Role of the *TMPRSS2–ERG* Gene Fusion in Prostate Cancer[1,2,3]

Scott A. Tomlins[*,4], Bharathi Laxman[*,4], Sooryanarayana Varambally[*,†,4], Xuhong Cao[*], Jindan Yu[*], Beth E. Helgeson[*], Qi Cao[*], John R. Prensner[*], Mark A. Rubin, Rajal B. Shah[*,†,§], Rohit Mehra[*,†] and Arul M. Chinnaiyan[*,†,§,¶]

*Michigan Center for Translational Pathology, Department of Pathology, †Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109; ‡Department of Pathology and Laboratory Medicine, Cornell University, New York, NY 10021; §Department of Urology, ¶Bioinformatics Program, University of Michigan Medical School, Ann Arbor, MI 48109

Abstract

TMPRSS2–ERG gene fusions are the predominant molecular subtype of prostate cancer. Here, we explored the role of *TMPRSS2–ERG* gene fusion product using *in vitro* and *in vivo* model systems. Transgenic mice expressing the *ERG* gene fusion product under androgen-regulation develop mouse prostatic intraepithelial *neoplasia* (PIN), a precursor lesion of prostate cancer. Introduction of the *ERG* gene fusion product into primary or immortalized benign prostate epithelial cells induced an invasion-associated transcriptional program but did not increase cellular proliferation or anchorage-independent growth. These results suggest that *TMPRSS2–ERG* may not be sufficient for transformation in the absence of secondary molecular lesions. Transcriptional profiling of *ERG* knockdown in the *TMPPRSS2–ERG*–positive prostate cancer cell line VCaP revealed decreased expression of genes over-expressed in prostate cancer *versus* PIN and genes overexpressed in ETS-positive *versus* -negative prostate cancers in addition to inhibiting invasion. *ERG* knockdown in VCaP cells also induced a transcriptional program consistent with prostate differentiation. Importantly, VCaP cells and benign prostate cells overexpressing *ERG* directly engage components of the plasminogen activation pathway to mediate cellular invasion, potentially representing a downstream ETS target susceptible to therapeutic intervention. Our results support previous work suggesting that *TMPRSS2–ERG* fusions mediate invasion, consistent with the defining histologic distinction between PIN and prostate cancer.

Neoplasia (2008) 10, 177–188

Introduction

Based on a bioinformatics strategy that nominated genes showing high expression in a subset of cancer cases, we identified fusions of the 5′-untranslated region of *TMPRSS2* (21q22) to *ERG* (21q22), *ETV1* (7p21), *ETV4* (17q21), or *ETV5* (3q27) in prostate cancer cases that over-expressed the respective *ETS* family member [1–3]. *TMPRSS2–ERG* fusions are the most predominant molecular subtype, with multiple studies showing that approximately 50% of prostate cancers from prostate-specific antigen (PSA) screened surgical cohorts are *TMPRSS2–ERG* fusion-positive, and greater than 90% of prostate cancers over-expressing *ERG* harbor *TMPRSS2–ERG* fusions [2,4–18].

As *TMPRSS2* had previously been characterized as an androgen-regulated gene [19], and *TMPRSS2* only contributes untranslated sequence to many *TMPRSS2–ERG* transcripts, we hypothesized that the androgen responsive regulatory elements of *TMPRSS2* drive *ERG* over-expression in fusion-positive cases. In support of this hypothesis, we observed that treatment of the *TMPRSS2–ERG*–positive prostate cancer cell line VCaP with the synthetic androgen R1881 resulted in increased expression of the *TMPRSS2–ERG* [2,20] fusion product. Additionally, castration of mice with androgen-dependent *TMPRSS2–ERG*–positive xenografts resulted in decreased expression of *ERG* in the xenograft [21].

Following the identification of *TMPRSS2* fusions to *ERG*, *ETV1*, and *ETV4*, we recently discovered additional 5′ fusion partners involved in *ETV1* and *ETV5* gene fusions, including the 5′ untranslated regions from *SLC45A3*, *HERV–K_22q11.3*, *C15ORF21*, and *HNRPA2B1* [3,22]. Presently, these additional 5′ partners have only been identified in *ETV1* and *ETV5* fusions, and it is unknown if they can fuse with *ERG* (in rare *TMPRSS2–ERG*–negative cases with *ERG* outlier expression) or additional ETS family members. *ETV1* and *ETV5* gene fusions are relatively rare and account for only 2% to 8% of prostate cancers. Interestingly, in these recent studies, we observed that *ETV1* or *ETV5* over-expression induces a cell invasion program [3,22]. Furthermore, androgen regulation and over-expression of the *ETV1* fusion product in the prostate induced mouse prostatic intraepithelial *neoplasia* (mPIN) in mice. Thus, whereas *ETV1* and *ETV5* are rare gene fusions in prostate cancer, it is unknown if the functional role of the most common aberration in prostate cancer, *TMPRSS2–ERG*, is similar. Here, we recapitulated *TMPRSS2–ERG* fusions *in vivo* and *in vitro* and used an integrative expression profiling strategy to determine functional roles for *TMPRSS2–ERG* in prostate cancer.

Materials and Methods

Transgenic ERG Mice

cDNA of *ERG* (exon 2 to base 1533 of NM_182918.2), was amplified by reverse transcription–polymerase chain reaction (RT-PCR) from the VCaP cell line and TOPO cloned into the Gateway entry vector pCR8/GW/TOPO (Invitrogen, Carlsbad, CA), yielding pCR8-*ERG*. 3XFLAG-epitope–tagged construct was generated by PCR using pCR8-*ERG* as the template with the reverse primer encoding a triple FLAG tag before the stop codon. The product was TOPO cloned into pCR8, generating pCR8-3xFLAG-ERG. To generate a prostate-specific *ERG* transgenic construct, 3xFLAG-*ERG* was inserted into pBSII (Stratagene, La Jolla, CA) downstream of a modified small composite probasin promoter (ARR2PB) and upstream of a bovine growth hormone polyA site (PA-BGH) [23,24].

The ARR2Pb-*ERG* plasmid was linearized with *Pvu*I/*Kpn*I/*Sac*II and microinjected into fertilized FVB mouse eggs and surgically transplanted into a pseudopregnant female by the University of Michigan Transgenic Animal Model Core. Transgenic founders were screened by PCR using genomic DNA isolated from tail snips. Multiple ARR2Pb-*ERG* transgenic founders were obtained and crossed with FVB mice, and transgene-positive male mice offspring were sacrificed at various time points.

Prostates from transgenic mice were dissected, stained with hematoxylin and eosin, and evaluated by three pathologists (R.M., M.A.R., and R.B.S.) as described earlier [22,25].

For immunohistochemical detection of Erg-FLAG, the basal cell marker p63, and smooth muscle actin, deparaffinized slides were subjected to microwave-citrate antigen retrieval and incubated with rabbit anti–FLAG polyclonal antibody (1:50 dilution, overnight incubation, #2368; Cell Signaling Technology, Danvers, MA), mouse monoclonal anti–p63 antibody (1:100 dilution, 45 minutes of incubation, MS1081P1; LabVision, Fremont, CA), and mouse monoclonal anti–smooth muscle actin antibody (1:50 dilution, 30 minutes of incubation, M0851; DakoAb, Carpinteria, CA), respectively. Visualization of p63 and SMA was performed using a standard biotin–avidin complex technique using M.O.M. Immunodetection kit (PK2200; Vector Laboratories, Burlingame, CA). FLAG was detected using Envision+System–HRP (DAB) kit (K4011; DakoCytomation, Carpinteria, CA).

Cell Lines and Samples

The benign immortalized prostate cell line RWPE was obtained from the American Type Culture Collection (Manassas, VA). Primary benign prostatic epithelial cells (PrEC) were obtained from Cambrex Bio Science (Walkersville, MD). VCaP was derived from a vertebral metastasis from a patient with hormone-refractory metastatic prostate cancer [26], and was provided by Kenneth Pienta (University of Michigan).

Prostate tissues were from the radical prostatectomy series at the University of Michigan and from the Rapid Autopsy Program, which are both part of the University of Michigan Prostate Cancer Specialized Program of Research Excellence Tissue Core. All samples were collected with informed consent of the patients and prior institutional review board approval. For all samples and cell lines, total RNA was isolated with Trizol (Invitrogen) according to the manufacturer's instructions.

In Vitro Over-expression of ERG

To generate adenoviral and lentiviral constructs, pCR8-*ERG* and a control entry clone (pENTR-*GUS*) were recombined with pAD/CMV/V5 (Invitrogen) and pLenti6/CMV/V5 (Invitrogen), respectively, using LR Clonase II (Invitrogen). Control pAD/CMV/*LACZ* clones were obtained from Invitrogen. Adenoviruses and lentiviruses were generated by the University of Michigan Vector Core. The benign immortalized prostate cell line RWPE was infected with lentiviruses expressing *ERG* or *GUS*, and stable clones were generated by selection with blasticidin (Invitrogen). Benign PrEC cells were infected with adenoviruses expressing *ERG* or *LACZ*. RWPE cells were also infected with *ERG* or *LACZ* adenoviruses for transient over-expression.

Immunoblot Analysis

Cell lysates transferred to polyvinylidene fluoride membranes were probed with rabbit polyclonal anti-ERG (sc-354; Santa Cruz Biotechnology, Santa Cruz, CA) at 1:500 dilution, mouse monoclonal

anti–matrix metalloproteinase 3 (MMP3) (IM36L; Calbiochem, San Diego, CA) at 1:500 dilution, mouse monoclonal anti-uPA (IM13L, Calbiochem) at 1:500 dilution, and mouse anti–GAPDH antibody (Abcam, Cambridge, MA) at 1:30,000 dilution for loading control.

Proliferation Assay

Cell counts were estimated by trypsinizing the cells and, analysis was done using a Coulter counter (Beckman Coulter, Fullerton, CA) at the indicated time points in triplicate.

FACS Cell Cycle Analysis

Propidium iodide–stained RWPE-*ERG* and RWPE-*GUS* cells were analyzed on a LSR II flow cytometer (BD Biosciences, San Jose, CA) running FACSDivia, and cell cycle phases were calculated using ModFit LT (Verity Software House, Topsham, ME).

Soft Agar Assay

A 0.6% (wt./vol.) bottom layer of low melting point agarose in normal medium was prepared in six-well culture plates. On top, a layer of 0.3% agarose containing 1×10^4 RWPE-*GUS*, RWPE-*ERG*, or *DU145* (positive control) cells was placed. After 12 days, foci were stained with crystal violet and counted.

Invasion Assays

Invasion assays were performed using PrEC and RWPE-*ERG* and -*LACZ* cells (48 hours after infection with adenoviruses), stable RWPE-*ERG* and -*GUS* cells, or VCaP cells as described earlier [22].

For inhibitor studies, amiloride (20 µM; EMD Biosciences, San Diego, CA), MMP3 inhibitor (10 µM; EMD Biosciences), MMP2/9 inhibitor (10 µM; EMD Biosciences), MMP8 inhibitor (10 µM; EMD Biosciences), the pan MMP inhibitor GM 6001 (10 µM; EMD Biosciences), the EWS:FLI inhibitor cytosine arabinoside (250 nM) [27], or vehicle control was added to VCaP and stable RWPE-*ERG* or -*GUS* cells for 24 hours, before trypsinization and seeding for invasion assays. For PAI-1, VCaP and stable RWPE-*ERG* or -*GUS* cells were trypsinized and treated with the indicated amount of recombinant PAI-1 (EMD Biosciences) for 15 minutes at indicated concentrations, before seeding.

ERG, PLAU, and PLAT Knockdown

For short interfering RNA (siRNA) knockdown of *ERG*, *PLAT*, or *PLAU*, the individual siRNA composed of the Dharmacon SMART-pool against *ERG* (MQ-003886-01; Lafayette, CO), *PLAT* (LQ-005999-00), or *PLAU* (LQ-006000-00), were tested for knockdown by quantitative polymerase chain reaction (qPCR), and the most effective single siRNA (*ERG*, D-003886-01; *PLAT*, J-005999-05; *PLAU*, J-006000-07) was used for further experiments. siCONTROL Non-Targeting siRNA #1 (D-001210-01) or siRNA against *ERG*, *PLAT*, or *PLAU* was transfected into VCaP or RWPE-*ERG* cells as indicated using Oligofectamine (Invitrogen). After 24 hours, we carried out a second identical transfection and cells were harvested 24 hours later for RNA isolation, invasion assays, or proliferation assays.

Expression Profiling

Expression profiling was performed using the Agilent Whole Human Genome Oligo Microarray (Santa Clara, CA) according to the manufacturer's protocol [22]. For all hybridizations involving *ERG* over-expression by adenovirus or lentivirus, the reference was the same cell line expressing *LACZ* or *GUS*, respectively. For profiling

of *ERG* knockdown in VCaP, the reference was VCaP treated with nontargeting siRNA. All hybridizations were performed in duplicate with duplicate dye flips, for a total of four arrays, except for transiently expressing RWPE-*ERG*, which consisted of duplicate hybridizations and a single dye flip. Over- and under-expressed signatures were generated by filtering to include only features with significant differential expression (Log ratio, $P < .01$) in all hybridizations and two-fold average over- or under-expression (Log ratio) after correction for the dye flip. For VCaP profiling, all features with significant differential expression (Log ratio, $P < .01$) and Cy5/Cy3 ratios of > or < 1 in all hybridizations were included in the over- and under-expressed signatures, respectively.

Molecular Concepts Analysis

All expression signatures were uploaded into the Oncomine Concepts Map (OCM, www.oncomine.org) [28] as molecular concepts, using all features on the Agilent Whole Human Genome Oligo Microarray as the null set. For the assessment of prostate-specific gene expression, the expO (GSE2109) and Shyamsundar normal tissue [29] data sets were accessed using the Oncomine database. Cancer and normal tissue types are defined in Table W3. For the assessment of prostate cell type expression, the *Prostate cell-specific expression* Affymetrix data set of Oudes et al. [30] was downloaded from the Gene Expression Omnibus (GSE3998). Data are reported as RMA-normalized fluorescent intensities.

Quantitative PCR

Quantitative PCR was performed using Power SYBR Green Mastermix (Applied Biosystems, Foster City, CA) on an Applied Biosystems 7300 Real Time PCR system as described [1,2]. All oligonucleotide primers were synthesized by Integrated DNA Technologies (Coralville, IA) and are listed in Table W2. All reactions were performed in duplicate unless otherwise indicated.

Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed according to published protocols using anti-ERG (sc-354x; Santa Cruz) or rabbit anti-IgG (sc-2027; Santa Cruz) antibodies [31]. For PCR analysis of enrichment of target gene promoters, 2 µl each of input DNA, ERG-enriched, or IgG-enriched DNA were subjected to PCR using Platinum PCR Supermix (Invitrogen) and primers specific for target gene promoters (Table W2).

Results

Transgenic Expression of ERG in the Mouse Prostate Induces mPIN

Fusion transcripts juxtaposing exon 1 of *TMPRSS2* (NM_005656.2) to exon 2 of *ERG* isoform 1 (NM_182918.2; identical to exon 4 of *ERG* isoform 2, NM_004449.3) are the most commonly detected transcripts in *TMPRSS2–ERG*–positive cases (*TMPRSS2–ERGa*) [2,5,9]. Because exon 1 of *TMPRSS2* is entirely noncoding, this fusion transcript likely results in a truncated *ERG* protein product. Thus, we generated transgenic mice expressing the truncated *ERG* product from *TMPRSS2–ERGa* (beginning at exon 2 through the reported stop codon (base 1533) of NM_182918.2, C-terminal FLAG-tagged) under the control of the modified probasin promoter (ARR2Pb-*ERG*) (Figure 1*a*), which drives androgen-regulated transgene expression exclusively
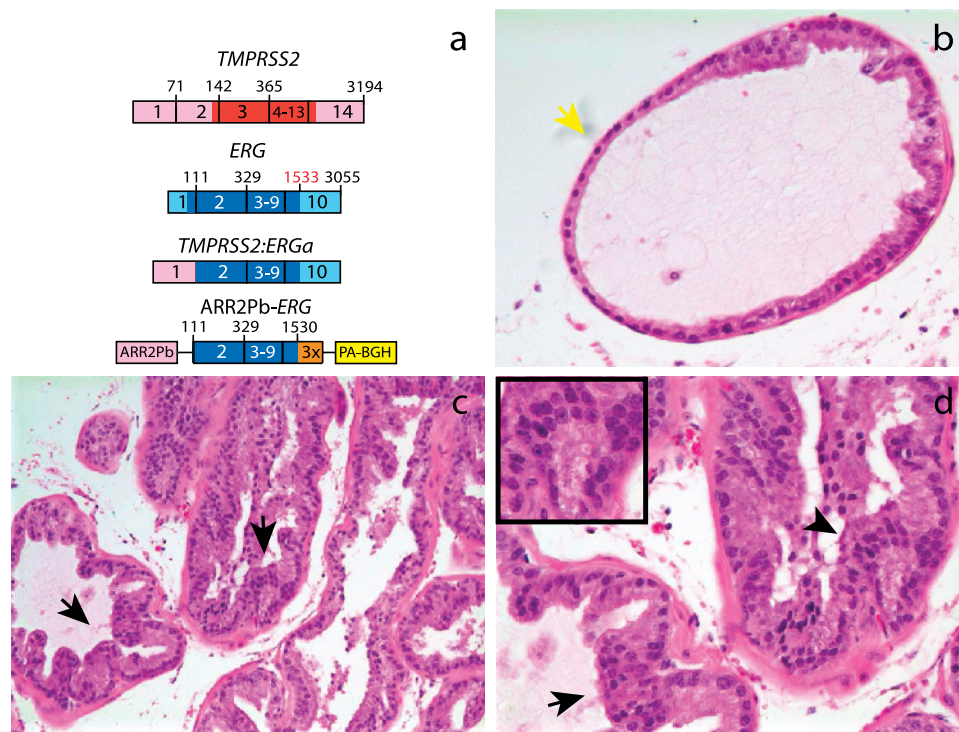
Figure 1. Transgenic mice recapitulating TMPRSS2–ERG in the prostate develop mPIN. (a) To recapitulate TMPRSS2–ERG in vivo, we generated transgenic mice over-expressing the ERG gene fusion product (exons 2 through the reported stop codon; 1533 of NM_182918.2, C-terminal 3X-FLAG epitope tag) with a bovine growth hormone polyA signal (PA-BGH) under the control of the enhanced probasin promoter (ARR2Pb). Mice were sacrificed at 12 to 14 weeks or >20 weeks, and mouse prostatic intraepithelial neoplasia (mPIN) was observed in 4 of 11 ARR2Pb-ERG mice as described in Table W1. Benign epithelia and areas of mPIN are indicated by yellow and black arrows, respectively. (b–d) Hematoxylin and eosin staining of ARR2Pb-ERG prostates for morphologic assessment. Consistent with the focality of mPIN, (b) benign glands and (c and d) mPIN were observed in the ventral prostate (VP) of ARR2Pb-ERG mice. Original magnification: (b) ×400, (c) ×200, and (d) inset showing area of mPIN with macronucleoli, ×400.

in the prostate [23,24]. This transgene is functionally analogous to the TMPRSS2–ERGa fusion product. We obtained multiple ARR2Pb-ERG founders and lines were expanded for phenotypic analysis. By 12 to 14 weeks of age, three of eight (37.5%) ARR2Pb-ERG mice developed mPIN (Table W1 and Figure 1), the candidate precursor lesion of prostate cancer [25].

We observed normal glands in the prostates of ARR2Pb-ERG mice containing focal proliferative lesions displaying nuclear atypia,

including stratification, hyperchromasia, and macronucleoli (Figure 1, b–d), consistent with the definition of mPIN [25]. In 12- to 14-week-old ARR2Pb-ERG mice, foci of mPIN were observed exclusively in the ventral lobe (Table W1). Immunohistochemistry in ARR2Pb-ERG mice demonstrated strong ERG-FLAG expression primarily in mPIN foci and not benign glands (Figure W1, a and b), and qPCR confirmed that transgene expression was limited to the prostate (data not shown).

Figure 2. Over-expression of ERG in RWPE cells increases invasion through the plasminogen activator pathway. (a) To recapitulate TMPRSS2–ERG in vitro, we generated adenoviruses and lentiviruses expressing the ERG gene fusion product (exons 2 through the reported stop codon). (b and c) Infected (b) RWPE and (c) PrEC cells as indicated were assayed for invasion through a modified basement membrane. Photomicrographs of invaded cells are shown below. (d) RWPE-ERG and RWPE-GUS (control vector) cells were profiled on Agilent Whole Genome microarrays and expression signatures were loaded into the Oncomine Concept Map. Molecular concept map analysis of the over-expressed in RWPE-ERG compared to RWPE-GUS signature (ringed yellow node). Each node represents a molecular concept, or set of biologically related genes. The node size is proportional to the number of genes in the concept. The concept color indicates the concept type according to the legend. Each edge represents a significant enrichment ($P < .005$). (e) qPCR confirmation of increased expression of genes involved in invasion. The amount of the indicated gene (normalized to the average of GAPDH and HMBS) in RWPE-GUS (white) and RWPE-ERG (black) is shown. Inset shows immunoblot confirmation of increased expression of PLAU and MMP3 in RWPE-ERG cells. (f) Chromatin immunoprecipitation shows enrichment of ERG binding to the proximal promoters of PLAU and MMP3 compared to IgG control. The promoter of KIAA0089 was used as a negative control. (g) RWPE-ERG cells were treated with PLAU inhibitors amiloride or ectopic PAI-1, MMP inhibitors (including the pan-MMP inhibitor GM-6001), or the EWS:FLI inhibitor ARA-C (EWS:FLI inhibitor) as indicated and assayed for invasion as in c. (h) RWPE-ERG cells were treated with transfection reagent alone (untreated), or transfected with nontargeting, PLAU or PLAT siRNA as indicated and assayed for invasion through a modified basement membrane. For all invasion assays, mean ($n = 3$) ± SEM are shown; *$P < .05$.
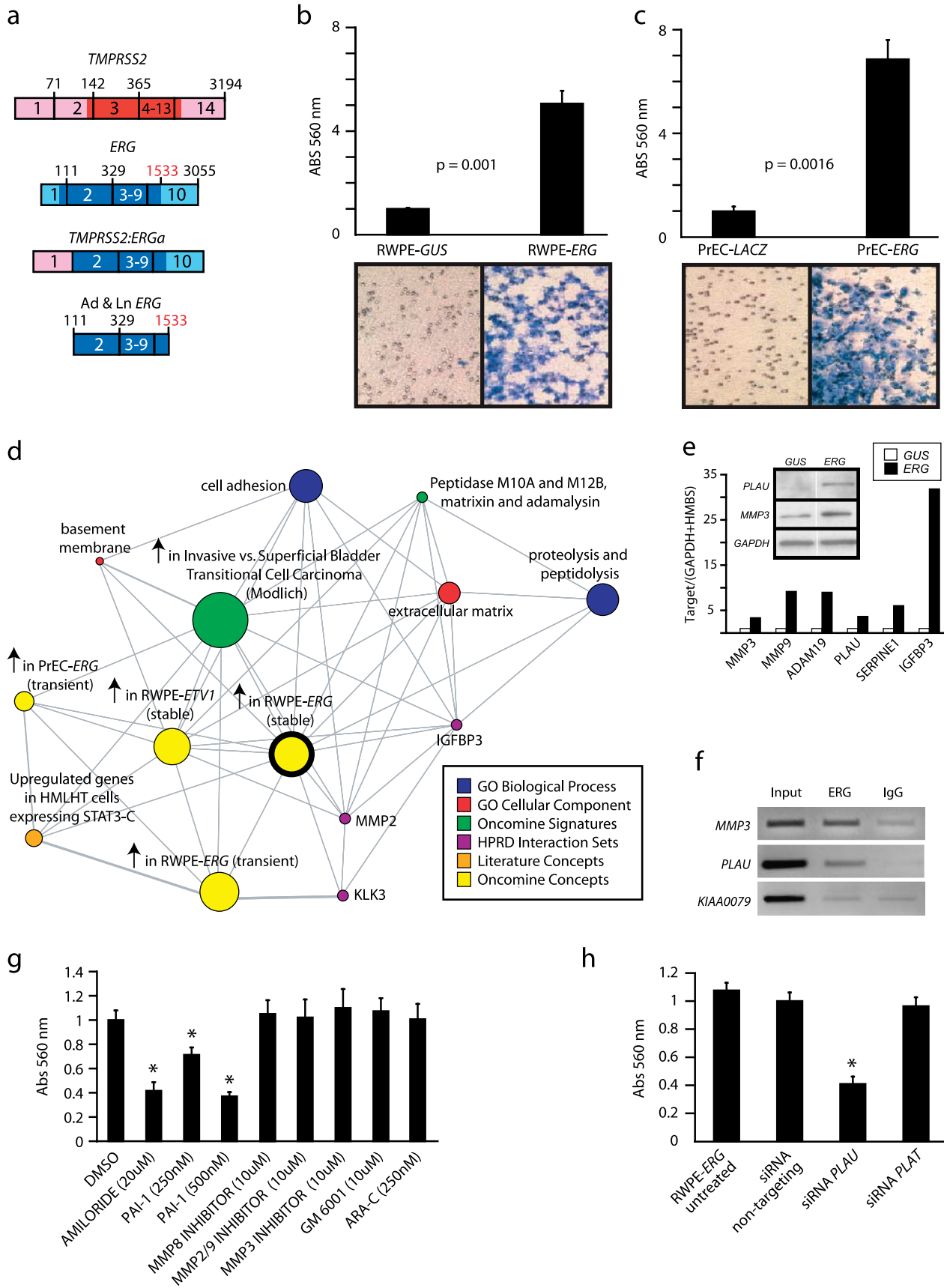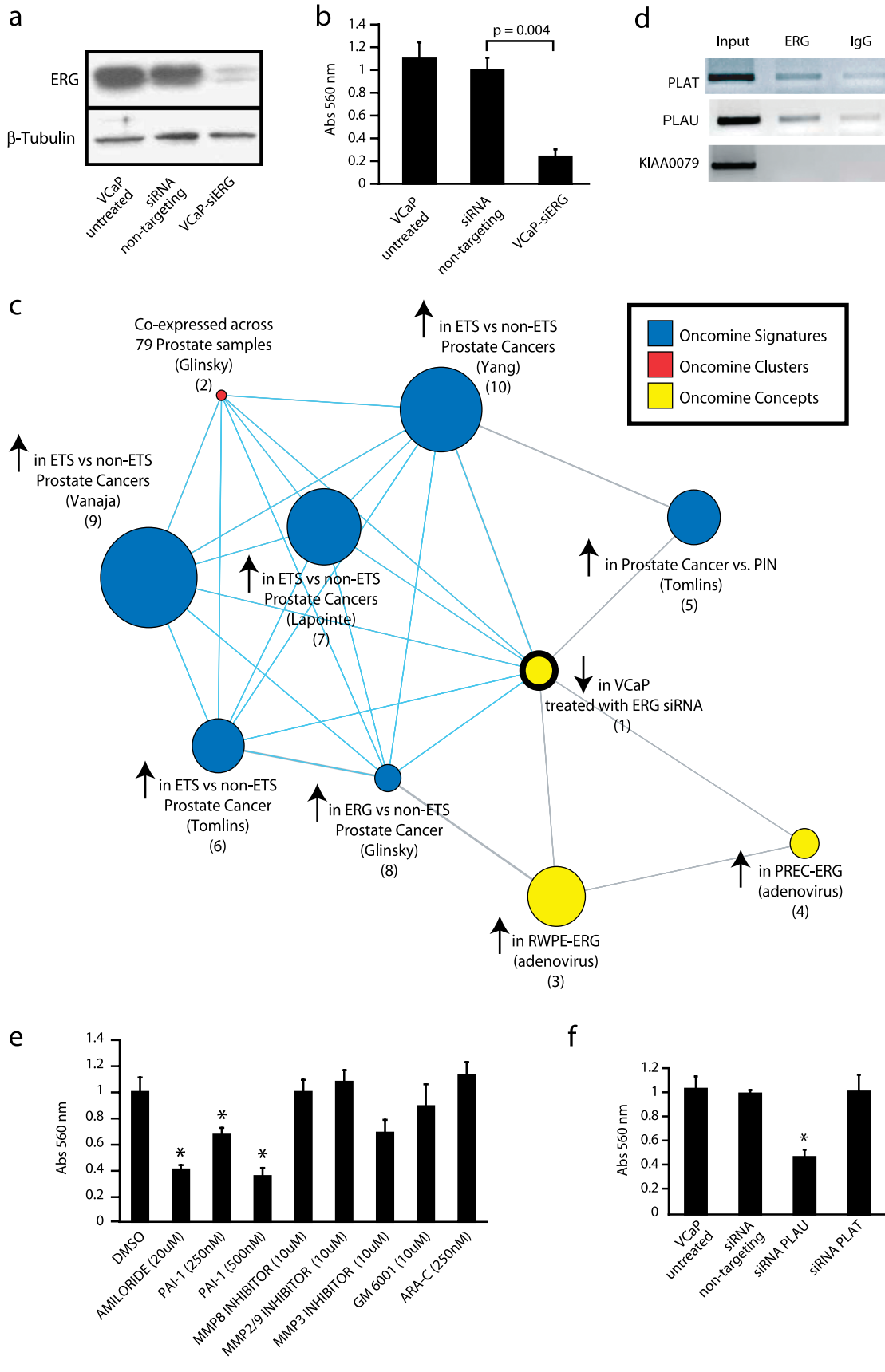
All lesions were confirmed to be *in situ* by the presence of an intact fibromuscular layer, as demonstrated by contiguous smooth muscle actin staining (Figure W1, *c* and *d*). However, immunohistochemistry with the basal cell marker p63 demonstrated loss of the circumferential basal epithelial layer in ARR2Pb-*ERG* mPIN compared to benign glands (Figure W1, *e* and *f*), indicating the disruption of the basal cell layer. Because loss of the basal layer is a hallmark of prostate carcinoma development in both mice and humans [32], ARR2Pb-*ERG* mice will be closely monitored for the development of invasive carcinoma at later time points. Whereas we have not observed progression to invasive carcinoma in ARR2Pb-*ERG* mice, we have only characterized three mice at greater than 20 weeks of age, one of which (33.3%) also had mPIN in both the ventral and dorsolateral lobes (Table W1). These results demonstrate that, although *ERG* induces a neoplastic phenotype in the mouse prostate, providing support for an oncogenic role in human prostate cancer, it is not sufficient for the development of prostate cancer in mice.

ERG Over-expression Induces an Invasion Program In Vitro

Next, we determined the effects of *ERG* over-expression *in vitro*, by generating adenoviruses and lentiviruses that express the same truncated *ERG* product from *TMPRSS2–ERGa* as in the ARR2Pb-*ERG* mice (Figure 2*a*). We infected the benign immortalized prostate epithelial cell line RWPE with lentivirus expressing *ERG* and selected for stable RWPE-*ERG* cells, and transiently over-expressed *ERG* in primary benign prostate epithelial cells (PrEC) by infection with adenovirus expressing *ERG*. By immunoblot analysis, we confirmed the expression of a protein product recognized by a commercial anti–ERG antibody in both RWPE and PrEC (Figure W2).

In both RWPE and PrEC cells, over-expression of *ERG* did not increase proliferation (Figure W2), and *ERG* did not affect the percentage of RWPE cells in S phase by cell cycle analysis (Figure W2*c*). Additionally, soft agar transformation assays showed that *ERG* over-expression was not sufficient to transform RWPE cells (Figure W2*d*). Finally, orthotopic xenograft assays using RWPE-*ERG* cells did not result in tumor formation (data not shown). However, *ERG* over-expression markedly increased invasion in a modified basement membrane invasion assay in both RWPE (5-fold, $P = .001$) (Figure 2*b*) and PrEC cells (6.9-fold, $P = .0016$) (Figure 2*c*). Transient over-expression of *ERG* in RWPE using *ERG* adenovirus similarly increased invasion (Figure W3). These results are similar to over-expression of *ETV1* and *ETV5*, which we have previously shown to increase invasion in PrEC and RWPE cells [3,22].

To investigate the transcriptional program regulated by *ERG*, we profiled stable RWPE-*ERG* and transiently expressing RWPE-*ERG* and PrEC-*ERG* cells using Agilent Whole Genome Oligo Expression Arrays, and identified 865, 854, and 221 features that were over-expressed in the respective cell lines (as described in the Materials and Methods section). We have recently developed a resource termed the Oncomine Concepts Map (OCM, www.oncomine.org) to look for associations between more than 20,000 biologically related gene sets by disproportionate overlap [28,33]. Thus, we uploaded these expression signatures into the OCM to identify transcriptional programs induced by *ERG*. We began by seeding the OCM analysis with the *over-expressed in stable RWPE-ERG* signature. OCM analysis identified the most significantly enriched concept as our previous *over-expressed in stable RWPE-ETV1 signature* [22] [odds ratio (OR) = 59.43, $P = 1 \times 10^{-100}$] (Figure 2*d*), consistent with their similar phenotypes and supporting the functional redundancy of these ETS family members in gene fusions.

The stable RWPE-*ERG* signature also shared significant enrichment with the *over-expressed in transient RWPE-ETV5* (OR = 3.9, $P = 1.2 \times 10^{-9}$), *over-expressed in transient RWPE-ERG* (OR = 19.43, $P = 1.1 \times 10^{-100}$), and *transient PrEC-ERG* (OR = 5.77, $P = 3.1 \times 10^{-10}$) signatures, demonstrating similarities in these transcriptional programs, as well as several molecular concepts related to invasion. These concepts include the Interpro concept of gene products containing *Peptidase M10A and M12B, matrixin or adamalysin domains* (OR = 5.27, $P = .002$), which includes MMPs and a disintegrin and metalloproteinase domains (ADAM), and a signature of genes *over-expressed in benign breast epithelial cells (HMLHT) over-expressing the STAT3-C oncogene* (OR = 4.04, $P = 6.3 \times 10^{-5}$). In this system, *STAT3-C* over-expression did not increase proliferation, but increased invasion in an *MMP9*-dependent manner [34].

ERG-Mediated Induction of the Plasminogen Activator Pathway

We identified several genes over-expressed in RWPE-*ERG* that were present in multiple concepts in this enrichment network and have been directly implicated in the invasion in multiple cancers and models, including the metalloproteinases *MMP3*, *MMP9*, and *ADAM19*, the urokinase plasminogen activator (*PLAU*), and the plasminogen activator inhibitor type 1 (*SERPINE1*, also known as *PAI-1*) [35,36]. Both MMPs and the urokinase plasminogen pathway have been reported to be direct targets of ETS transcription factors [35–37]. By qPCR, we confirmed the over-expression of these genes, as well as the MMP cleavage target *IGFBP3* in RWPE-*ERG* cells (Figure 2*e*).

Figure 3. Knockdown of ERG in VCaP cells attenuates a transcriptional program over-expressed in *TMPRSS2–ETS*–positive prostate cancers. (a) SiRNA knockdown of *ERG* in the *TMPRSS2–ERG*–positive prostate cancer cell line VCaP. VCaP cells were treated with transfection reagent alone (untreated), or transfected with nontargeting or *ERG* siRNA (VCaP-si*ERG*) as indicated. *ERG* knockdown was confirmed by immunoblot analysis. (b) VCaP cells as indicated were assayed for invasion through a modified basement membrane. (c) VCaP-si*ERG* and VCaP cells treated with nontargeting siRNA were profiled and a molecular concept map of the *under-expressed in VCaP-siERG* signature (ringed yellow node) was generated. Each edge represents a significant enrichment ($P < .001$). Blue edges indicate enrichments with *in vivo ETS*–positive *versus* negative prostate cancer signatures. (d) Chromatin immunoprecipitation identifies *PLAT* and *PLAU* as direct targets of *ERG* in VCaP cells, by enrichment of *ERG* binding to the proximal promoters of *PLAT* and *PLAU* compared to IgG control. The promoter of KIAA0089 was used as a negative control. (e) VCaP cells were treated with the indicated inhibitors (as in Figure 2*g*) and assessed for invasion. (f) VCaP cells were treated with transfection reagent alone (untreated), or transfected with nontargeting, *PLAU* or *PLAT* siRNA as indicated and assayed for invasion. For all invasion assays, mean ($n = 3$) ± SEM are shown; *$P < .05$.
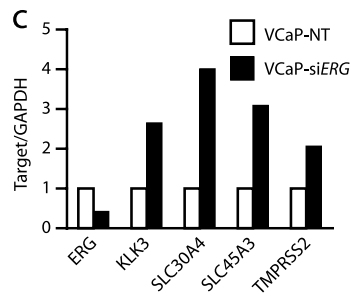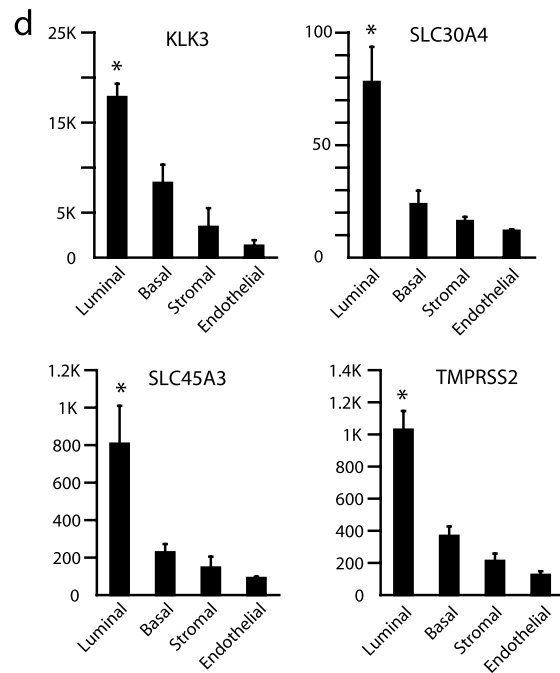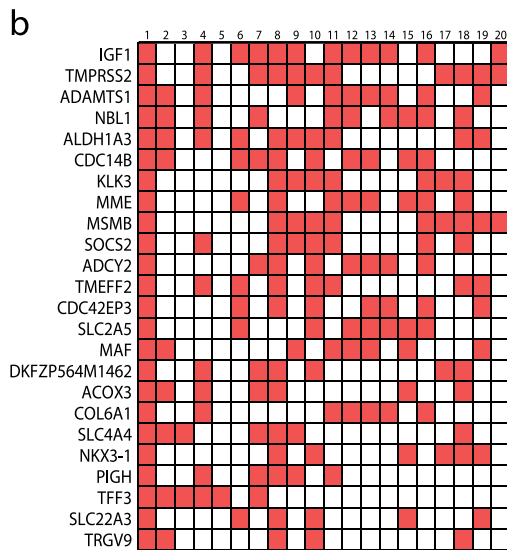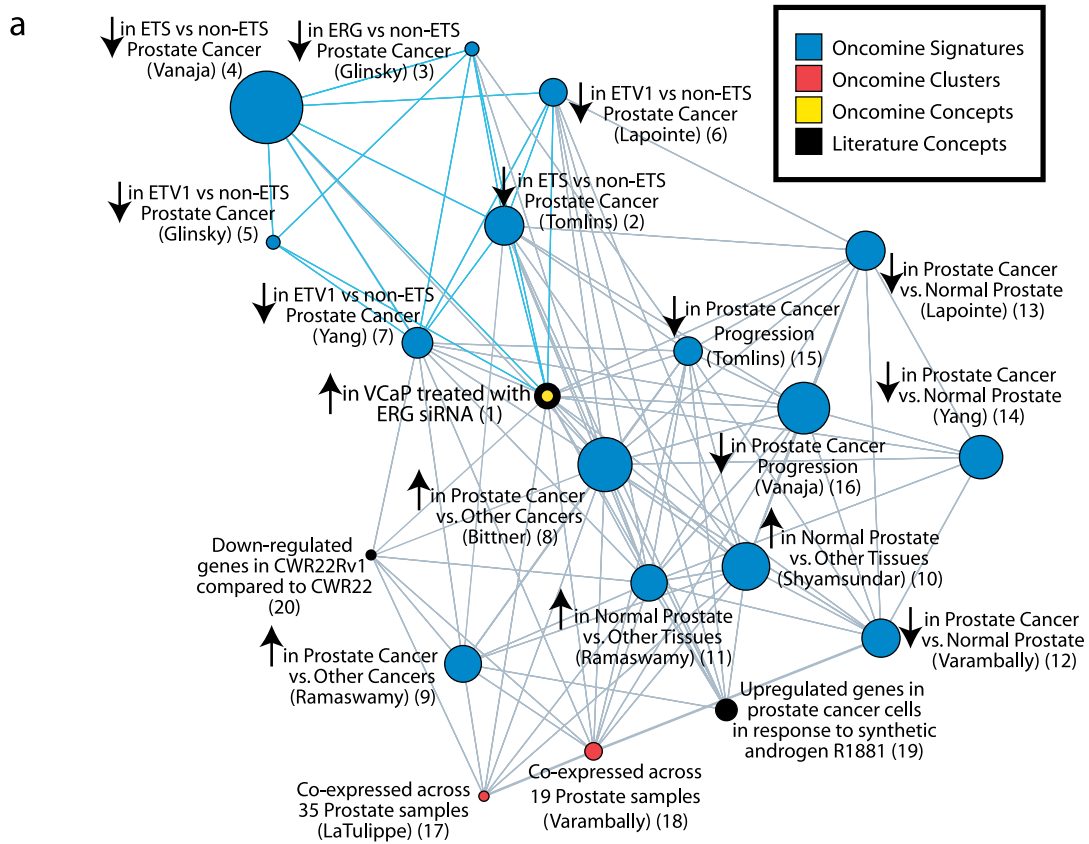
By immunoblot analysis, we confirmed the over-expression of PLAU and MMP3 in RWPE-*ERG* cells (Figure 2*e*). To determine if these genes are direct targets of ERG, we performed ChIP, which demonstrated that ERG binds to the proximal promoter of both *PLAU* and *MMP3* (Figure 2*f*). No enrichment of ERG binding was observed in RWPE-*GUS* cells or LNCaP (*ETV1* rearrangement–positive [22]) for *MMP3* or *PLAU* (Figure W4).

We next assessed the role of both MMPs and the plasminogen activator pathways in the invasive phenotype of RWPE-*ERG* cells using small molecule MMP inhibitors, amiloride (a specific PLAU inhibitor [38]), ectopic PAI-1 (which inhibits plasminogen activators [39]) and siRNA knockdown of *PLAU*. As shown in Figure 2*g*, whereas MMP inhibitors did not significantly inhibit invasion, both amiloride and PAI-1 significantly inhibited the invasiveness of RWPE-*ERG* cells. Similarly, siRNA knockdown of *PLAU* significantly inhibited the invasion of RWPE-*ERG* cells, whereas siRNA knockdown of the tissue plasminogen activator (*PLAT*) had no effect on RWPE-*ERG* invasion (Figure 2*h*). Similar effects on invasion were seen with independent siRNA duplexes directed against *PLAU*. Cytosine arabinoside (ARA-C), which has recently been identified as an inhibitor of the EWS–FLI fusion found in Ewing's sarcoma [27], also showed no effect on RWPE-*ERG* invasion (Figure 2*g*). Together, this work demonstrates that *ERG* directly induces *PLAU* expression in RWPE cells and that inhibition of *PLAU* blocks *ERG*-mediated invasion.

Knockdown of ERG in VCaP Cells Inhibits Invasion

Together, our *in vivo* and *in vitro* studies show that the most common *TMPRSS2–ERG* fusion product is unable to transform benign prostatic epithelial cell lines or induce the development of frank adenocarcinoma in the mouse prostate. However, our previous work, including expression profiling on laser-captured microdissected cell populations and a fluorescence *in situ* hybridization (FISH)–based study on prostate cancer progression, suggest that *TMPRSS2–ERG* gene fusions occur in the context of preexisting genetic lesions, often during the PIN to carcinoma transition [7,33].

To investigate the role of *TMPRSS2–ERG* in this context of preexisting genetic lesions, we used siRNA to knockdown *ERG* in VCaP (VCaP-si*ERG*) cells that are known to harbor the TMPRSS2–ERG gene fusion [2]. Immunoblot analysis confirmed that siRNA directed against *ERG* reduced expression compared to nontargeting control siRNA (Figure 3*a*). Quantitative PCR also demonstrated a 63% decrease in ERG transcript expression in VCaP-si*ERG* (Figure W5*a*, *P* = .009). *ERG* knockdown also significantly inhibited the invasion of VCaP cells (Figure 3*b*) without affecting proliferation (Figure W5*b*), similar to *ERG* over-expression in RWPE cells. Similar results were seen using a second siRNA targeting an independent sequence in *ERG* (data not shown).

To determine the transcriptional profile mediated by *TMPRSS2–ERG* in VCaP, we profiled VCaP-si*ERG* cells. We identified 265 and 291 features over- and under-expressed (as described in the Materials and Methods section), respectively, in VCaP-si*ERG* compared to VCaP treated with nontargeting siRNA, and uploaded these signatures into the OCM. The two most significantly enriched concepts in our *under-expressed in VCaP-siERG* signature were two signatures of genes *over-expressed in ETS-positive versus -negative prostate cancers* (GSE8218, OR = 5.73, *P* = 2.5 × 10^{-19} and Vanaja et al. [40], OR = 3.49, *P* = 3.9 × 10^{-11}) (Figure 3*c*). All other *over-expressed in ETS-positive versus -negative prostate cancer* signatures [33,41,42] in the Oncomine database were also enriched in our *under-expressed in VCaP-siERG* signature, supporting VCaP as a highly relevant model of *TMPRSS2–ERG*–positive prostate cancers. Our under-expressed in VCaP-si*ERG* signature also shared enrichment with our previous signature of genes *over-expressed in laser-captured prostate cancer versus PIN* (OR = 3.79, *P* = 4.5 × 10^{-6}). In that study, we observed that PIN and prostate cancer had very similar expression signatures and hypothesized that *TMPRSS2–ERG* fusions occurred during the PIN to prostate cancer transition and dysregulated a limited number of transcripts, likely involved in invasion [33].

The Role of the Plasminogen Activator Pathway in VCaP Cells

Our under-expressed in VCaP-si*ERG* signature also shared significant enrichment with our *over-expressed in transient PrEC-ERG* and *transient RWPE-ERG* signatures (Figure 3*c*; OR = 6.89 and 3.21, *P* = 1.4 × 10^{-5} and 7 × 10^{-5}, respectively), suggesting common transcriptional programs controlled by *ERG* across cell types and genetic context. Interestingly, although PLAU was not significantly dysregulated in VCaP-si*ERG* cells, the most strongly down regulated feature in VCaP-si*ERG* cells was tissue plasminogen activator (*PLAT*). Similar to *PLAU*, which we showed to be strongly over-expressed and a direct target of *ERG* in RWPE cells, we confirmed that *PLAT* was strongly downregulated (Figure W5*c*). Quantitative PCR analysis showed that VCaP-si*ERG* cells expressed very low levels of *PLAU* at baseline (Figure W6), however ChIP identified both *PLAU* and *PLAT* as direct targets of *ERG* in VCaP-si*ERG* cells (Figure 3*d*). Whereas ectopic PAI-1, amiloride (which inhibits *PLAU* but not *PLAT* [38]) (Figure 3*e*), and siRNA knockdown of *PLAU* inhibited the invasion of VCaP cells, siRNA knockdown of *PLAT* had no effect on VCaP invasion (Figure 3*f*). Additionally, inhibitors of MMPs and ARA-C had no significant effect on VCaP invasion (Figure 3*e*), similar to RWPE-*ERG*. Together, these results support plasminogen activators as direct targets of *ERG* across multiple *TMPRSS2–ERG* model systems and demonstrate that inhibition of *PLAU* blocks *ERG*-induced invasion across *TMPRSS2–ERG* cell line models.

Figure 4. *ERG* knockdown in VCaP cells derepresses a transcriptional program associated with normal prostatic epithelial cell differentiation. (a) VCaP-si*ERG* and VCaP cells treated with nontargeting siRNA were profiled and a molecular concept map of the *over-expressed in VCaP-siERG* signature (ringed yellow node) was generated. Each edge represents a significant enrichment (*P* < .001). Blue edges indicate enrichments with *in vivo ETS-positive versus -negative prostate cancer* signatures. (b) Overlay map identifying genes present (red cells), including *KLK3* (PSA), across multiple concepts in the over-expressed in VCaP-si*ERG* enrichment network (indicated by number). (c) qPCR confirmation of increased expression in VCaP-si*ERG* cells (black) compared to VCaP-NT cells (white) of transcripts strongly expressed in prostatic epithelial cells. (d) Analysis of prostate cell type specificity using a microarray data set profiling magnetically sorted prostate cell populations. Mean RMA normalized fluorescent intensities (*n* = 5 ± SEM) are shown. **P* < .05, for all pairwise *t* tests involving luminal cells.

Transcriptional Signatures of ERG in VCaP Cells

Our *under-expressed in VCaP-siERG* signature also shared significant enrichment with a cluster of 18 genes coexpressed across 72 prostate cancer tissue samples [41], with eight genes shared (OR = 56.65, P = 7.2×10^{-10}). Because this cluster contains *ERG* (Figure W5*d*), this result supports *ERG* knockdown in VCaP-modulating genes regulated by *ERG* in *TMPRSS2–ERG*–positive tumors. To identify such genes, we examined genes coexpressed with *ERG* across multiple prostate cancer profiling studies in the Oncomine database. We identified four genes, *CACNA1D*, *KCNS3*, *LAMC2*, and *PLA1A*, that were down-regulated in VCaP-si*ERG* cells and also showed greater than 0.5 correlation with *ERG* across multiple prostate cancer profiling studies (Figure W7). *CACNA1D* was significantly down regulated in three of four arrays, with the fourth array showing 0.54-fold expression in VCaP-si*ERG* (P = .06). In addition, we also identified decreased expression of *ARGHDIB* in VCaP-si*ERG* cells and over-expression in all ETS-positive *versus* -negative expression signatures (Figure W5*d*). By qPCR, we confirmed the decreased expression of these genes in VCaP-si*ERG* cells (Figure W5*e*) and ChIP identified *LAMC2*, *KCNS3*, and *PLA1A* as direct targets of *ERG* (Figure W5*f*). By qPCR, we also confirmed the coexpression of *ERG* and *PLA1A* (R = 0.72, P = 6.1×10^{-8}) in an independent set of prostate tissues (Figure W5*g*). Thus, our work provides direct *ERG* target genes over-expressed in *TMPRSS2–ERG*–positive prostate cancers for further functional study.

We next examined our *over-expressed in VCaP-siERG* signature using the OCM. Consistent with the results described above, all *under-expressed in ETS-positive versus -negative prostate cancer* signatures in the Oncomine database (GSE8218 and [33,40–42]) were enriched in our *over-expressed in VCaP-siERG* signature (OR = 6.41–2.71, P = 5.2×10^{-15} to 7.0×10^{-5}). Intriguingly, OCM analysis revealed that the most significantly enriched concept in our *over-expressed in VCaP-siERG* signature was a signature of genes *over-expressed in prostate cancers compared to 28 other cancer types* (GSE2109) (OR = 4.46, P = 5.8×10^{-18}) (Figure 4*a*). Several other concepts representing genes over-expressed in prostate cancer compared to other cancers, normal prostate tissue compared to other normal tissues and normal prostate compared to prostate cancer were also strongly enriched in our signature. Examining the genes common to these concepts and VCaP-si*ERG*, we identified numerous archetypal prostate epithelial cell transcripts, including *KLK3* (*PSA*), *MSMB*, *NKX3-1*, *TMPRSS2*, *TRGV9* (also known as *TARP*) [43], *SLC30A4* (also known as ZnT4) [44], and *SLC45A3* [22] (Figure 4*b* and Figure W8). We confirmed the over-expression of this transcriptional program by qPCR (Figure 4*c*), and confirmed that these genes are normally expressed specifically in luminal epithelial prostate cells using an independent data set containing expression profiling data from magnetically sorted prostate luminal epithelial, basal epithelial, stromal fibromuscular, and endothelial cells (Figure 4*d* and Figure W8). Because *ERG* knockdown in VCaP results in the increased expression of genes associated with differentiated luminal prostate epithelial cells, we hypothesize that *TMPRSS2–ERG* fusion may function to keep prostate cancer cells in a dedifferentiated state. Future experiments will be needed to address this hypothesis.

Discussion

Our *in vitro* and *in vivo* studies on the *TMPRSS2–ERG* fusion described here support the functional similarity between *ETS* gene fusions, consistent with our initial observation of mutually exclusive *ERG* or *ETV1* over-expression in prostate cancers [2]. This includes the similar phenotypic and transcriptional programs induced by *ERG*, *ETV1*, and *ETV5* over-expression in benign prostate cells, the similar phenotype of transgenic mice expressing *ERG* or *ETV1* in the prostate [22], and the enrichment of genes over-expressed in *ERG* or *ETV1*-positive *versus* ETS-negative prostate cancers in our VCaP-si*ERG* signature (see Figure 3*c*).

Importantly, our *in vivo* and *in vitro* studies show that the most common *TMPRSS2–ERG* fusion product is unable to transform benign prostatic epithelial cell lines or induce the development of frank adenocarcinoma in the mouse prostate. However, both of these results are consistent with the occurrence of *TMPRSS2–ERG* fusions in the context of preexisting genetic lesions during the course of human prostate cancer development.

Similar to the expression of the *TMPRSS2–ETV1* fusion product in the mouse prostate [22], expression of the *TMPRSS2–ERG* fusion product in the mouse prostate resulted in the development of PIN, without the development of frank adenocarcinoma. As described below, in human prostate cancer development, *TMPRSS2–ERG* fusions occur in the context of earlier lesions, such as loss of single *NKX3-1* and/or *PTEN* alleles [45]. Importantly, mouse models of such early lesions, such as *NKX3-1*$^{+/-}$ and *PTEN*$^{+/-}$ mice [46–48], also only develop mPIN without frank adenocarcinoma. Together, these results are consistent with the development of invasive prostate cancer requiring multiple genetic lesions. Importantly, these results also suggest that crosses between ARR2Pb-*ERG* mice and transgenic mice modeling earlier lesions should produce highly relevant oncogene/tumor suppressor models mimicking early events in human prostate cancer development.

In this study, over-expression of *ERG* in benign prostate cells markedly increased invasion but did not result in transformation, similar to experiments with *ETV1* and *ETV5* [3,33]. These results support our previous hypothesis that *ETS* gene rearrangements mediate invasion in human prostate cancer development. For example, using expression profiling on laser captured microdissected cell populations, we demonstrated that whereas benign prostatic epithelial cells and epithelial cells in PIN lesions have distinct expression profiles, PIN and cancerous epithelium share remarkably similar expression profiles [33]. This suggests that PIN and cancerous cells share many genetic lesions, with a limited number of genetic events likely mediating the PIN to prostate cancer transition (defined histologically by the presence of invasion). Importantly in our profiled samples, *TMPRSS2–ERG* fusions only occurred in prostate cancer and not in PIN lesions (as evidenced by *ERG* outlier expression), suggesting that it might be the key lesion driving the invasive transition.

Further supporting a role for *TMPRSS2–ERG* in invasion, we previously demonstrated in a FISH-based study that *TMPRSS2–ERG* fusion was not identified in benign prostate cells or proliferative inflammatory atrophy, which may be an early precursor of PIN/prostate cancer. However, *TMPRSS2–ERG* fusion could be detected in 19% of PIN lesions, but these foci were intermingling with cancerous glands that were similarly *TMPRSS2–ERG*–positive [7]. *TMPRSS2–ERG* fusion was not identified in PIN lesions distant to prostate cancer, even if the cancerous lesion from the same individual demonstrated the *TMPRSS2–ERG* fusion. Together, this FISH-based study suggested that *TMPRSS2–ERG* fusions may directly mediate the development of prostate cancer from PIN lesions.

Thus, to study *TMPRSS2–ERG* function in a more realistic cellular context, we investigated the effects of *ERG* knockdown in the *TMPRSS2–ERG*–positive VCaP cell line. These experiments confirmed VCaP as a highly relevant prostate cancer cell line model,

as siRNA knockdown of *ERG* inhibited invasion and modulated transcriptional programs activated in *TMPRSS2–ERG*–positive tumors. Additionally, *ERG* knockdown also modulated the transcriptional program that differentiated our laser captured PIN and prostate cancer cell populations, consistent with *TMPRSS2–ERG* driving this important transition. Importantly, these programs were not modulated by *ERG* over-expression in RWPE cells, further supporting VCaP as a more realistic model of *TMPRSS2–ERG* prostate cancer. Interestingly, in both *RWPE-ERG* and VCaP cells, we demonstrate that the plasminogen activator pathway is crucial to *ERG*-mediated invasion, similar to *ETV5*-mediated invasion [3]. Thus, this pathway warrants further investigation as a therapeutic target for *TMPRSS2–ERG*–positive prostate cancer.

Acknowledgments

References
[1] Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, Cao X, Wei JT, Rubin MA, Shah RB, et al. (2006). *TMPRSS2:ETV4* gene fusions define a third molecular subtype of prostate cancer. *Cancer Res* **66**, 3396–3400.

[2] Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, et al. (2005). Recurrent fusion of *TMPRSS2* and *ETS* transcription factor genes in prostate cancer. *Science* **310**, 644–648.

[3] Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR, Cao X, Singla N, Montie JE, Varambally S, et al. (2008). Characterization of *TMPRSS2:ETV5* and *SLC45A3:ETV5* gene fusions in prostate cancer. *Cancer Res* **68**, 73–80.

[4] Demichelis F, Fall K, Perner S, Andren O, Schmidt F, Setlur SR, Hoshida Y, Mosquera JM, Pawitan Y, Lee C, et al. *TMPRSS2:ERG* gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene*.

[5] Lapointe J, Kim YH, Miller MA, Li C, Kaygusuz G, van de Rijn M, Huntsman DG, Brooks JD, and Pollack JR (2007). A variant *TMPRSS2* isoform and *ERG* fusion product in prostate cancer with implications for molecular diagnosis. *Mod Pathol* **20**, 467–473.

[6] Perner S, Demichelis F, Beroukhim R, Schmidt FH, Mosquera JM, Setlur S, Tchinda J, Tomlins SA, Hofer MD, Pienta KG, et al. (2006). *TMPRSS2: ERG* fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res* **66**, 8337–8341.

[7] Perner S, Mosquera JM, Demichelis F, Hofer MD, Paris PL, Simko J, Collins C, Bismar TA, Chinnaiyan AM, De Marzo AM, et al. (2007). *TMPRSS2–ERG* fusion prostate cancer: an early molecular event associated with invasion. *Am J Surg Pathol* **31**, 882–888.

[8] Soller MJ, Isaksson M, Elfving P, Soller W, Lundgren R, and Panagopoulos I (2006). Confirmation of the high frequency of the *TMPRSS2/ERG* fusion gene in prostate cancer. *Genes Chromosomes Cancer* **45**, 717–719.

[9] Wang J, Cai Y, Ren C, and Ittmann M (2006). Expression of variant *TMPRSS2/ERG* fusion messenger RNAs is associated with aggressive prostate cancer. *Cancer Res* **66**, 8347–8351.

[10] Cerveira N, Ribeiro FR, Peixoto A, Costa V, Henrique R, Jeronimo C, and Teixeira MR (2006). *TMPRSS2–ERG* gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. *Neoplasia* **8**, 826–832.

[11] Clark J, Merson S, Jhavar S, Flohr P, Edwards S, Foster CS, Eeles R, Martin FL, Phillips DH, Crundwell M, et al. (2006). Diversity of *TMPRSS2–ERG* fusion transcripts in the human prostate. *Oncogene* **26**, 2667–2673.

[12] Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, and Trapman J (2006). *TMPRSS2:ERG* fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor–negative prostate cancer. *Cancer Res* **66**, 10658–10663.

[13] Iljin K, Wolf M, Edgren H, Gupta S, Kilpinen S, Skotheim RI, Peltola M, Smit F, Verhaegh G, Schalken J, et al. (2006). *TMPRSS2* fusions with oncogenic *ETS* factors in prostate cancer involve unbalanced genomic rearrangements and are associated with *HDAC1* and epigenetic reprogramming. *Cancer Res* **66**, 10242–10246.

[14] Yoshimoto M, Joshua AM, Chilton-Macneill S, Bayani J, Selvarajah S, Evans AJ, Zielenska M, and Squire JA (2006). Three-color FISH analysis of *TMPRSS2/ERG* fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. *Neoplasia* **8**, 465–469.

[15] Mehra R, Tomlins SA, Shen R, Nadeem O, Wang L, Wei JT, Pienta KJ, Ghosh D, Rubin MA, Chinnaiyan AM, et al. (2007). Comprehensive assessment of *TMPRSS2* and *ETS* family gene aberrations in clinically localized prostate cancer. *Mod Pathol* **20**, 538–544.

[16] Nam RK, Sugar L, Wang Z, Yang W, Kitching R, Klotz LH, Venkateswaran V, Narod SA, and Seth A (2007). Expression of *TMPRSS2 ERG* gene fusion in prostate cancer cells is an important prognostic factor for cancer progression. *Cancer Biol Ther* **6**, 40–45.

[17] Rajput AB, Miller MA, De Luca A, Boyd N, Leung S, Hurtado-Coll A, Fazli L, Jones EC, Palmer JB, Gleave ME, et al. (2007). Frequency of the *TMPRSS2: ERG* gene fusion is increased in moderate to poorly differentiated prostate cancers. *J Clin Pathol* **60**, 1238–1243.

[18] Winnes M, Lissbrant E, Damber JE, and Stenman G (2007). Molecular genetic analyses of the *TMPRSS2–ERG* and *TMPRSS2–ETV1* gene fusions in 50 cases of prostate cancer. *Oncol Rep* **17**, 1033–1036.

[19] Lin B, Ferguson C, White JT, Wang S, Vessella R, True LD, Hood L, and Nelson PS (1999). Prostate-localized and androgen-regulated expression of the membrane-bound serine protease *TMPRSS2*. *Cancer Res* **59**, 4180–4184.

[20] Mertz KD, Setlur SR, Dhanasekaran SM, Demichelis F, Perner S, Tomlins S, Tchinda J, Laxman B, Vessella RL, Beroukhim R, et al. (2007). Molecular characterization of *TMPRSS2–ERG* gene fusion in the NCI-H660 prostate cancer cell line: a new perspective for an old model. *Neoplasia* **9**, 200–206.

[21] Hendriksen PJ, Dits NF, Kokame K, Veldhoven A, van Weerden WM, Bangma CH, Trapman J, and Jenster G (2006). Evolution of the androgen receptor pathway during progression of prostate cancer. *Cancer Res* **66**, 5012–5020.

[22] Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, et al. (2007). Distinct classes of chromosomal rearrangements create oncogenic *ETS* gene fusions in prostate cancer. *Nature* **448**, 595–599.

[23] Zhang J, Thomas TZ, Kasper S, and Matusik RJ (2000). A small composite probasin promoter confers high levels of prostate-specific gene expression through regulation by androgens and glucocorticoids *in vitro* and *in vivo*. *Endocrinology* **141**, 4698–4710.

[24] Ellwood-Yen K, Graeber TG, Wongvipat J, Iruela-Arispe ML, Zhang J, Matusik R, Thomas GV, and Sawyers CL (2003). Myc-driven murine prostate cancer shares molecular features with human prostate tumors. *Cancer Cell* **4**, 223–238.

[25] Shappell SB, Thomas GV, Roberts RL, Herbert R, Ittmann MM, Rubin MA, Humphrey PA, Sundberg JP, Rozengurt N, Barrios R, et al. (2004). Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor Meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. *Cancer Res* **64**, 2270–2305.

[26] Korenchuk S, Lehr JE, MClean L, Lee YG, Whitney S, Vessella R, Lin DL, and Pienta KJ (2001). VCaP, a cell-based model system of human prostate cancer. *In Vivo* **15**, 163–168.

[27] Stegmaier K, Wong JS, Ross KN, Chow KT, Peck D, Wright RD, Lessnick SL, Kung AL, and Golub TR (2007). Signature-based small molecule screening identifies cytosine arabinoside as an EWS/FLI modulator in Ewing sarcoma. *PLoS Med* **4**, e122.

[28] Rhodes DR, Kalyana-Sundaram S, Tomlins SA, Mahavisno V, Kasper N, Varambally R, Barrette TR, Ghosh D, Varambally S, and Chinnaiyan AM (2007). Molecular concepts analysis links tumors, pathways, mechanisms, and drugs. *Neoplasia* **9**, 443–454.

[29] Shyamsundar R, Kim YH, Higgins JP, Montgomery K, Jorden M, Sethuraman A, van de Rijn M, Botstein D, Brown PO, and Pollack JR (2005). A DNA microarray survey of gene expression in normal human tissues. *Genome Biol* **6**, R22.

[30] Oudes AJ, Campbell DS, Sorensen CM, Walashek LS, True LD, and Liu AY (2006). Transcriptomes of human prostate cells. *BMC Genomics* **7**, 92.

[31] Boyd KE, Wells J, Gutman J, Bartley SM, and Farnham PJ (1998). c-Myc target gene specificity is determined by a post-DNA binding mechanism. *Proc Natl Acad Sci USA* **95**, 13887–13892.

[32] DeMarzo AM, Nelson WG, Isaacs WB, and Epstein JI (2003). Pathological and molecular aspects of prostate cancer. *Lancet* **361**, 955–964.

[33] Tomlins SA, Mehra R, Rhodes DR, Cao X, Wang L, Dhanasekaran SM, Kalyana-Sundaram S, Wei JT, Rubin MA, Pienta KJ, et al. (2007). Integrative molecular concept modeling of prostate cancer progression. *Nat Genet* **39**, 41–51.

[34] Dechow TN, Pedranzini L, Leitch A, Leslie K, Gerald WL, Linkov I, and Bromberg JF (2004). Requirement of matrix metalloproteinase-9 for the transformation of human mammary epithelial cells by Stat3-C. *Proc Natl Acad Sci USA* **101**, 10602–10607.

[35] Laufs S, Schumacher J, and Allgayer H (2006). Urokinase-receptor (u-PAR): an essential player in multiple games of cancer: a review on its role in tumor progression, invasion, metastasis, proliferation/dormancy, clinical outcome and minimal residual disease. *Cell Cycle* **5**, 1760–1771.

[36] Fingleton B (2006). Matrix metalloproteinases: roles in cancer and metastasis. *Front Biosci* **11**, 479–491.

[37] de Launoit Y, Baert JL, Chotteau-Lelievre A, Monte D, Coutte L, Mauen S, Firlej V, Degerny C, and Verreman K (2006). The Ets transcription factors of the PEA3 group: transcriptional regulators in metastasis. *Biochim Biophys Acta* **1766**, 79–87.

[38] Vassalli JD and Belin D (1987). Amiloride selectively inhibits the urokinase-type plasminogen activator. *FEBS Lett* **214**, 187–191.

[39] Gils A and Declerck PJ (2004). Plasminogen activator inhibitor-1. *Curr Med Chem* **11**, 2323–2334.

[40] Vanaja DK, Cheville JC, Iturria SJ, and Young CY (2003). Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression. *Cancer Res* **63**, 3877–3882.

[41] Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, and Gerald WL (2004). Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest* **113**, 913–923.

[42] Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K, Ferrari M, Egevad L, Rayford W, Bergerheim U, et al. (2004). Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci USA* **101**, 811–816.

[43] Wolfgang CD, Essand M, Vincent JJ, Lee B, and Pastan I (2000). TARP: a nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor gamma chain locus. *Proc Natl Acad Sci USA* **97**, 9437–9442.

[44] Henshall SM, Afar DE, Rasiah KK, Horvath LG, Gish K, Caras I, Ramakrishnan V, Wong M, Jeffry U, Kench JG, et al. (2003). Expression of the zinc transporter ZnT4 is decreased in the progression from early prostate disease to invasive prostate cancer. *Oncogene* **22**, 6005–6012.

[45] Tomlins SA, Rubin MA, and Chinnaiyan AM (2006). Integrative biology of prostate cancer progression. *Annu Rev Pathol* **1**, 243–271.

[46] Kim MJ, Cardiff RD, Desai N, Banach-Petrosky WA, Parsons R, Shen MM, and Abate-Shen C (2002). Cooperativity of Nkx3.1 and Pten loss of function in a mouse model of prostate carcinogenesis. *Proc Natl Acad Sci USA* **99**, 2884–2889.

[47] Abdulkadir SA, Magee JA, Peters TJ, Kaleem Z, Naughton CK, Humphrey PA, and Milbrandt J (2002). Conditional loss of Nkx3.1 in adult mice induces prostatic intraepithelial neoplasia. *Mol Cell Biol* **22**, 1495–1503.

[48] Di Cristofano A, De Acetis M, Koff A, Cordon-Cardo C, and Pandolfi PP (2001). Pten and p27KIP1 cooperate in prostate cancer tumor suppression in the mouse. *Nat Genet* **27**, 222–224.

Table W1. Prostate Pathology in ARR2Pb-*ERG* Transgenic Mice.

ARR2Pb-*ERG* Mouse	Founder No.	Age	AP	VP	DLP	Diagnosis	Liver
1	272	12 weeks	Normal	Hyperplasia	Normal	Hyperplasia	NA
2	272	12 weeks	Normal	Hyperplasia	Adipose tissue	Hyperplasia	NA
3	285	12 weeks	Normal	mPIN	Normal	mPIN	NA
4	285	12 weeks	Normal	mPIN	Hyperplasia	mPIN	Normal
5	429	12–14 weeks	Hyperplasia	Hyperplasia	Hyperplasia	Hyperplasia	Normal
6	429	12–14 weeks	Normal	No tissue	Hyperplasia	Hyperplasia	Normal
7	429	12–14 weeks	No tissue	mPIN	Normal	mPIN	NA
8	457	12–14 weeks	Normal	Hyperplasia	Normal	Hyperplasia	Normal
9	285	20 weeks	Normal	Hyperplasia	Hyperplasia	Hyperplasia	Normal
10	282	34 weeks	Hyperplasia	mPIN	mPIN	mPIN	Normal
11	302	15 months	Hyperplasia	Hyperplasia	Hyperplasia	Hyperplasia	Normal
		Total:	0/10 (0%)	4/10 (40%)	1/11 (9.1%)	4/11 (36.3%)	

For ARR2Pb-*ERG* transgenic mice, the founder number is indicated, along with the age of sacrifice. Observed pathology from H&E–stained sections from the anterior (AP), ventral (VP), and dorsolateral (DLP) prostatic lobes and an overall diagnosis of prostate pathology are given. The liver from the indicated mice was also dissected, and H&E–stained sections were observed for any pathology. NA, not available.

Table W2. Oligonucleotide Primers.

Assay	Gene/Region	Sequence	Bases	Primer	Sequence 5′ to 3′
Expression qPCR	*ERG*	NM_004449.3	574–597	ERG_exon 5–6_f	CGCAGAGTTATCGTGCCAGCAGAT
Expression qPCR	*ERG*	NM_004449.3	659–636	ERG_exon 5–6_r	CCATATTCTTTCACCGCCCACTCC
Expression qPCR	*SERPINE1*	NM_000602.1	1181–1200	SERPINE1-f	GCATGGCCCCCGAGGAGAT
Expression qPCR	*SERPINE1*	NM_000602.1	1270–1248	SERPINE1-r	CTTGGCCCATGAAAAGGACTGTT
Expression qPCR	*IGFBP3*	NM_000598.4	738–762	IGFBP3-f	CGAGTCCAAGCGGGAGACAGAATA
Expression qPCR	*IGFBP3*	NM_000598.4	837–814	IGFBP3-r	TACACCCCTGGGACTCAGCACATT
Expression qPCR	*MMP3*	NM_002422.3	1055–1080	MMP3-f	TTCATTTTGGCCATCTCTTCCTTCAG
Expression qPCR	*MMP3*	NM_002422.3	1181–1155	MMP3-r	TATCCAGCTCGTACCTCATTTCCTCT
Expression qPCR	*ADAM19*	NM_023038.3	2146–2165	ADAM19-f	GCCTATGCCCCCTGAGAGTG
Expression qPCR	*ADAM19*	NM_023038.3	2271–2245	ADAM19-r	GCTTGAGTTGGCCTAGTTTGTTGTTC
Expression qPCR	*MMP9*	NM_004994.2	1181–1201	MMP9-f	TGCCCGGACCAAGGATACAGT
Expression qPCR	*MMP9*	NM_004994.2	1239–1221	MMP9-r	AGCGCGTGGCCGAACTCAT
Expression qPCR	*PLAU*	NM_002658.2	1169–1194	PLAU-f	TACGGCTCTGAAGTCACCACCAAAT
Expression qPCR	*PLAU*	NM_002658.2	1308–1286	PLAU-r	CCCCAGCTCACAATTCCAGTCAA
Expression qPCR	*ARHGDIB*	NM_001175.4	250–273	ARHGDIB-f	AGAAAACGCTGCTGGGAGATGGT
Expression qPCR	*ARHGDIB*	NM_001175.4	326–307	ARHGDIB-r	CAGGGTGAGCCGGGTGACAA
Expression qPCR	*KCNS3*	NM_002252.3	1576–1599	KCNS3-f	CCCTTCCCATCACCATCATCTTCA
Expression qPCR	*KCNS3*	NM_002252.3	1659–1635	KCNS3-r	CCTCACTGCACTGGTCCACATCAAT
Expression qPCR	*LAMC2*	NM_005562.1	3317–3345	LAMC2-f	GGTGATTACAGAAGCCCAGAAGGTTGATA
Expression qPCR	*LAMC2*	NM_005562.1	3408–3385	LAMC2-r	GCAGGAGGCCGTCTAATGTGTTGA
Expression qPCR	*F5*	NM_000130.4	6560–6583	F5-f	CAGGGCTGCAAGTCTCTGTCCTCT
Expression qPCR	*F5*	NM_000130.4	6641–6617	F5-r	GTTTCCATTCCACTCCCTGCTCACT
Expression qPCR	*CACNA1D*	NM_000720.1	5776–5805	CACNA1D-f	CTACTACAGCAGATACCCAGGCAGAAACAT
Expression qPCR	*CACNA1D*	NM_000720.1	5885–5861	CACNA1D-r	GTGAATCATAGCAAACGGGCGAGTC
Expression qPCR	*CD44*	NM_000610.3	3702–3727	CD44-f	TGTTATCCCTGGGGCCCTATTTCAT
Expression qPCR	*CD44*	NM_000610.3	3820–3791	CD44-r	ATCTCTTTCATTTCCATTGGCTTCTTCTCT
Expression qPCR	*PLAU*	NM_002658.2	1169–1194	PLAU-f	TACGGCTCTGAAGTCACCACCAAAT
Expression qPCR	*PLAU*	NM_002658.2	1308–1286	PLAU-r	CCCCAGCTCACAATTCCAGTCAA
Expression qPCR	*PLA1A*	NM_015900.1	1194–1216	PLA1A-f	CCACCCCACAATGCCAGATAAAC
Expression qPCR	*PLA1A*	NM_015900.1	1283–1258	PLA1A-r	TCCCAATAATGGTAGTCCGGTCTTTT
Expression qPCR	*PLAT*	NM_033011.1	843–863	PLAT-f	CACTGGGCCTGGGCAAACATA
Expression qPCR	*PLAT*	NM_033011.1	933–913	PLAT-r	CACGTCAGCCTGCGGTTCTTC
Expression qPCR	*KLK3*	NM_001648.2	826–849	KLK3-f	GAGCACCCCTATCAACCCCCTATT
Expression qPCR	*KLK3*	NM_001648.2	944–921	KLK3-r	AGCAACCCTGGACCTCACACCTAA
Expression qPCR	*SLC30A4*	NM_013309.4	1608–1637	SLC30A4-f	TGTATTTTGGGAACTCCTGCCTTATTTATC
Expression qPCR	*SLC30A4*	NM_013309.4	1696–1668	SLC30A4-r	CAGGGATTCCATTTTCTCATTTAGGTTTG
Expression qPCR	*SLC45A3*	NM_033102.2	1223–1242	SLC45A3-f	TCGTGGGCGAGGGGCTGTA
Expression qPCR	*SLC45A3*	NM_033102.2	1308–1284	SLC45A3-r	CATCCGAACGCCTTCATCATAGTGT
Expression qPCR	*TMPRSS2*	NM_005656.2	1539–1563	TMPRSS2-f	CAGGAGTGTACGGGAATGTGATGGT
Expression qPCR	*TMPRSS2*	NM_005656.2	1608–1585	TMPRSS2-r	GATTAGCCGTCTGCCCTCATTTGT

Table W2. (*continued*)

Assay	Gene	Location (to TSS)	Predicted ETS Site	Primer	Sequence
ChIP PCR	*PLAU*	−1458	−1410 & 135	PLAU_pF2	ATTTGCAAGGCAGGAAAATG
ChIP PCR	*PLAU*	−1282		PLAU_pR2	GTGATTCTGTCACCCCCATC
ChIP PCR	*PLAT*	−217	−57	PLAT_pF1	TGTCATCACAGGGTCCTGAA
ChIP PCR	*PLAT*	−27		PLAT_pR1	TAAAGCAGGGGGAGGAAGTT
ChIP PCR	*MMP3*	−227	−223	MMP3_pF1	CCTCTACCAAGACAGGAAGCA
ChIP PCR	*MMP3*	−93		MMP3_pF1	GCAGGACCATTTCCAAACAT
ChIP PCR	*PLA1A*	−287	−246	PLA1A_pF1	TATCACGGGAAGTGGGAGAG
ChIP PCR	*PLA1A*	−143		PLA1A_pR1	TGCCAGAGTTTTCGGTTTCT
ChIP PCR	*LAMC2*	−561	−535	LAMC2_pF1	CCCTGGTGAGCAGGAAGTTA
ChIP PCR	*LAMC2*	−474		LAMC2_pR1	CACCCTCCAGTTTAGGGTCA
ChIP PCR	*KCNS3*	−1325	−1144	KCNS3_pF1	TAGCCTCTCCTCTGGACCAA
ChIP PCR	*KCNS3*	−1083		KCNS3_pR1	GCAGATTCAAGCTCCAGACC
ChIP PCR	*ARHGDIB*	−1692	−1733	ARHGDIB_pF1	TGCTCTCTCATCCCCCAATA
ChIP PCR	*ARHGDIB*	−1604		ARHGDIB_pR1	CACCCCTTCCCAGAAAAATC
ChIP PCR	*KIAA0079*	Within exon 23	NA	KIAA0079_Exon23	TCTGTCATGTCCTGCTGATGGA
ChIP PCR	*KIAA0079*	Within exon 23		KIAA0079_Exon23	GCCCAAGAAGGACTGACCACTT

Oligonucleotide primers for all assays described in the Materials and Methods section are listed. The assayed gene expression qPCR for all primers is indicated, along with the bases from the corresponding GenBank sequence. All primers are listed 5′ to 3′. For primers for ChIP PCR, the gene, primer location (in relation to the transcriptional start site (TSS)), and the location of predicted ETS binding sites (in relation to the TSS) are given.

Table W3. Cancer Types and Normal Tissues from the expO and Shyamsundar Normal Tissue Datasets.

International Genomics Consortium's expO Data Set (GSE2109) (Bittner_Multi-cancer at www.oncomine.org)			Shyamsundar Normal Tissue Data Set (GSE2193) (Shyamsundar_Normal at www.oncomine.org)		
No.	Cancer Type	*n*	No.	Normal Tissue Type	*n*
1	Bladder papillary carcinoma	4	1	Adrenal	4
2	Bladder transitional cell carcinoma	10	2	Bladder	2
3	Breast ductal carcinoma	95	3	Brain	8
4	Cervix squamous cell carcinoma	10	4	Buffycoat	2
5	Colon adenocarcinoma	104	5	Cervix	3
6	Metastatic colon carcinoma	16	6	Colon	3
7	Mucinous colon carcinoma	12	7	Esophagus	3
8	Endometrial adenocarcinoma	5	8	Fallopian tube	4
9	Endometrial endometrioid carcinoma	45	9	Heart	6
10	Endometrial mixed mullerian tumor	6	10	Kidney	5
11	Metastatic endometrial carcinoma	7	11	Liver	5
12	Soft tissue sarcoma	13	12	Lung	4
13	Clear cell renal carcinoma	78	13	Lymph node	5
14	Papillary renal cell carcinoma	6	14	Muscle	2
15	Lung adenocarcinoma	19	15	Ovary	5
16	Bronchioloalveolar carcinoma	7	16	Pancreas	2
17	Squamous cell lung carcinoma	17	17	Parathyroid	3
18	Ovarian adenocarcinoma	20	18	Salivary Gland	4
19	Ovarian endometrioid carcinoma	13	19	Seminal Vesicle	3
20	Metastatic ovarian carcinoma	36	20	Small Bowel	3
21	Ovarian mucinous carcinoma	4	21	Spleen	3
22	Ovarian papillary carcinoma	38	22	Stomach	4
23	Pancreatic ductal carcinoma	3	23	Testes	3
24	Rectosigmoid adenocarcinoma	15	24	Thymus	2
25	Rectal adenocarcinoma	13	25	Thyroid	6
26	Renal pelvis transitional cell carcinoma	4	26	Tonsil	4
27	Metastatic melanoma	5	27	Uterus	5
28	Papillary thyroid carcinoma	10		Prostate	5
	Prostate adenocarcinoma	15			

For the expO multicancer data set accessed in the Oncomine database, the 29 cancer types displayed in Figure W7 are indicated with the number of profiled samples per type. For the Shyamsundar normal tissue data set, the 28 normal tissue types displayed in Figure W7 are indicated.
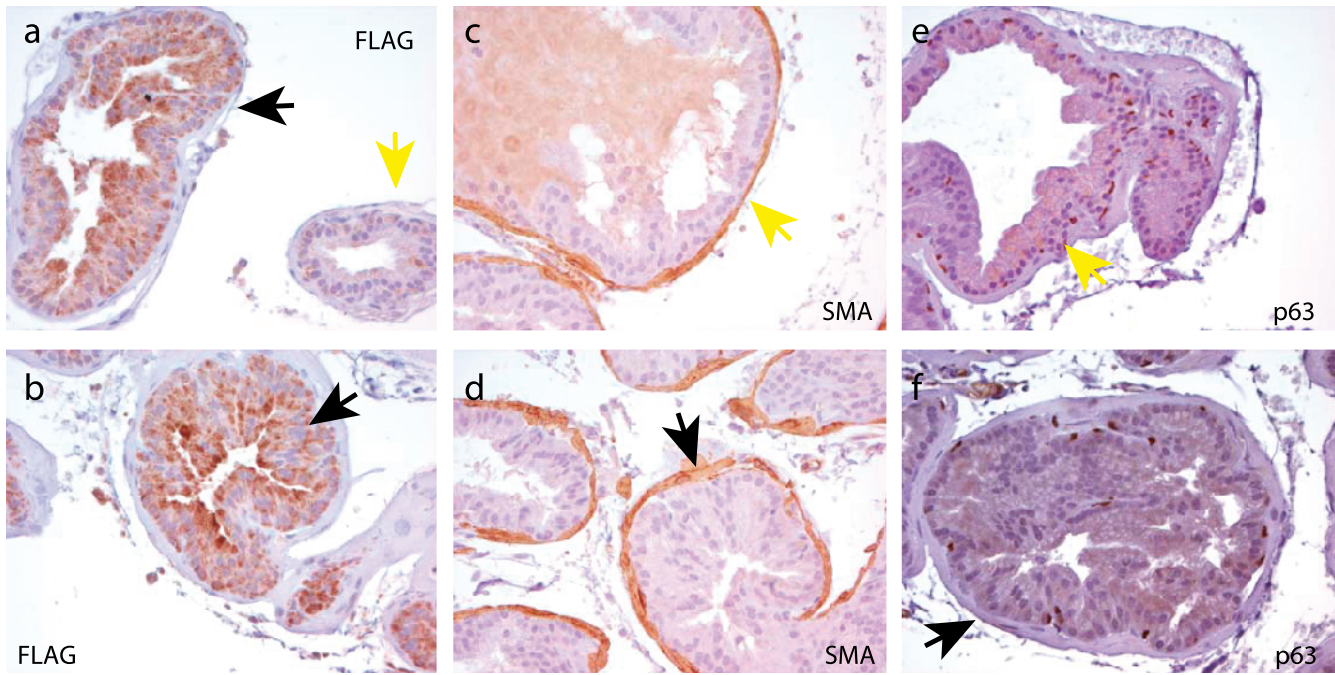
Figure W1. Development of mPIN in *TMPRSS2–ERG* transgenic mice. (a and b) Immunohistochemistry confirmed ERG-FLAG expression exclusively in areas of mPIN and not benign glands in ARR2Pb-*ERG* mice. Benign epithelia and areas of mPIN are indicated by yellow and black arrows, respectively. (c and d) Immunohistochemistry with smooth muscle actin (SMA) demonstrates a continuous fibromuscular layer around (c) benign glands and (d) all mPIN lesions, whereas the basal cell markers (e and f) p63 demonstrate loss of circumferential basal cells in mPIN foci (f) compared to normal glands (e). Original magnification, ×400.
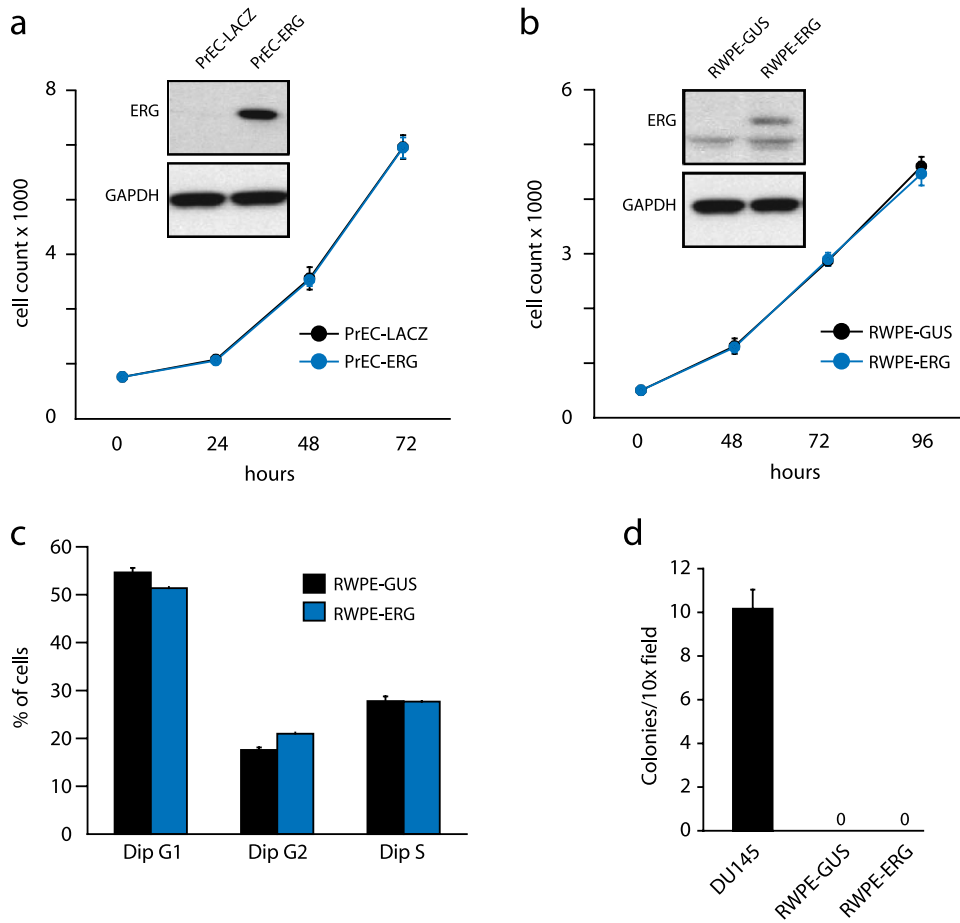
Figure W2. Over-expression of *ERG* does not affect proliferation or transform benign prostatic epithelial cells. (a) Primary prostatic epithelial cells (PrEC) were infected with *ERG* or *LACZ* adenovirus as indicated and assayed for proliferation. Mean ($n = 3$) ± SEM are shown. Results are representative of three independent experiments. (b) The benign immortalized prostate cell line RWPE was infected with *ERG* or control (*GUS*) lentivirus as indicated, and stable clones were generated and assayed for proliferation. Insets of a and b show ERG over-expression by immunoblot analysis. (c) *ERG* over-expression does not increase the percentage of RWPE cells in S phase. RWPE-*GUS* and RWPE-*ERG* cells were analyzed for cell cycle distribution by fluorescence activated cell sorting (FACS). The distributions of cells in the G_1, S, and G_2 phases are indicated. Mean ($n = 4$) ± SEM are shown. (d) *ERG* over-expression does not enhance the anchorage independent growth of RWPE cells. RWPE-*GUS*, RWPE-*ERG*, and DU145 (positive control) cells were assessed for anchorage-independent growth by assaying colony formation in soft agar. After 12 days, the plates were stained, and colonies counted. The number of colonies per high-power field was assessed. Mean colonies per field ($n = 6$) ± SEM are shown.
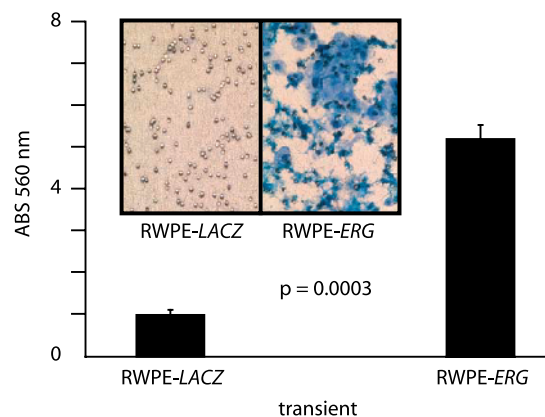

Figure W3. Transient over-expression of *ERG* increases invasion in RWPE cells. We infected the benign immortalized prostate cell line RWPE with *ERG* or *LACZ* adenovirus and assayed for invasion through a modified basement membrane, mean ($n = 3$) ± SEM. Inset shows photomicrograph of invaded cells.
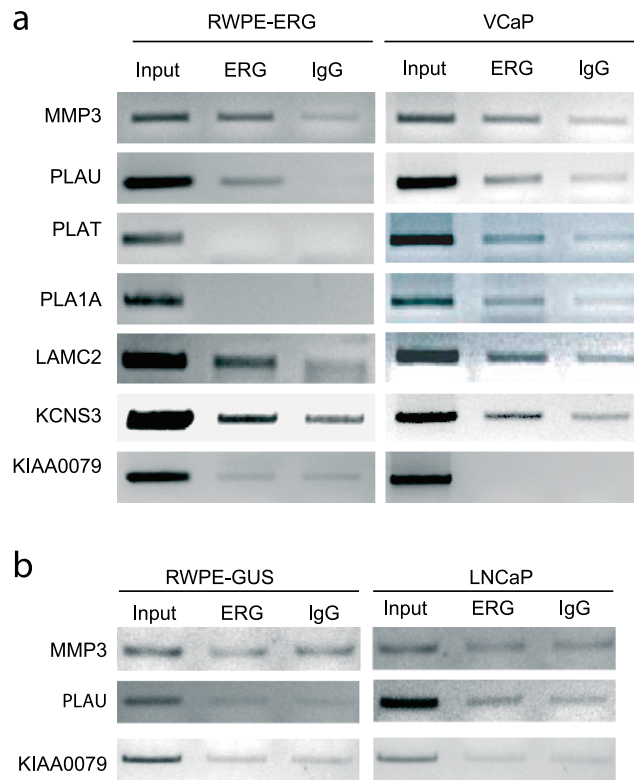
Figure W4. Chromatin immunoprecipitation across *TMPRSS2–ERG* model systems. (a) Chromatin immunoprecipitation to detect enrichment of ERG binding to the proximal promoters of indicated genes compared to IgG control in RWPE-*ERG* and VCaP cells. The promoter of KIAA0089 was used as a negative control. (b) RWPE-*GUS* and LNCaP failed to show any enrichment of ERG binding to assayed promoters.
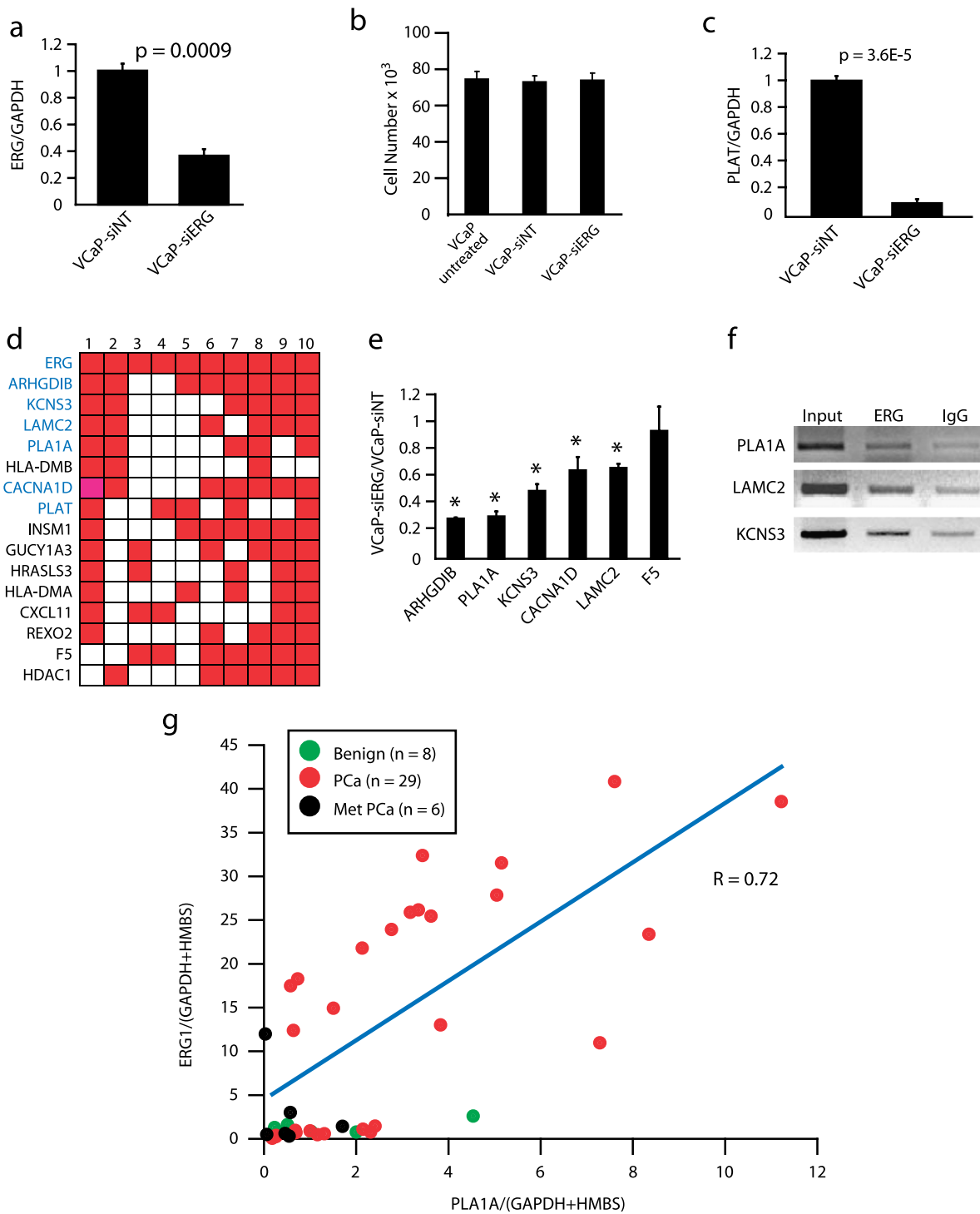
Figure W5. *ERG* knockdown in VCaP attenuates a transcriptional program over-expressed in *TMPRSS2:ETS*–positive tumors. (a) siRNA knockdown of *ERG* in the *TMPRSS2–ERG*–positive prostate cancer cell line VCaP. VCaP cells were either treated with transfection reagent alone (untreated) or transfected with nontargeting or *ERG* siRNA (VCaP-si*ERG*) as indicated. *ERG* knockdown was confirmed by qPCR. (b) ERG knockdown in VCaP does not affect cell proliferation. VCaP cells as indicated were assayed for proliferation by cell counting 72 hours after siRNA transfection. Mean ($n = 3$) ± SEM are shown. (c) qPCR confirmation of decreased *PLAT* expression in VCaP-si*ERG* compared to VCaP-si*NT* cells. (d) Overlay map identifying genes present (red cells) across multiple concepts in the VCaP-si*ERG* enrichment network (indicated by number). *CACNA1D*, in magenta, was identified as differentially expressed in three of four replicate VCaP-si*ERG* arrays. Genes confirmed as under-expressed in VCaP-si*ERG* cells by qPCR are indicated in blue. (e) qPCR confirmation of downregulated genes in VCaP-si*ERG* cells; $*P < .05$, compared to VCaP treated with nontargeting siRNA. (f) Chromatin immunoprecipitation identification of direct *ERG* targets. (g) *ERG* and *PLA1A* show correlated expression across prostate tissues. *ERG* and *PLA1A* expression (normalized to *GAPDH*) was determined by qPCR in benign prostate (green), localized prostate cancer (PCa, red), and metastatic prostate cancer (Met PCa, black) tissue samples. The trend line is shown in blue.
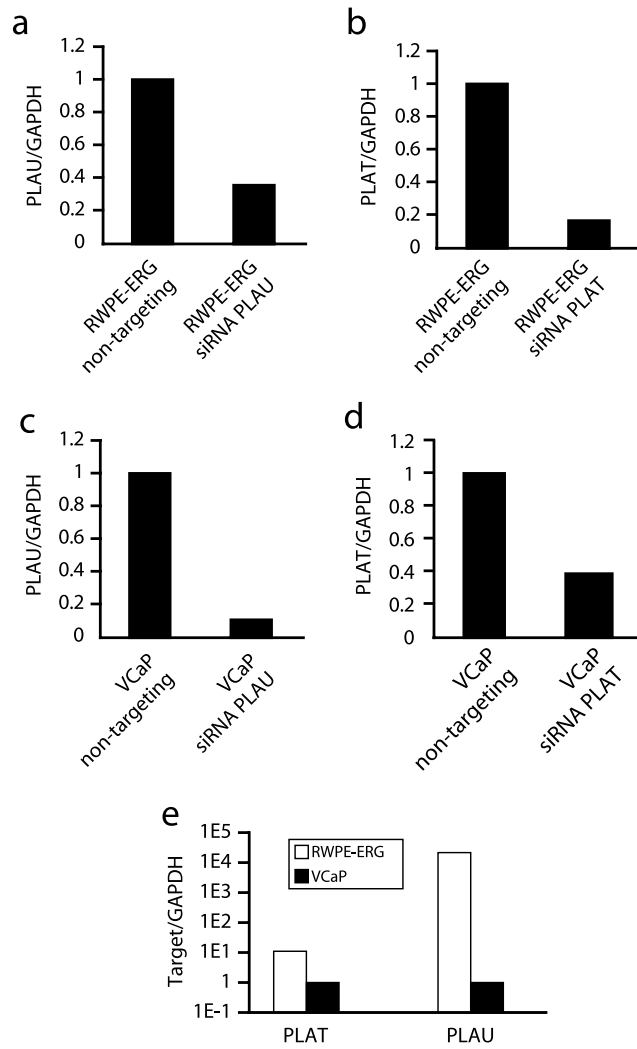
Figure W6. qPCR confirmation of *PLAU* and *PLAT* knockdown in RWPE-*ERG* and VCaP cells. (a and b) RWPE-*ERG* cells were treated with non-targeting siRNA or siRNA against (a) *PLAU* or (b) *PLAT,* and knockdown was confirmed by qPCR. (c and d) VCaP cells were treated with nontargeting siRNA or siRNA against (c) *PLAU* or (d) *PLAT*, and knockdown was confirmed by qPCR. (e) The relative amount of *PLAT* and *PLAU* in RWPE-*ERG* (white) compared to VCaP (black) was determined by qPCR.

GSE 8218 (Yang et al.)

R Rank	GENE	Feature ID	R
1	ERG	211626_x_at	0.89
2	ERG	213541_s_at	0.89
3	ERG	222079_at	0.89
4	PLA1A	219584_at	0.75
5	PLA2G7	206214_at	0.66
6	EST	221018_s_at	0.64
7	COL2A1	213492_at	0.60
8	COL2A1	217404_s_at	0.60
9	PEX10	206351_s_at	0.57
10	EST	219695_at	0.57
11	FAM77C	219438_at	0.57
12	PEX10	206352_s_at	0.57
13	CACNA1D	210108_at	0.57
14	OGDHL	219277_s_at	0.57
15	CACNA1D	207998_s_at	0.57
16	CRISP3	207802_at	0.54
17	LAMC2	202267_at	0.53
18	KCNS3	205968_at	0.53
19	FOXD1	206307_s_at	0.51
20	DLX2	207147_at	0.51
21	TNRC9	215108_x_at	0.51
22	NETO2	218888_s_at	0.51
23	TNRC9	216623_x_at	0.51
24	TNRC9	214774_x_at	0.51
25	INSM1	206502_s_at	0.51

Glinsky et al.

R Rank	GENE	Feature ID	R
1	ERG	211626_x_at	0.91
2	ERG	222079_at	0.91
3	ERG	213541_s_at	0.90
4	KCNS3	205968_at	0.76
5	CACNA1D	207998_s_at	0.74
6	PDE3B	222317_at	0.74
7	EST	214582_at	0.74
8	CACNA1D	210108_at	0.74
9	EST	214596_at	0.69
10	MAGED4	221261_x_at	0.68
11	ITPR3	201188_s_at	0.67
12	ITPR3	201189_s_at	0.67
13	LAMC2	202267_at	0.64
14	HDAC1	201209_at	0.61
15	AMPD3	207992_s_at	0.61
16	NCALD	211685_s_at	0.61
17	ARHGDIB	201288_at	0.57
18	ANKRD6	204671_s_at	0.57
19	ANKRD6	204672_s_at	0.57
20	HLA-DMB	203932_at	0.53
21	PLA1A	219584_at	0.53

Lapointe et al.

R Rank	GENE	Feature ID	R
1	ERG	IMAGE:123755	0.71
2	CACNA1D	IMAGE:49630	0.71
3	EST	IMAGE:1709503	0.68
4	NPR3	IMAGE:1762111	0.68
5	MYO6	IMAGE:470216	0.64
6	MYO6	IMAGE:744944	0.64
7	CACNA1D	IMAGE:757337	0.64
8	GPR110	IMAGE:1492202	0.62
9	PLA1A	IMAGE:250673	0.62
10	MEG3	IMAGE:206907	0.62
11	SH3RF1	IMAGE:1573665	0.51
12	EST	IMAGE:1926387	0.51
13	JOSD3	IMAGE:193122	0.51
14	SH3RF1	IMAGE:811101	0.51
15	C20orf199	IMAGE:796309	0.51
16	C20orf199	IMAGE:745296	0.51
17	CBR4	IMAGE:454795	0.51
18	CBR4	IMAGE:359713	0.51

Tomlins et al.

R Rank	GENE	Feature ID	R
1	ERG	IMAGE:123755	0.57
2	PLA1A	IMAGE:250673	0.57

Vanaja et al.

R Rank	GENE	Feature ID	R
1	ERG	213541_s_at	0.58
2	KCTD6	238077_at	0.58

Figure W7. Identification of genes showing coexpression with *ERG* across multiple prostate cancer profiling studies. Genes showing coexpression with *ERG* ($R > 0.5$) from prostate cancer profiling studies in the Oncomine database. *ERG* was queried in the Oncomine database using the coexpression module. For each study, all genes showing $R > 0.5$ are listed, along with the corresponding feature identification. *ERG* is indicated in red. Genes showing $R > 0.5$ in multiple studies are indicated in blue.
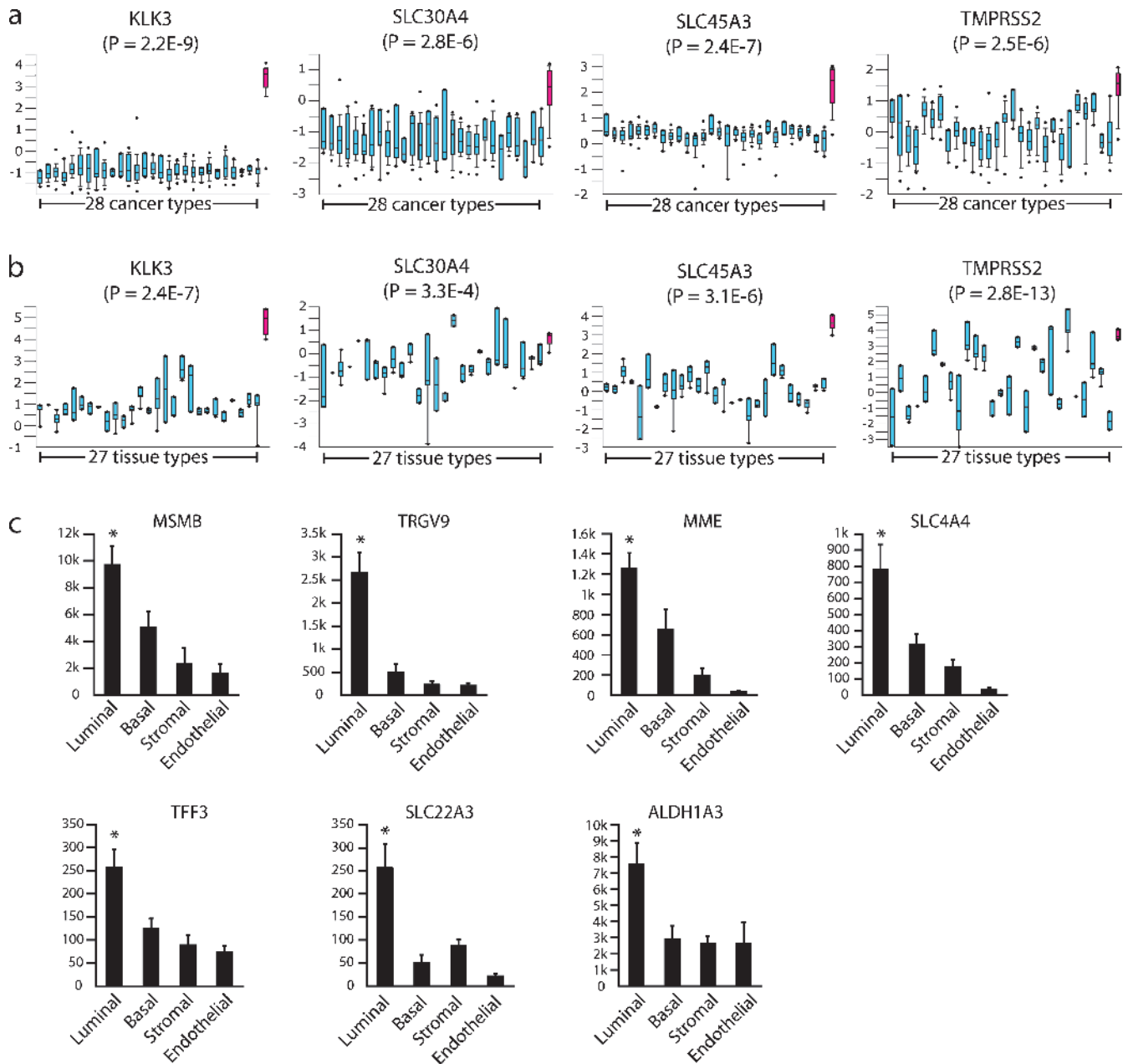
Figure W8. Prostate epithelial specificity of genes induced in VCaP on ERG knockdown. (a) Genes confirmed by qPCR to be over-expressed in VCaP cells treated with *ERG* siRNA were interrogated in the expO multicancer data set, containing expression profiles from 28 cancer types (blue) and prostate cancer (magenta). The significance of prostate cancer *versus* all other cancer types is indicated. (b) The same genes were also interrogated in the Shyamsundar et al. [29] normal tissue data set, containing expression profiles from 27 normal tissue types (blue) and normal prostate tissue (magenta). For both a and b, box and whisker plots show the median and 10th and 90th percentiles in normalized expression units (*z* scores). All cancer and normal tissue types are defined in Table W3. (c) Analysis of prostate cell type specificity using a microarray data set profiling magnetically sorted prostate cell populations for additional genes identified as over-expressed in VCaP cells on *ERG* knockdown (see Figure 4*b*). Mean RMA–normalized fluorescent intensities (*n* = 5 ± SEM) are shown. *P* < .05, for all pairwise *t* tests involving luminal cells.