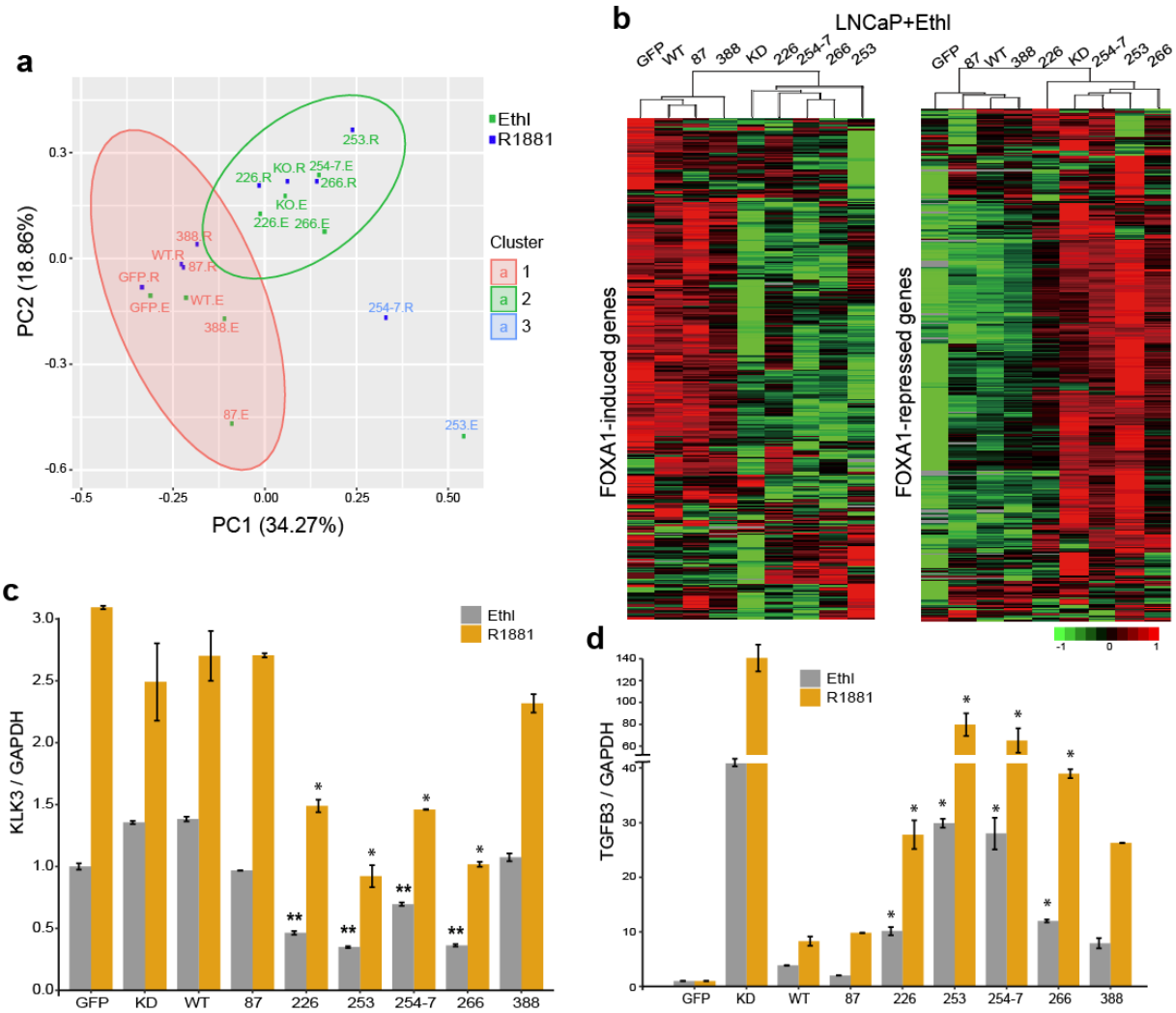
Supplementary information

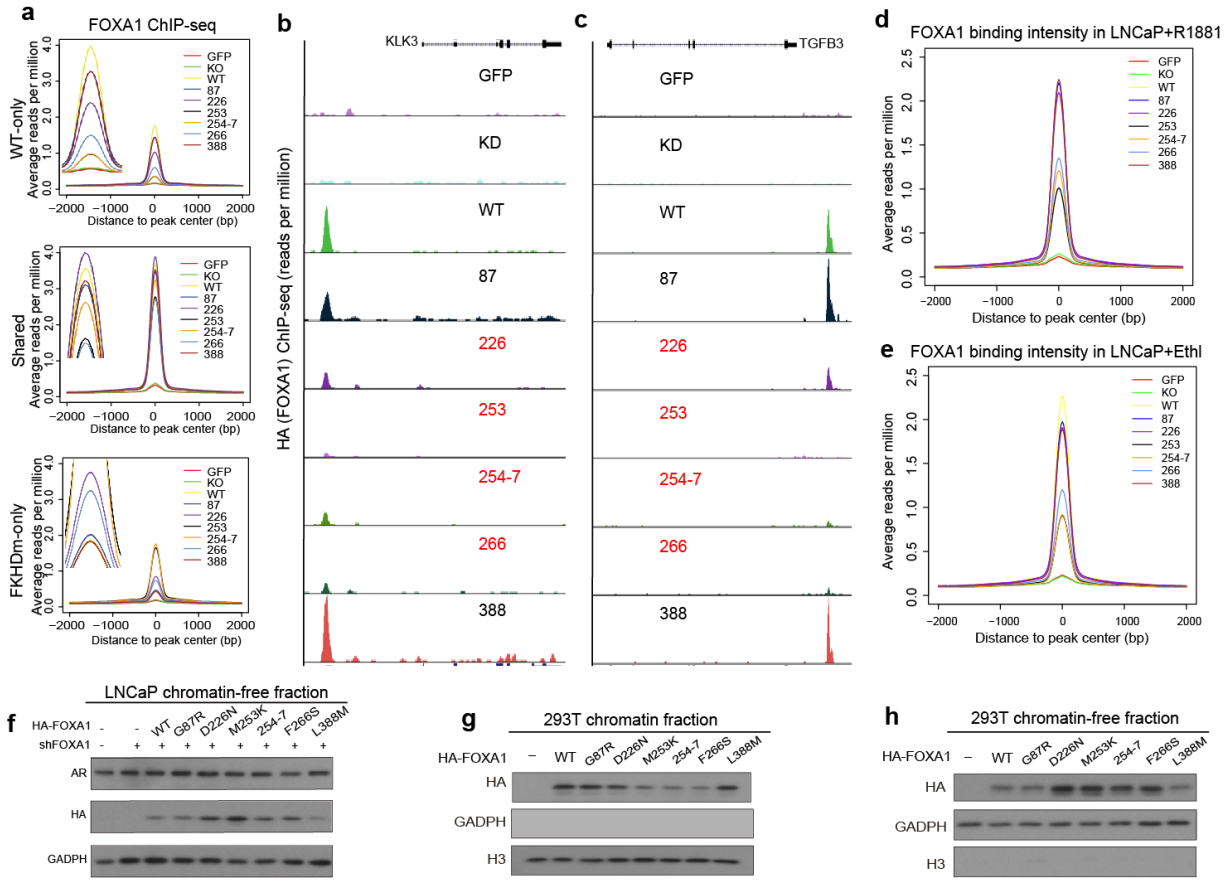
Table of Contents:

Fig. S1



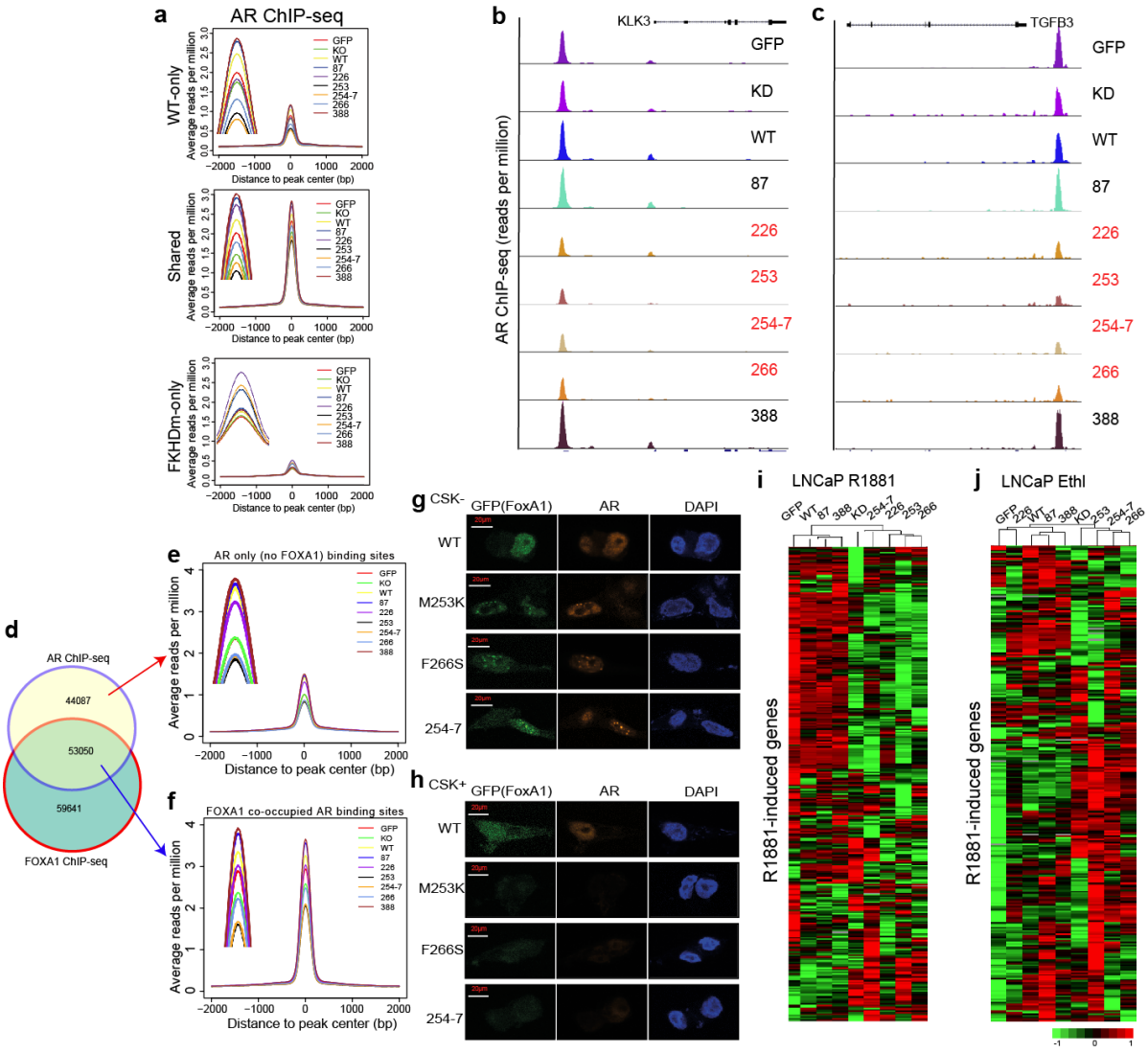
Supplementary information, Fig. S1 FKHDm FOXA1 has impaired ability to rescue FOXA1-mediated gene expression in FOXA1 KD cells. **a** Principle component analysis of the transcriptomes of the various LNCaP cells that were hormone-deprived for 3 days followed with ethanol (E, Ethl) or 1 nM R1881 (R, R1881) treatment for 24 h. Hierarchical clustering shows the relatedness of each sample. The ovals and font colors indicate clusters. **b** Heatmap showing gene expression across Ethl-stimulated LNCaP cells. FOXA1-induced (n=359, left panel) and -repressed genes (n=386, right panel) were derived by comparing GFP vs KD cells. Hierarchical clustering was utilized to group genes and samples. **c-d**. QRT-PCR analysis of KLK3 (**c**) and TGFB3 (**d**) mRNA levels in hormone-starved LNCaP cell line panels stimulated with Ethl or R1881 for 24 h. Target gene expression was normalized to GAPDH control. *****$P < 0.05$, ******$P < 0.005$, FKHDm compared to their respective WT under R1881 or Ethl condition.

Fig. S2



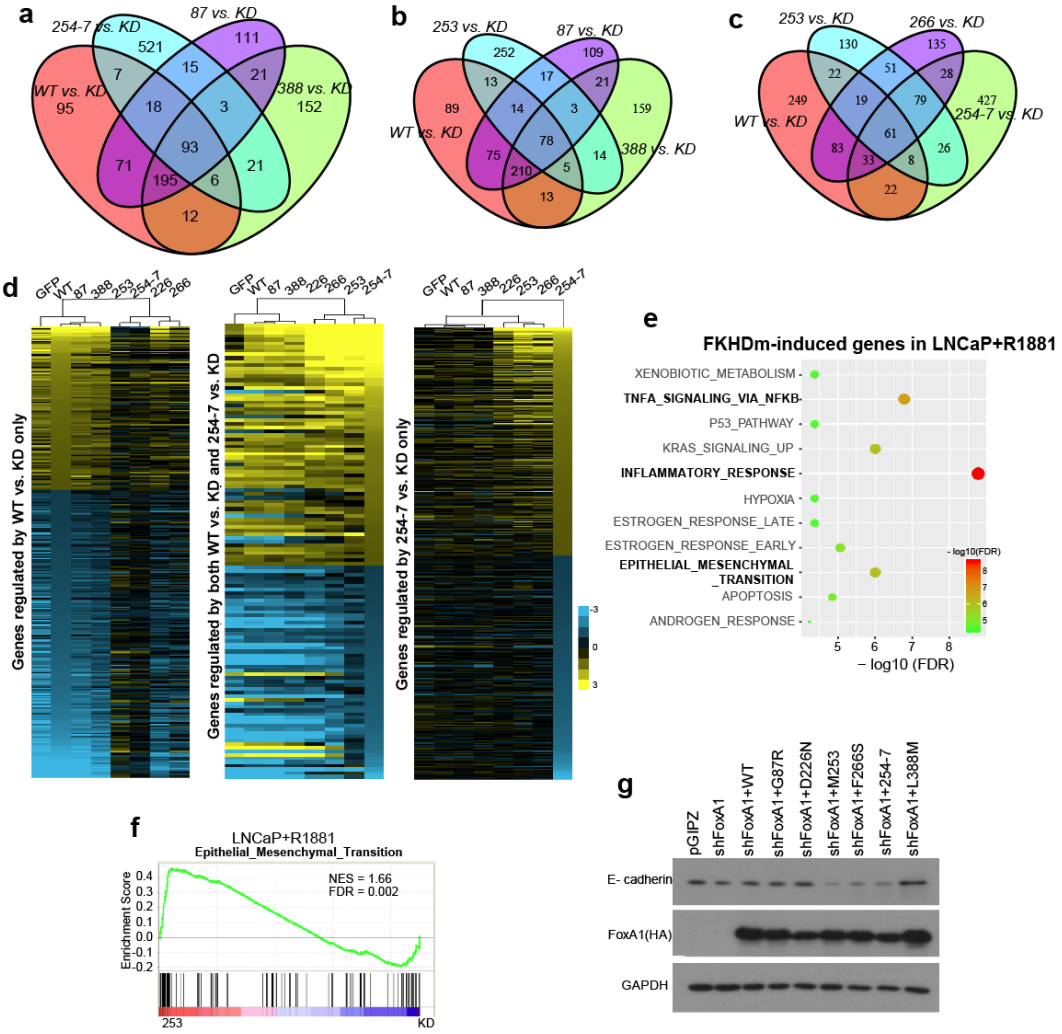
Supplementary information, Fig. S2 FKHDm FOXA1 has an overall reduced chromatin and DNA binding ability. **a** Average FOXA1 WT and mutant ChIP-seq intensities around FOXA1-binding sites (± 2 kb) that were identified in WT only, 254-7 only (FKHDm-only), or shared in both cells. Inset: zoomed-in plots. **b, c** Genome browser tracks showing HA-FOXA1 occupancy on KLK3 and TGFB3 enhancers in FOXA1-WT, KD, and mutant LNCaP cells. Y-axis: (**b**) 0 to 225 and (**c**) 0 to 245. **d, e** Average ChIP-seq intensity around endogenous FOXA1-binding sites (± 2 kb)[1] in hormone-deprived LNCaP cells stimulated with androgen R1881 (**d**) or ethanol (Ethl) (**e**) for 24 h. **f** Western blot analysis of chromatin-free AR and ectopic FOXA1 using anti-AR and anti-HA antibodies, respectively. GAPDH was used as a loading control. **g, h** Nuclear lysates from FOXA1 WT or mutant 293T cells were subjected to fractionation followed by immunoblotting of ectopic FOXA1 levels in chromatin-bound (**g**) and chromatin-free fractions (**h**) using anti-HA antibodies. H3 and GAPDH were used as controls for chromatin and nucleus and cytoplasm factions, respectively.

Fig. S3



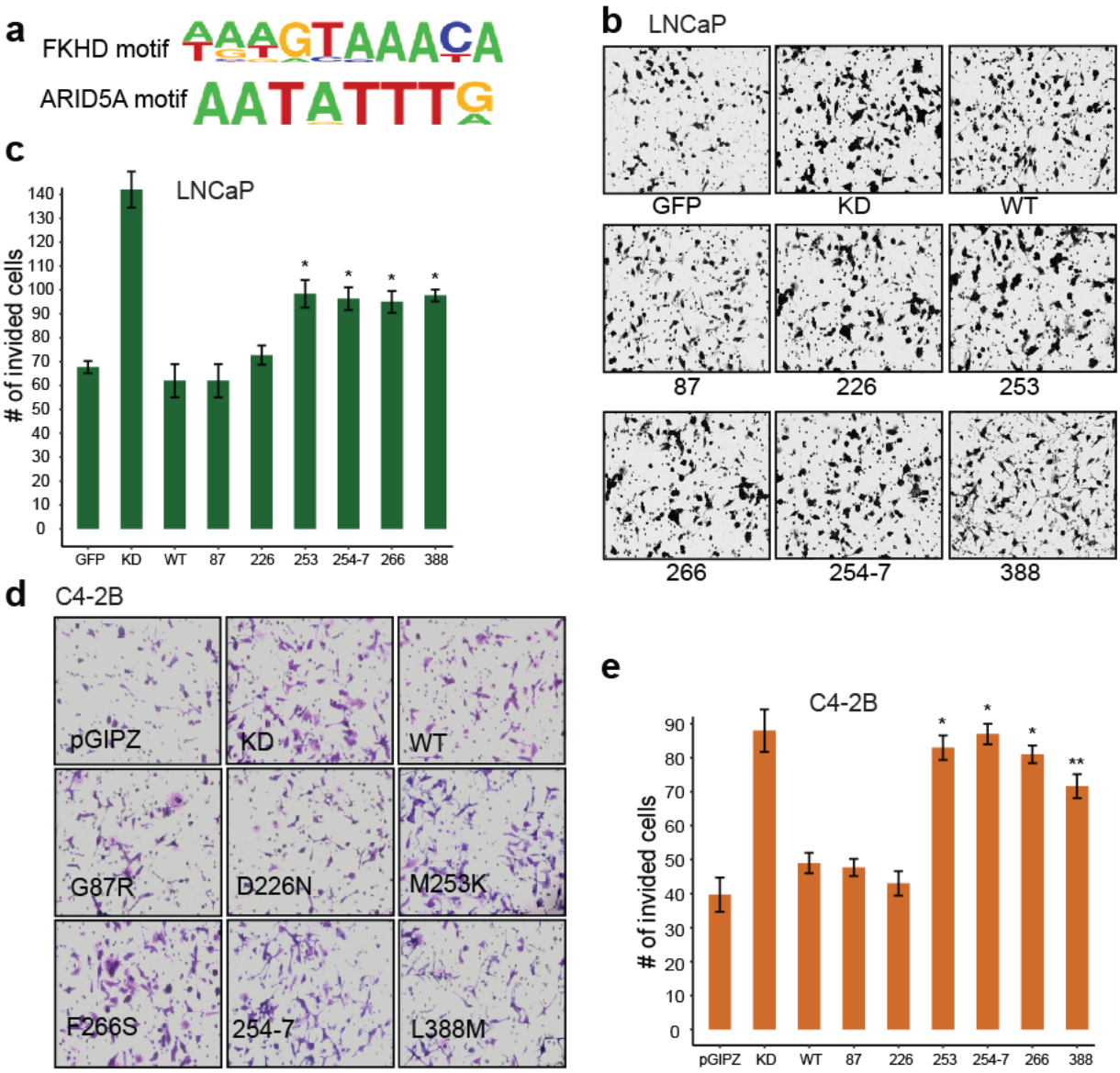
Supplementary information, Fig. S3 FKHDm reduced chromatin-bound AR and inhibited androgen response. **a** Average AR ChIP-seq intensity around FOXA1-binding sites (± 2 kb) identified in WT and/or 254-7 cells. **b, c** Genome browser tracks showing AR occupancy on KLK3 and TGFB3 enhancers in FOXA1-WT, KD, and mutant LNCaP cells. Y-axis (**b**) 0 to 600 and y-axis (**c**) 0 to 120. **d** Overlap of FOXA1- and AR-binding sites. **e, f** Average AR ChIP-seq intensity around AR-only or FOXA1-co-occupied AR-binding sites (± 2 kb) in various cells. **g, h** LNCaP cells were infected with GFP-tagged WT and FKHDm FOXA1 plasmids and stained for AR using an anti-AR antibody, which showed co-localization with FOXA1 (GFP) (**g**). To remove the proteins that were loosely bound or not bound on chromatin, the cells were pre-treated with CSK buffer (**h**) before IF co-staining. **i**, **j** Heatmap showing the expression of androgen-induced genes in control (GFP), FOXA1 knockdown (KD) cells with or without WT or mutant FOXA1 rescue. Cells were hormone-starved for 3 days before stimulation by androgen R1881 (**i**) or ethanol (**j**) for 24 h. Hierarchical clustering was utilized to group genes and samples.

Fig. S4



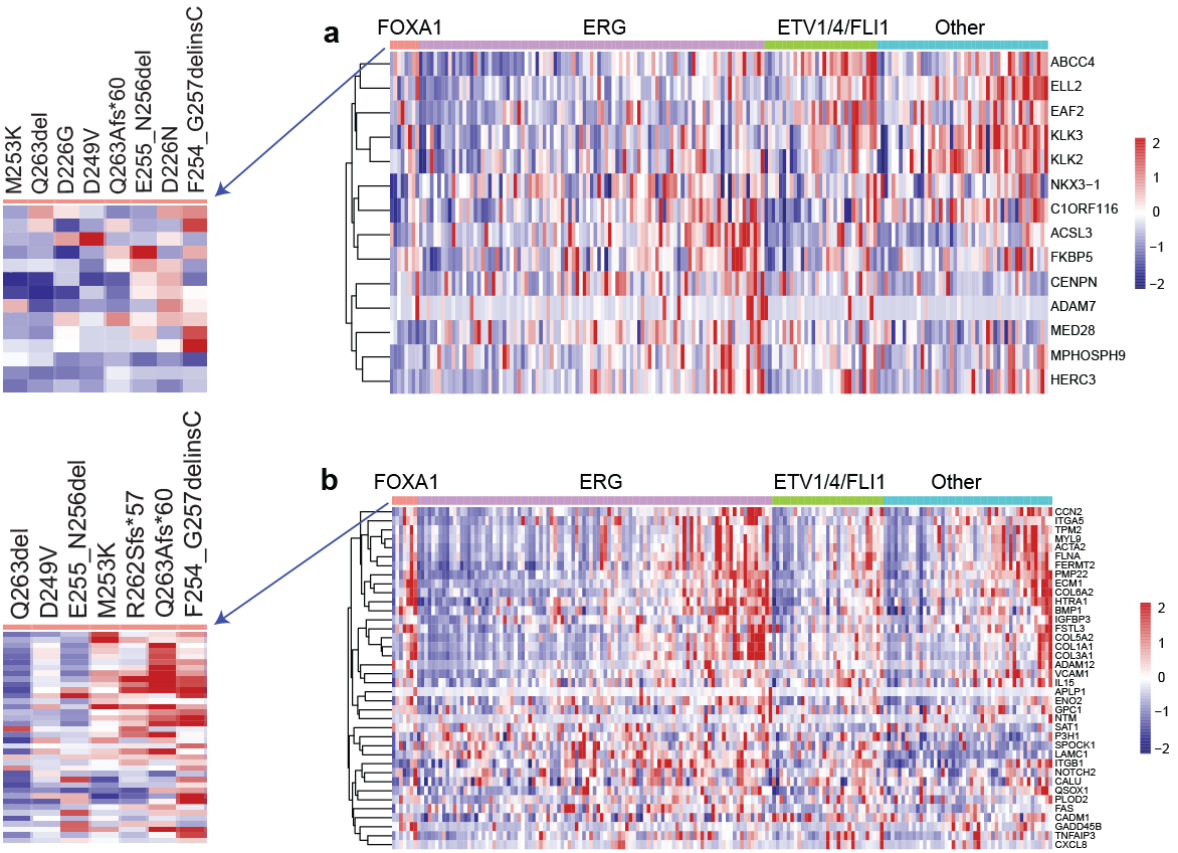
Supplementary information, Fig. S4 FKHDm FOXA1 regulates a unique transcriptome that promotes EMT and cell invasion. **a-c** Venn diagram showing genes that are differentially regulated (fold change >= 2) by re-expression of WT or FKHDm FOXA1 in LNCaP cells with FOXA1 KD. **d** Heatmap showing the expression of the genes that were regulated by WT only (left panel), 254-7 only (right panel), or both (middle panel) in all cell lines as compared to FOXA1 KD cells. Values shown are log2 fold change in the indicated sample vs KD cells. **e**. Gene set enrichment analysis was used to compare curated hallmark gene sets with the genes that were regulated by 253 vs KD cells. Cells were hormone-starved for 3 days and then treated with 1 nM R1881 for 24 h. Y-axis shows the enriched hallmark concepts, while X-axis indicates FDR values of the enrichment. Molecular concepts shown in bold are related to inflammatory response or EMT. **f** GSEA of EMT gene set in 253 vs FOXA1 KD LNCaP cells. **g** Immunoblot analysis of epithelial marker E-cadherin expression in LNCaP cell line panels. GAPDH was used as a loading control.

Fig. S5



Supplementary information, Fig. S5 FKHDm FOXA1 increased LNCaP and C4-2B cell invasion. **a** De novo motif analysis identified FKHD as the most enriched motif in both WT and FKHDm FOXA1 binding sites. In the latter, another strongly enriched motif, the ARID5A motif, was found, second to the FKHD motif. **b, c** LNCaP cell panel was subjected to cell invasion assays (**b**) and data were quantified (**c**). **d, e** C4-2B cell panel was subjected to cell invasion assays (**d**) and data were quantified (**e**). *$P < 0.005$, **$P < 0.001$, all samples compared to WT.

Fig. S6



Supplementary information, Fig. S6 Expression of AR signature and hallmark EMT genes in TCGA PCa dataset. **a** The expression of AR signature genes were obtained from TCGA[2], except the removal of four genes for the following reasons. TMPRSS2 gene is deleted in tumors with TMPRSS2-ERG gene fusion and thus has low expression in these tumors independent of AR signaling. Our data showed that NNMT and MAF were directly bound and repressed by FOXA1, a function that was lost in FKHDm, leading to their up-regulation. On the other hand, GNMT is a FOXA1-induced gene and is further induced by FKHDm, potentially being one of the gain-of-function targets of FKHDm. The expression of these four genes thus does not simply reflect AR signaling and were removed from the heatmap analysis. TCGA RNA-Seq dataset of PCa samples were obtained from cBioportal. Nine samples were found to have FKHDm mutation. One sample was removed due to outlier FOXA1 expression and the left eight samples were compared with samples of ERG fusion, ETV1/4/FLI1 fusion, and other. Samples were clustered by rows (genes). Within each group, columns (patient samples) were ordered by AR output score (sum of Z-scores of all genes for each sample) from low to high (left to right). **b** EMT genes were curated from hallmark gene sets. Out of the nine samples with FKHDm FOXA1, there were two D226 mutations, which does not increase EMT and thus were removed from the analysis. The rest seven FKHDm samples were compared with other groups based on EMT gene expression.

Supplementary information, Materials and Methods

Plasmids and shRNA

Human FOXA1 cDNA was amplified by reverse transcription PCR from LNCaP cells. C-terminal 2×HA-tagged full-length FOXA1 was cloned into the entry vector pCR8/GW/TOPO (Thermo Fisher). The pCR8-FoxA1 mutants (G87R, D226N, M253K, 254-257FENG>C(del), F266S, L388M) were generated using QuickChange II Site Directed Mutagenesis Kit (Agilent Technologies). Lentiviral overexpression constructs were generated by LR recombination between pCR8-FOXA1 constructs and pLenti-CMV-puro-Dest (Addgene plasmid 17293). The pGIPZ lentiviral shRNAmir targeting FOXA1 (Clone ID# V2LHS_16780) and control vector were purchased from Open Biosystems. For production of lentivirus, HEK293T cells were transfected by the PEI transfection method (Sigma) using 2 μg of lentiviral vector, 1.5 μg of psPAX2 and 0.5 μg of pMD2G. Supernatant containing lentiviruses was harvested at 48 h after transfection and filtered through a 0.45 μm filter. Lentiviruses, supplemented with 8 μg/mL polybrene, were used to infect cells. 48 h after infection, cells were selected with 2 μg/mL puromycin.

RNA Isolation and Quantitative RT-PCR

Total cellular RNAs were isolated using the Invitrogen Trizol reagent. For cDNA synthesis, 500 ng of total RNA were reverse transcribed with qScript cDNA SuperMix (Quanta BioSciences). qRT-PCR analysis was done using a StepOnePlus™ Real-Time PCR System (Applied Biosystems). Results were analyzed using StepOne Software v2.1 (Applied Biosystems), and the relative expression of mRNA was determined using GAPDH as the loading control. Data are from triplicate samples.

Gene expression microarray analysis

LNCaP cells were infected with control (GFP) or FOXA1 overexpression lentiviruses (WT, G87R, D226N, M253K, 254-257FENG>C(del), F266S, L388M), followed by infection of FOXA1 knockdown lentivirus a day after. 2 μg/mL puromycin were added 48 h post infection. Hormone starvation for 3 days, treat with or without 1 nM R1881 for 24 h. Total RNAs were isolated using TRIzol reagent (Invitrogen). The integrity of the RNA was monitored using the Bioanalyzer 2100 (Agilent). Microarray profiling was conducted using the HumanHT-12 v 4.0 Expression BeadChip

(Illumina). Bead-level data were preprocessed using GenomeStudio (Illumina), and the expression values were quantile-normalized using the bead array package from Bioconductor. Differentially expressed genes were identified using a 2-fold cutoff.

ChIP-seq and bioinformatics analysis

ChIP was performed as previous described [3]. LNCaP cells were infected with control (GFP) or FOXA1 overexpression lentiviruses (WT, G87R, D226N, M253K, 254-257FENG>C(del), F266S, L388M) together with shCtr or shFOXA1 knockdown lentivirus. 2 μg/mL puromycin were added 48 h post infection. Hormone starvation for 3 days, treat with or without 10 nM R1881 for 16 h. Antibodies used in this study include HA (Abcam, ab9110), FOXA1(Abcam, ab23738), AR (Santa Cruz, sc-816x). ChIP-seq libraries were prepared according to standard protocols using Bioo Scientific's DNA Sample Kit (Cat# 514101)[3,4] and sequenced using Illumina Hi-Seq platforms. Sequence reads were aligned to the Human Reference Genome (assembly hg19) using Burrows-Wheeler Alignment Tool (bwa) version 0.6.1[5].

Electrophoretic mobility shift assays (EMSA)

A 31-nt oligonucleotide was synthesized to contain FKHD motif:
Sense: 5'biotin-TTACAACAGATTTGTTTACTGTCAAGGACTG 3',
Antisense: 5'biotin-CAGTCCTTGACAGTAAACAAATCTGTTGTAA 3'.
Double-stranded oligonucleotides were generated by incubating complementary oligonucleotides at 95 °C for 5 min followed by gradual cooling to room temperature (RT). EMSA assays were performed using the LightShift Chemiluminescent EMSA kit from Thermo Fisher according to the manufacturer's standard protocol. In brief, nuclear extract of AR or FOXA1 WT or Mutants transfected 293T cells was mixed with Biotin-labeled double-stranded oligonucleotides. The protein and DNA binding reaction products were then run on a 5% TBE gel, transferred to nylon membrane, crosslinked and detected by chemiluminescence.

Co-Immunoprecipitation (Co-IP)

293T Cells were co-transfected with AR and HA-tagged FOXA1 WT and various mutants. Cells were lysed in IP lysis buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-

100, Roche protease inhibitor cocktail) 48 h post transfection. An aliquot of the cell lysate was kept as input for western blot analysis. Cell lysate was first pre-cleared with protein G-magnetic beads at 4 °C for 2 h followed by incubation with HA antibody (1 µg/sample, Santa Cruz, F-7) overnight. Then, protein G-magnetic beads were added and incubated at 4 °C for 2 h and washed four times with IP lysis buffer before boiling in SDS sample buffer. Western blot analysis was performed using AR (1:3000, PG-21, Millipore) and HA antibody (1:5000, Abcam, ab9910) according to protocol as described previously.

Chromatin fractionation

Chromatin was isolated as previously reported with the modifications [1], Briefly, cells were re-suspended in buffer A (10 mM HEPES, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 1 mM DTT, 1× Roche protease inhibitor cocktail) and incubated on ice for 10 min. Then, final concentration of 0.1% Triton X-100 was added to cell suspension and vortexed for 15 s and spun down at 4 °C for 5 min at 1,000× *g*. The supernatant was saved as cytoplasmic fraction. Nuclei pellet was washed once with buffer A, and then resuspended in buffer B (3 mM EDTA, 75mM NaCl, 0.1% TritonX-100, 1 mM DTT, protease cocktail) for 30 min on ice. Nuclei was spun down for 5 min at 1,000× *g* at 4 °C and supernatant was saved as nuclear fraction. Insoluble chromatin was re-suspended in 1× SDS sample buffer.

Immunofluorescence (IF)

FOXA1 WT- or mutant-infected LNCaP cells were incubated with CSK buffer (10 mM PIPES, pH 7.0, 100 mM NaCl, 3 mM $MgCl_2$ and 300 mM sucrose) with 0.1% Triton X-100 for 3 min at RT to remove chromatin-free and loosely bound proteins. Then both pre-extracted and non-pre-extracted cells were fixed with 4% PFA for 10 min at RT, then permeabilized in 0.1% Triton X-100 for 15 min at RT. Wash the cells three times with PBS, followed by incubation with 5% normal goat serum for 30 min at RT. Cells were incubated with anti-AR (Santa Cruz, SC-7305) overnight at 4 °C. Then cells were washed three times with PBS and incubate with secondary antibody for 1 h at RT. Lastly, cells were washed three times with PBS and mounted using Prolong Gold Antifade Reagent (Invitrogen). Images were acquired using a Nikon A1 Confocal Laser Microscope. Representative pictures were selected in each case.

WST-1 cell proliferation and cell invasion assay

Cell proliferation assay was performed using WST-1 (Clontech Laboratories) as described by the manufacturer. In brief, 10,000 cells were seeded in 24-well plate. Cells were then hormone starved for 3 days prior to WST-1 assay. 50 µl of WST-1 was added to the cultures and incubated for 2 h before harvesting, and the absorbance was measured at 450 nm using a spectrophotometer. Cell invasions assays were carried out as previously reported [6]. In brief, the cell suspension containing 300,000 (LNCaP) or 100,000 (PC3) cells/mL in serum-free RPMI medium were prepared, 100 µL of cell suspension were transferred into the upper chamber. The lower chamber contained 500 µL of complete growth medium with 40% FBS. After incubation for 72 h, non-invading cells and matrigel were gently removed using a cotton-tipped swab. The inserts were fixed and stained for 15 min in 25% methanol containing 0.5% Crystal Violet. The images of invaded cells were captured under a bright-field microscope, and the number of invaded cells per field view was counted using the cell counter plugins in Image J.

References:

1 Jin, H. J., Zhao, J. C., Wu, L., Kim, J. & Yu, J. Cooperativity and equilibrium with FOXA1 define the androgen receptor transcriptional program. *Nat Commun* **5**, 3972, doi:10.1038/ncomms4972 (2014).

2 Cancer Genome Atlas Research, N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* **163**, 1011-1025, doi:10.1016/j.cell.2015.10.025 (2015).

3 Yu, J. *et al.* An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* **17**, 443-454, doi:10.1016/j.ccr.2010.03.018 (2010).

4 Zhao, J. C. *et al.* Cooperation between Polycomb and androgen receptor during oncogenic transformation. *Genome Res* **22**, 322-331, doi:10.1101/gr.131508.111 (2012).

5 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:btp324 [pii] 10.1093/bioinformatics/btp324 (2009).

6 Jin, H. J., Zhao, J. C., Ogden, I., Bergan, R. C. & Yu, J. Androgen receptor-independent function of FoxA1 in prostate cancer metastasis. *Cancer Res* **73**, 3725-3736, doi:10.1158/0008-5472.CAN-12-3468 (2013).