Supplementary Methods:

ChIP

LNCaP shControl, shMAP3K7, shCHD1 and shDouble cells were seeded onto 150-mm dishes, 5x 10^6 cells per dish, in standard growth medium. The next day, the medium was changed to phenol red-free RPMI-1640 containing 10% charcoal-stripped FBS and 1% PS. After 72 hours, cells were treated with 1 nM R1881 or 0.1% EtOH for 4 hours. Cells were crosslinked with 1% formaldehyde/PBS for 15 minutes at room temperature followed by 0.125 M glycine for 5 minutes. Dishes were washed twice with cold PBS and collected in Szak RIPA buffer (150 mM NaCl, 1% v/v Nonidet P-40, 0.5% w/v deoxycholate, 0.1% SDS w/v, 50 mM Tris-HCl pH 8.0, 5 mM EDTA) with cocktail protease and phosphatase inhibitors on ice. Lysates were sonicated with Bioruptor for 3-15 minute cycles and centrifuged for 30 minutes. Supernatants were pre-cleared with 20 µl Salmon Sperm DNA Protein A or G agarose (Millipore #16-201, #16-157) for 1 hour with rocking at 4°C. Aliquots of cleared lysate were taken as input controls. Lysates were incubated with 10 µg of AR N-20 antibody (Santa Cruz #sc-816, RRID:AB_1563391) and 50 µl of Salmon Sperm DNA Protein A or G agarose with nutation overnight at 4°C. Beads were washed two times with Szak RIPA buffer, four times with Szak IP wash buffer (100 mM Tris-HCl pH 8.0, 500 mM LiCl, 1% v/v Nonidet P-40, 1% w/v deoxycholic acid), two times again with Szak RIPA buffer and twice with TE buffer. The protein-antibody complexes were eluted and reverse cross-linked in Talianidis elution buffer (70 mM Tris-HCl pH 8.0, 1 mM EDTA, 1.5% w/v SDS), containing 200 mM NaCl, by heating at 65°C overnight. The next morning, supernatants were removed and 20 µg of Proteinase K (Thermo Scientific #EO0491) was added for 1 hour at 37°C, followed by 10 µg of RNAse A (Qiagen #19101) for 30 minutes at 37°C. DNA was purified using the Wizard SV Gel and PCR

Clean-Up System (Promega #A9281), eluted with 50 µl of 10 mM Tris-HCl pH 8, and measured with the Qubit dsDNA HS kit (Invitrogen #Q32851).

Clinical cohort of radical prostatectomy samples (Stanford Database)

Patient cohort selection: The Stanford University database of surgically removed prostates (1986-2003) was previously subjected to a comprehensive histopathologic review. In this current study we aimed to only evaluate high grade cases from this database, defined as having at least 2% total cancer area consisting of Gleason grade 4 or 5. We selected a total of 250 high grade cases with 5+ years of follow-up, 125 with biochemical recurrence (BCR) and 125 without BCR that were matched by age at surgery (± 3 years) and by percent total high grade (± 20 percentage points). 206 of these matched cases, 112 with BCR and 94 without, had sufficient tumor tissue for all histologic analyses.

Sample preparation and clinical features: Each prostatectomy specimen was fixed in formalin, serially blocked at 3-mm intervals, and embedded in paraffin. A 5-mm-thick section was cut from each block and H&E stained for histologic assessment. PCa volume was calculated by tracing the exact tumor outline on each slide, determining the area of tumor at each level of section with a digitizing pad, summing the tumor areas at different levels multiplied by the section thickness, and correcting the volume for tissue shrinkage during processing. Tumor Gleason grade was determined. Every histologic pattern, Gleason grades 1 to 5, present in each cancer was recorded by using all the sections that contained tumor. For each patient, an estimate was made of the percentage of the total cancer represented by Gleason grade 4 and 5 (high

grade). Cribriform patterning was graded for each case as either present or absent in any of the tumor sample. The linear extent of any full-thickness capsular penetration into periprostatic fat was quantified in centimeters parallel to the prostatic capsule. Any positive margins were likewise identified and quantified in centimeters of extent. Percent seminal vesicle invasion was estimated on a single coronal section. Biochemical (PSA) failure is used here as a longitudinal outcome measure. Biochemical failure is defined as two consecutive PSA values above a cutoff point of 0.07 ng/ml for the Tosoh method and 0.2 ng/ml for less sensitive methods. Time to failure is calculated as the number of months between the date of surgery and the first of the two consecutive PSA values that exceeded the cutoff point. Subjects whose PSA level never dropped below the detectable level are classified as PSA failure at time zero.

Immunohistochemistry for ERG, CHD1 and MAP3K7: Five-micron thick paraffin sections were deparaffinized, antigens unmasked and immunohistochemically stained for CHD1 at 1:100 (Sigma Aldrich #HPA022236, RRID:AB_2670915), ERG at 1:100 (Cell Marque (EP111) #434R) and MAP3K7 at 1:50 (Sigma Aldrich # HPA007633, RRID:AB_ 1079359). Antigens for CHD1 were unmasked in Cell Conditioner #2 (Ventana Medical Systems) using the standard protocol. Antigens for MAP3K7 and ERG were unmasked in a sodium citrate solution (10 mM, pH 6.0 + 0.1% Tween 20) and BORG Decloaker (Biocare Medical) for 10 minutes at 110°C (NxGen Decloaker, Biocare Medical), respectively, with an ambient cool down for 10 minutes. Immunodetection was performed on the Benchmark XT immunostainer (Ventana) at an operating temperature of 37°C. CHD1 required additional 3% hydrogen peroxide blocker for 8 minutes. Both CHD1 and MAP3K7 antibodies were incubated for 60 minutes and ERG for 30 minutes. Primary antibodies were detected with an UltraView DAB (Ventana) detection kit. All

sections were counterstained in Harris hematoxylin for 2 minutes, blued in 1% ammonium hydroxide (v/v), dehydrated in graded alcohols, cleared in xylene and cover-glass mounted using synthetic resin. Primary antibodies for CHD1 and MAP3K7 were validated by immunocytochemistry and western blotting of LNCaP cells with knockdown of *CHD1* and *MAP3K7* (Fig. S4A).

Immunohistochemical stains were scored by a genitourinary pathologist (MSL) for each grade pattern observed in the tumor section. Nuclear ERG expression was scored as positive or negative. Since some degree of CHD1 and MAP3K7 staining was present in almost all tumors, CHD1 and MAP3K7 were assigned H-scores as a combination of stain intensity (0=none, 1=weak, 2=moderate, 3=strong) and percent of tumor cells using the following formula: (% tumor cells intensity 3 X 3) + (% tumor cells intensity 2 X 2) + (% tumor cells intensity X 1). At least 200 cells in each grade pattern were counted for each case.