTTGTCCTCCTGCA
STK39_nuc1_F	GAATGTTCTGGAAGTTGGAAG	STK39_nuc1_R	GTTTGATAGGACTGATGCAGAATAA
STK39_nuc2_F	CACACATCATGTAAGACCAAAGG	STK39_nuc2_R	CTTAACTTATTGCTGTATCCTAAGC
STK39_nuc3_F	CAGGGCTGGAACCAGAAGA	STK39_nuc3_R	GCACAGTCAACCTGACAACTAG
TMPRSS2_tile1_F	TCTGAGCCCCCTACAATTGCA	TMPRSS2_tile1_R	TAGGGCAGGGGTGAGGAAGT
TMPRSS2_tile2_F	ACTTCCTCACCCCTGCCCTA	TMPRSS2_tile2_R	GAAAATGTTGGTCCTGGATG
TMPRSS2_tile3_F	CATCCAGGACCAACATTTTC	TMPRSS2_tile3_R	TACGGCAGGTACTCATATCTGA
TMPRSS2_tile4_F	TTTAGACAACAAATGGCCACC	TMPRSS2_tile4_R	TGGAGCTAGTGCTGCATGTC
TMPRSS2_tile5_F	GCCACCTGGTGAAGTGCAGA	TMPRSS2_tile5_R	GGAGCTAGTGCTGCATGTCTT
TMPRSS2_tile6_F	TCAGATATGAGTACCTGCCGTA	TMPRSS2_tile6_R	GGAAGCTGACCTTTAATGAAG
TMPRSS2_tile7_F	AAGACATGCAGCACTAGCTCC	TMPRSS2_tile7_R	CAGGATGTGAGCTTGAAATGT
TMPRSS2_tile8_F	ACATTTCAAGCTCACATCCTG	TMPRSS2_tile8_R	GACACTTCACCCATCTTTGACA
TMPRSS2_tile9_F	TGTCAAAGATGGGTGAAGTGTC	TMPRSS2_tile9_R	GATCTTATCTTGTAGGCCCTTT
TMPRSS2_tile10_F	AAAGGGCCTACAAGATAAGATC	TMPRSS2_tile10_R	TCAACCATGGAAAGCAGGTG
STK39_tile1_F	GAATGTTCTGGAAGTTGGAAG	STK39_tile1_R	GTTTGATAGGACTGATGCAGAATAA
STK39_tile2_F	GGAACTATTCTGAGGAGGCAG	STK39_tile2_R	GGGAAATGCCATATTGGAG
STK39_tile3_F	GCACAGAGAAGCAATTAGATTT	STK39_tile3_R	GTTCTTTGTTGCCCTCACTT
STK39_tile4_F	TTATTCTGCATCAGTCCTATCAAAC	STK39_tile4_R	GGCTCTTAGCTGTTGCCA
STK39_tile5_F	CTAAAGCTCCTGCTGGAAAG	STK39_tile5_R	GTGCTATTTGTCTTACCAGGAC
STK39_tile6_F	GTCCTGGTAAGACAAATAGCAC	STK39_tile6_R	AACCTCGACTCTGCTCTGGA
STK39_tile7_F	CACACATCATGTAAGACCAAAGG	STK39_tile7_R	CTTAACTTATTGCTGTATCCTAAGC
STK39_tile8_F	AAACCAAATGTTTGGCTCC	STK39_tile8_R	CCTCTTTCTTCTGCAGGTAAGC
STK39_tile9_F	TCCAGAGCAGAGTCGAGGTT	STK39_tile9_R	GTTGGCTGTGTGCTTCTCTT
STK39_tile10_F	GCTTAGGATACAGCAATAAGTTAAG	STK39_tile10_R	TCTTCTGGTTCCAGCCCTG
STK39_tile11_F	GCTTACCTGCAGAAGAAAGAGG	STK39_tile11_R	GTTGCTTGTCAGAAATCTTCTGG
STK39_tile12_F	AAGAGAAGCACACAGCCAAC	STK39_tile12_R	AACTGACCATCGCAGTGTGT
STK39_tile13_F	CAGGGCTGGAACCAGAAGA	STK39_tile13_R	GCACAGTCAACCTGACAACTAG
Primers used in Fig. 4a			
AR_1_F	TCAACATCAAGATTCAAGCAAC	AR_1_R	TCTGCACATGTATCCTGGAA
AR_2_F	GGGCAGCACCAACATCACA	AR_2_R	TGTCCCAGTATTCAGACCAGAG
AR_3_F	ATAGAAAGTTTGTAACTTGCCAAGG	AR_3_R	GCACAATCCAGAAAATGCTATT
AR_4_F	TGGAAGGCAAGTGGCAGA	AR_4_R	CCCATTACCCTGTACCTTGTT
AR_5_F	CACAGCATTAAGTGGAAGCTC	AR_5_R	TGGGAAACAAATGACAGGA
AR_6_F	TAAGGTGCCCTGGGATGGT	AR_6_R	GGAGGGGAACCAAGATTAAC
AR_7_F	CCTCCACACCAAGGAAATAA	AR_7_R	ACCAATAAAATAGGCTAAAGGC
AR_8_F	GCAGAAATGCAACACTGGT	AR_8_R	ACCAGGGAAATCGGAAGC
AR_9_F	ATCTCTAGTAACACATACTGAGGGC	AR_9_R	CCTCAGGGTGAGAAATTAGACAT
AR_10_F	TGGTGATGAGATTGCGAAG	AR_10_R	AGCCCCTTGCAGTCTTCT
KLK3	TTATCTAGGACAGTAAGCAAGCC	KLK3	GGATGTTTGTAAAGCAGGCAT
TMPRSS2	GCCACCTGGTGAAGTGCAGA	TMPRSS2	TGGAGCTAGTGCTGCATGTC
Primers used in Fig. 4b			
OCT1_1_F	ATTTCTTTCCAAACATGTGAGAGTC	OCT1_1_R	GACTGAGTTGTTCTAGACAAGTGTT

OCT1_2_F	ATCATCTGATCTCTGCTGGATC	OCT1_2_R	CGTGACACTTTAACTTCCATAGTG
OCT1_3_F	GGAAAGCTGCTAAACATGGT	OCT1_3_R	GCAAAGCCAGCTCTGAAAA
OCT1_4_F	CAAGCCCCCCATGACATGAG	OCT1_4_R	TGCCTGGCTCTTCCCACTA
OCT1_5_F	AGCCACAAATTCTATACTTACAACG	OCT1_5_R	CTCATGCCACCAGAATTTATTT
OCT1_6_F	CTTGGGACAGGGTATGTTTT	OCT1_6_R	CCTTGTTTTTCCTTCCAATCTT
OCT1_7_F	CAGGGTCCCTAGAGTCCAGA	OCT1_7_R	TTAAAGAGACAGCGTCTCGG
OCT1_8_F	AGACTGATGAAGGCAGAGAGAG	OCT1_8_R	CATGGCATATGTAGGGACTTT
OCT1_9_F	TTTTGGTGACTGACTGGCC	OCT1_9_R	AACGCGGAGGATCCTAAG
B38_F	TGTAGGGAGGGAGCCACACT	B38_R	AATGGCCGGGTGTGCC
B41_F	GCCTCCCCCGTGCAG	B41_R	TGCAAGGCACGTCTCAATTC
Primers used in Fig. 4c			
NBS3_F	AAGAATGAGCTGATCCTCCTATGC	NBS3_R	GGTTTAGTAATAGACTGGGCACCA
NBS1_F	GATTCTTTGACTGGTCTGGCACAC	NBS1_R	TTATCCATTGTGCCTGGAGCTGAG
NKX3.1-1_F	ATTGTGTAACTATCGCCTCTGC	NKX3.1-1_R	CAAGTAGGTGATGGAAACAATGAAT
NKX3.1-2_F	TGTGACCAAGACAGTCATGC	NKX3.1-2_R	GAGAACTGCTCAGCAAGAAAA
NKX3.1-3_F	TGATCTGTTTTTGTTGGCACAT	NKX3.1-3_R	CAGTTCTCTGCCTACGAATTATTTA
NKX3.1-4_F	GCTCTTGAATCATTGCACTAAG	NKX3.1-4_R	GAAGAAAAACAAATATGCCCAG
NKX3.1-5_F	CTCCTTCATGAGAGATGCGG	NKX3.1-5_R	GGCAGGGCCTATGTAATATTG
NKX3.1-6_F	TAGTTTAAATACAGTGCACACACAC	NKX3.1-6_R	GGGCAGTATTTCTATTTAAACACCA
NKX3.1-7_F	CATGTACCCACATTCCATAAAA	NKX3.1-7_R	AAATTGTGACACCGCTGAGT
NKX3.1-8_F	GGTCACCACTCAGGTAGTATTGTT	NKX3.1-8_R	GTGCTAATAAGTATGTAAAAGACGG
NKX3.1-9_F	CTTTTTGCTTTGGCCACTA	NKX3.1-9_R	CACAGGCAAAAGAAGACTTAAA
NKX3.1-10_F	GGTGCAATAATTACACATAGTCCT	NKX3.1-10_R	CTTAAGAACCTGTTACTAAGCAAGT
NKX3.1-11_F	CCAATATACACTGTGACCTTCATC	NKX3.1-11_R	CTGACAGCTATACTAAAACTCAGTG
NKX3.1-12_F	TCATCAGCTCCAAGGAAAAC	NKX3.1-12_R	CATATGCGTAGGATTCACAGC
NKX3.1-13_F	AGGTCGATTTGATTGAGGAA	NKX3.1-13_R	CCCAACATCAACATCATACATT
NKX3.1-14_F	CTGCTGCAGGTCTGGGTTT	NKX3.1-14_R	CAGTGTAGTTTGTGCACCCA
NKX3.1-15_F	GAGTGAAACCCTGTACCAAACA	NKX3.1-15_R	CAGTAAAAAAGAGGAATAACTGGTG
NKX3.1-16_F	CCAAACATAATTTGGGAAGTGAACC	NKX3.1-16_R	CACACTCCAAAATCCACACTAATAA
NKX3.1-17_F	ATTGCATTGGAGGTTAGGTT	NKX3.1-17_R	TTCTCCATCGTCTCTGCTATC
NKX3.1-18_F	CAGGAAACTGACTGTAAGTACAAGA	NKX3.1-18_R	GCCCCTAAGAACATCGTTTC
NKX3.1-19_F	TGCAATTGGTCAAAGTGAGA	NKX3.1-19_R	CATCACAGGTGCCCAAAA
NKX3.1-20_F	AGTCTCTACTGCACTTTAGACCTCT	NKX3.1-20_R	CCCCGTTGAGTTAGATATTATTTTC
NKX3.1-21_F	GTCAAAGAGAAAGAGAGAGACAGTC	NKX3.1-21_R	CTTTTAAAGGAATAGGGTACACTG
NKX3.1-22_F	CAGTTTCCTCACCTTTAAATCTTAG	NKX3.1-22_R	AGTCTGATGATGGCATTACAAT
NKX3.1-23_F	GCTATTCTATAAAAAGGACACTTGG	NKX3.1-23_R	GGGCCATTTAAAATATACAGTG

References for Supplementary Figures and Tables

1. **Liu, W. et al. Characterization of two functional NKX3.1 binding sites upstream of the PCAN1 gene that are involved in the positive regulation of PCAN1 gene transcription. BMC Mol Biol 9, 45 (2008).**
2. **Wang, Q. et al. A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. Mol Cell 27, 380-92 (2007).**