Supplementary Methods

RACE, Quantitative reverse-transcription PCR and immunoblot analysis

5' Rapid amplification of cDNA ends (5' RACE) was performed using the GeneRacer kit (Invitrogen) according to the manufacturer's instructions. The 5' RACE product was gel purified, cloned into the pCR4-TOPO vector and sequenced. Quantitative reverse-transcription PCR (qRT-PCR) was performed using Power SYBR Green PCR master mix (Applied Biosystems) on a 7900HT Fast Real-Time PCR system (Applied Biosystems) using the following primers: GAPDH (Forward: TGCACCACCAACTGCTTAGC, Reverse: GGCATGGACTGTGGTCATGAG), TMPRSS2-ERG (Forward: TAGGCGCGAGCTAAGCAGGAG, Reverse: GTAGGCACACTCAAACAACGACTGG), ERG-WT (Forward: GCTCTAAACAACCTCATCAAAACTACTT, Reverse: CTTAATAGTGCTGGCCATAATGCG), ERG-ALL (Forward: CGCAGAGTTATCGTGCCAGCAGAT, Reverse: CCATATTCTTTCACCGCCCACTCC). For immunoblot analysis, rabbit monoclonal antibodies against ERG (Epitomics, #2805-1) and rabbit polyclonal antibodies against β-Tubulin (Santa Cruz Biotechnology, #sc-9104) were used.

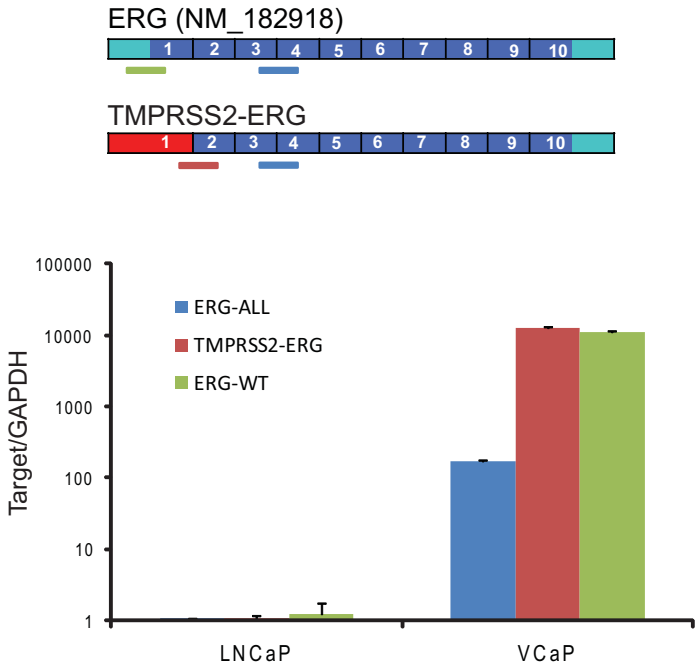
Androgen stimulation and RNA interference

For androgen stimulation experiments, the cells were cultured in 10% charcoal-stripped serum (Hyclone) containing phenol-red free media for 2 days, followed by treatment with the synthetic androgen, R1881 (1 nM) or ethanol for 24 hours. For RNA interference studies, non-targeting siRNA (D-001810-10-20) and ERG-ALL siRNA (D-003886-01) were obtained from Dharmacon Inc. The TMPRSS2-ERG siRNA (CAG GAA GCC UUA UCA GUU G) and ERG-WT siRNA

(UGG UCA GAG AGA AGC AAU A) were custom designed. siRNA transfections were carried out using Oligofectamine transfection reagent (Invitrogen).

Basement Membrane Matrix Invasion Assays

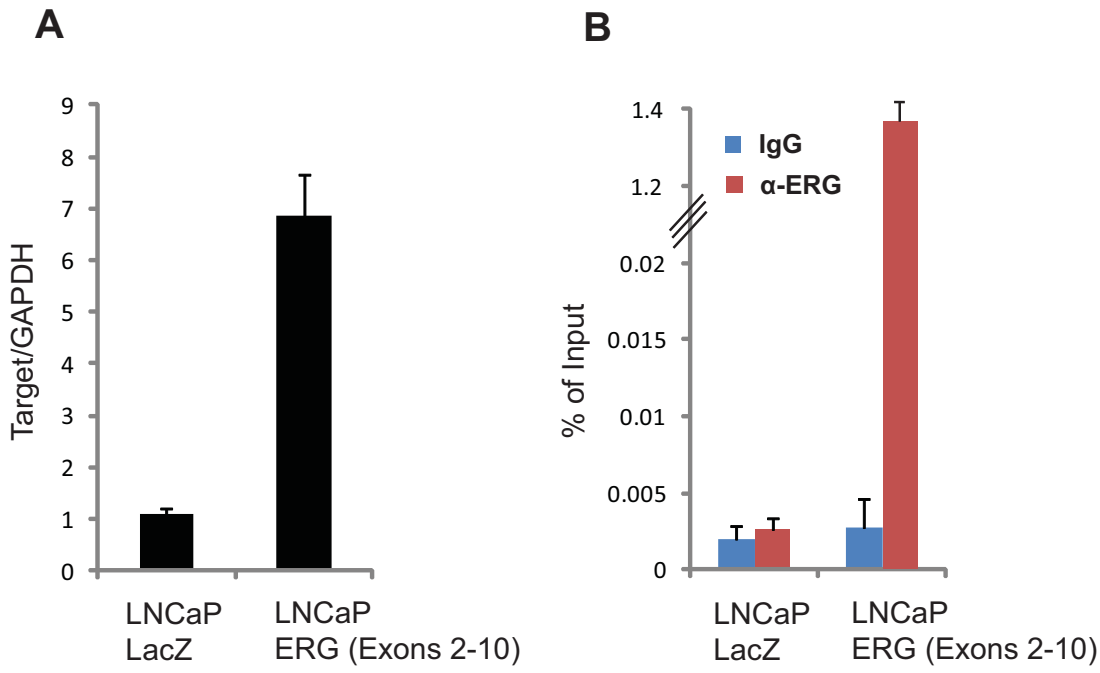
Cells were seeded onto the basement membrane matrix (Millipore) present in the insert of a 24 well culture plate. Fetal bovine serum was added to the lower chamber as a chemoattractant. After 48 hours, the noninvading cells and EC matrix were gently removed with a cotton swab. Invasive cells located on the lower side of the chamber were stained with crystal violet, air dried and photographed. For colorimetric assays, the inserts were treated with 150 μl of 10% acetic acid and the absorbance measured at 560nm using a spectrophotometer.



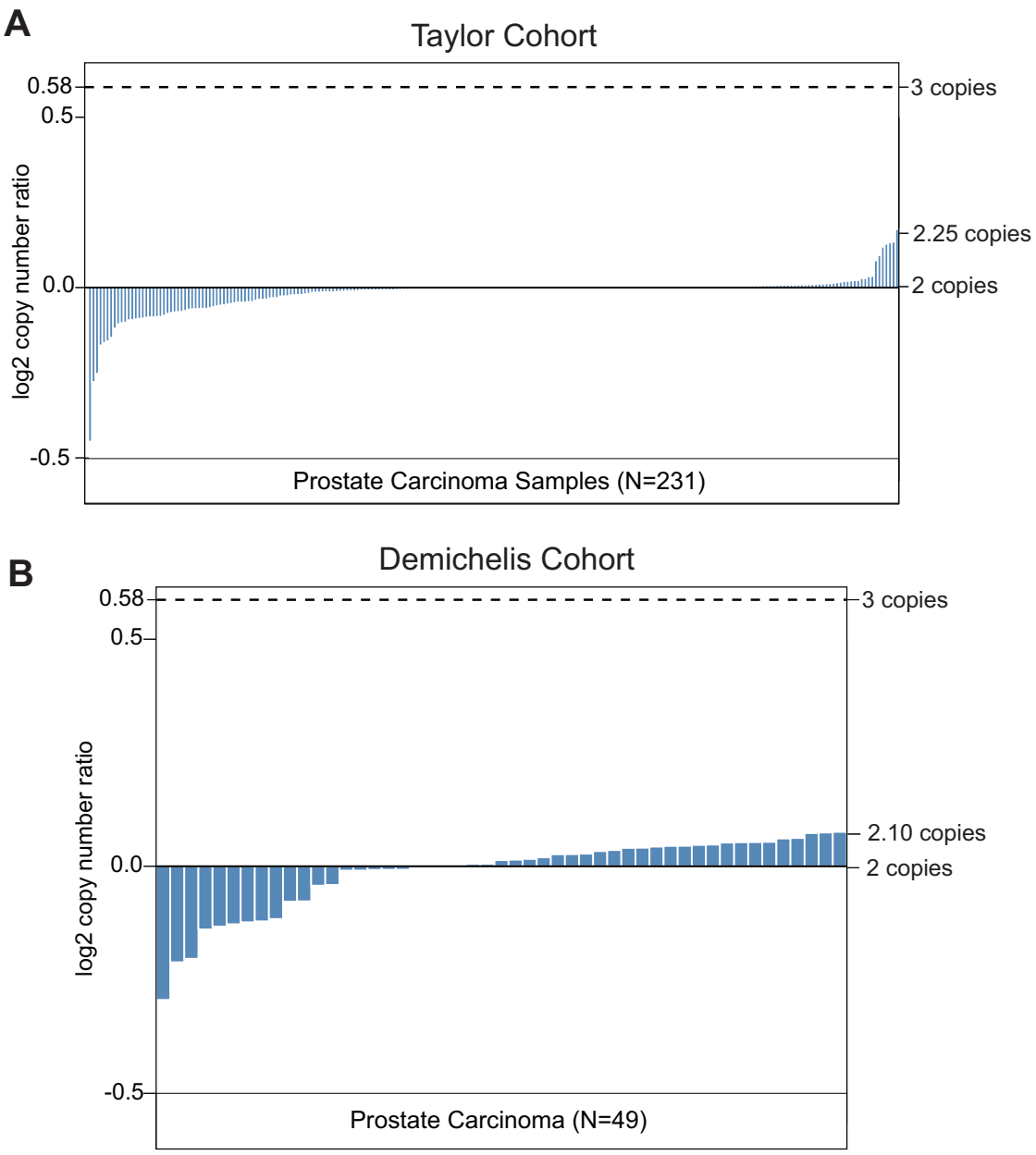
Supplementary Figure 1. High levels of *TMPRSS2-ERG* and wild-type *ERG* transcripts in VCaP cells. Schematic representation of the wild-type *ERG* (ERG-WT, accession: NM_182918) and *TMPRSS2-ERG* transcript (top panel). The locations of primers for gene expression analysis (solid line) are shown. The LNCaP cells that do not express ERG and are represented as controls.

GCANCNNATTACCCTTCACTAAAGGGACTAGTCCTGCAGGGTTTAAACGAATTCGCCCTTCGACTGGAGCACGAGGACAC
TGACATGGACTGAAGGAGTAGAAACTGTTATTGAACATGGCCATCTATTAACATGAAATAAACTCCACCTGCAAAGTTTC
CTTTTAGGAGAAAGGAGCTCAAGTCCTCTGTCCTAGTAATTCTGCGAAGAATTGGGCTGCGAAGGGCTGACACCCAGGGC
AGTGGGGTGCCCTCAACGCCTTGGCTCCAAGGTGCTTGCAGCCCGTGTGAAGCCATGCGGAGTCCCATCACATCTGCATC
CTGCCTTGGTCTCAGTGCAGTGTGTCACACGGCGGCACACCGGTCACACCTGCTGTGGCAGTCCCCATCCCAATTGATGC
CTTTTCTTTTTTTAATTAAAAAAAAAAAATCTCCTGGCTGCATATCATCAAGAAAAACAACATTGTTTTCAGGGTCTCTT
ATCTCTCCACTCCCAAATATCCTGTTTCTTTGGGCTGGACAGCCAGGACCAATAGAAACCTCTTCCTGCCATTGGCTGAC
TTCATTTCCCAGACTTAGCACAATCTCATCCGCTCTAAACAACCTCATCAAAACTACTTTCTGGTCAGAGAGAAGCAATA
ATTATTATTAACATTTATTAACGATCAATAAACTTGATCGCATTATGGCCAGCACTATTAAG

Supplementary Figure 2. 5’ RACE analysis of wild-type *ERG* (NM_182918) in VCaP cells. The complete 5’ end of the first exon of wild-type *ERG* transcript in VCaP cells was determined by rapid amplification of cDNA ends (RACE) method. The sequences highlighted yellow and green represent the vector backbone (pCR4-TOPO) and GeneRacer Oligo sequence respectively.



Supplementary Figure 3. Adenovirus mediated ectopic over-expression of TMPRSS2-ERG fusion product (comprising of ERG exons 2 to 10) in LNCaP cells is associated with the up-regulation of wild-type *ERG* transcript (left panel), and binding of ERG protein to the promoter of wild-type *ERG* locus (right panel).



Supplementary Figure 4. DNA copy number studies suggest that ERG is not amplified in prostate cancer. Two published prostate cancer copy number datasets were screened for ERG gene amplification using the Oncomine database. A, Barplot showing ERG copy number changes for the 231 samples in the Taylor et. al. dataset. Values are plotted in log2 copy number ratio units. The vast majority of samples exhibited minimal or no ERG amplification. The largest log2 copy number ratio observed in any sample was 0.17, corresponding to an absolute copy number of 2.25. B, Barplot showing ERG copy number changes in the Demichelis et. al. study of 49 samples. No samples appeared to harbor ERG amplification in this study.

Sample Code	TMPRSS2-ERG	ERG-WT	Age	Serum PSA	Clinical Stage	Race	Gleason 1	Gleason 2	Gleason 3	Gleason score	Maximum tumor dimension (cm)	Extra prostatic extension (1 yes)	Surgical margin (1 yes)	Seminal ves. invasion (1 yes)
PCA-01	–	–	51	14.9	T1c	Black	3	4		7	1.2	1	1	0
PCA-02	–	–	75	6.1	T1c	White	3	4		7	2.4	0	1	0
PCA-03	–	–	61	5	T1c	White	3	3		6	0.8	0	0	0
PCA-04	–	–	62	5.5	T1c	White	3	3		6	2.3	0	0	0
PCA-05	–	–	77	4.8	T2a	White	3	4		7	1.1	0	0	0
PCA-06	–	–	58	4.8	T1c	White	3	4		7	1.6	0	0	0
PCA-07	–	–	54	10.1	T1c	White	4	3		7	1.5	0	0	0
PCA-08	–	–	58	8	T1c	White	3	4		7	3.5	0	1	0
PCA-09	–	–	64	46.7	T2b	White	4	5		9	3.5	1	1	1
PCA-10	–	–	74	7.3	T2b	White	4	3	5	7	1.1	0	0	0
PCA-11	–	–	59	5.2	T1c	White	3	4		7	1.1	0	0	0
PCA-12	–	–	57	4.2	T1c	Unknown	3	4		7	2.2	0	0	0
PCA-13	–	–	58	7.6	T2a	Black	3	4		7	1.8	0	0	1
PCA-14	–	–	62	8.6	T1c	White	3	4		7	1.4	0	0	0
PCA-15	–	–	52	4.13	T1c	White	3	4		7	2.4	0	0	0
PCA-16	–	–	55	16.1	T2a	White	3	4		7	2.5	0	0	0
PCA-17	–	–	57	4.4	T1c	White	3	4		7	1.5	0	1	0
PCA-18	–	–	68	3.4	T1c	White	3	3		6	1.8	0	0	0
PCA-19	–	–	62	8.3	T1c	White	3	4		7	2.2	0	0	0
PCA-20	–	–	58	6.7	T2b	White	3	4		7	2.1	0	1	0
PCA-21	–	–	61	6.1	T1c	White	4	5		9	3.5	1	1	0
PCA-22	–	–	72	14.6	T1c	White	3	4		7	1.5	0	0	0
PCA-23	–	–	69	6.9	T1c	White	3	4	5	7	1.8	0	0	0
PCA-24	–	–	59	6.9	T2a	White	3	4		7	1.2	0	0	0
PCA-25	–	–	49	5.4	T1c	White	3	4		7	2	1	0	0
PCA-26	–	–	68	4.4	T1c	White	3	4		7	1.9	0	0	0
PCA-27	–	–	72	2.6	T1c	White	3	4		7	1.2	0	0	0
PCA-28	–	–	57	10	T1c	Unknown	4	3	5	7	1.7	1	0	0
PCA-29	–	–	64	11.2	T2c	White	3	4		7	3	0	0	0
PCA-30	–	–	61	13	T2a	White	4	5		9	1.7	1	0	0
PCA-31	–	–	60	2.4	T1c	White	3	4		7	1.9	0	0	0
PCA-32	–	–	64	21	T1c	Black	3	4		7	2.5	0	0	0
PCA-33	–	–	55	13.4	T1c	Black	3	4		7	2.3	1	0	0
PCA-34	–	–	51	13.7	T1c	White	3	4		7	1.6	0	0	0
PCA-35	–	–	70	10.22	T1c	White					4	1	1	0
PCA-36	+	–	51	7.3	T1c	Black	3	4		7	2.5	0	0	0
PCA-37	+	–	52	7.6	T1c	Unknown	3	4		7	1.8	1	0	0
PCA-38	+	–	52	9.7	T1c	White	4	3	5	7	1.6	0	0	0
PCA-39	+	–	59	18.2	T2a	White	4	3	5	7	2.1	1	0	0
PCA-40	+	–	57	1.5	T2a	White	3	4		7	1.5	0	0	0
PCA-41	+	–	64	6.9	T2b	White	4	3		7	1.8	1	0	0
PCA-42	+	–	69	8.1	T1c	White	4	3		7	2.6	1	1	0
PCA-43	+	–	62	8	T1c	White	4	3		7	1.4	0	0	0
PCA-44	+	–	68	5.6	T1c	White	3	4		7	1.1	0	0	0
PCA-45	+	–	63	6.4	T1c	White	3	4		7	0.9	0	0	0
PCA-46	+	–	71	14.4	T1c	White	4	3		7	2.1	0	0	0
PCA-47	+	–	69	7	T1c	White	3	4	5	7	1.7	1	1	0
PCA-48	+	–	60	9.6	T1c	White	4	3		7	4	1	1	1
PCA-49	+	–	52	8.6	T1c	White	3	4		7	1.7	0	0	0
PCA-50	+	–	63	10.6	T2b	White	4	3		7	3	1	1	1
PCA-51	+	–	51	6.1	T2b	White					1.4	1	0	0
PCA-52	+	–	62	6.1	T1c	White	3	4		7	1.8	1	0	1
PCA-53	+	–	57	4.3	T1c	White	3	4		7	1.5	0	0	0
PCA-54	+	–	59	4	T1c	White	3	4		7	2.3	0	0	0
PCA-55	+	–	73	4.4	T1c	White	4	4		8	2.5	0	0	0
PCA-56	+	+	55	7	T1c	White	3	4		7	2.3	0	1	0
PCA-57	+	+	61	3.5	T2b	White					3.3	1	1	1
PCA-58	+	+	49	9.2	T2c	White	3	4		7	2	1	0	0
PCA-59	+	+	49	4.5	T2a	White	3	4		7	3.5	0	0	0
PCA-60	+	+	58	12	T1c	White	4	3		7	2.2	1	1	1
PCA-61	+	+	51	6.1	T1c	Unknown	3	4		7	2.5	0	0	0
PCA-62	+	+	66	6.8	T1c	White	3	4		7	2	0	0	0
PCA-63	+	+	48	8.4	T1c	White	4	3		7	0.8	0	0	0
PCA-64	+	+	69	7.4	T1c	White	3	4		7	1.5	0	1	0
PCA-65	+	+	43	8	T1c	White	3	4		7	1.1	0	0	0
PCA-66	+	+	63	5.2	T1c	Black	3	4		7	2	0	0	0
PCA-67	+	+	55	8.1	T2a	White	3	4		7	1.5	0	0	0

Supplementary Table 1. Summary of clinicopathologic features, TMPRSS2-ERG and wild-type ERG status in the 67 clinically localized prostate cancers.