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	TACGGCTCTGAAGTCACCACCAAAAT
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	TACGGCTCTGAAGTCACCACCAAAAT
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)

No.	Cancer Type	*n*
1	Bladder papillary carcinoma	4
2	Bladder transitional cell carcinoma	10
3	Breast ductal carcinoma	95
4	Cervix squamous cell carcinoma	10
5	Colon adenocarcinoma	104
6	Metastatic colon carcinoma	16
7	Mucinous colon carcinoma	12
8	Endometrial adenocarcinoma	5
9	Endometrial endometrioid carcinoma	45
10	Endometrial mixed mullerian tumor	6
11	Metastatic endometrial carcinoma	7
12	Soft tissue sarcoma	13
13	Clear cell renal carcinoma	78
14	Papillary renal cell carcinoma	6
15	Lung adenocarcinoma	19
16	Bronchioloalveolar carcinoma	7
17	Squamous cell lung carcinoma	17
18	Ovarian adenocarcinoma	20
19	Ovarian endometrioid carcinoma	13
20	Metastatic ovarian carcinoma	36
21	Ovarian mucinous carcinoma	4
22	Ovarian papillary carcinoma	38
23	Pancreatic ductal carcinoma	3
24	Rectosigmoid adenocarcinoma	15
25	Rectal adenocarcinoma	13
26	Renal pelvis transitional cell carcinoma	4
27	Metastatic melanoma	5
28	Papillary thyroid carcinoma	10
	Prostate adenocarcinoma	15

Shyamsundar Normal Tissue Data Set (GSE2193) (Shyamsundar_Normal at www.oncomine.org)

No.	Normal Tissue Type	*n*
1	Adrenal	4
2	Bladder	2
3	Brain	8
4	Buffycoat	2
5	Cervix	3
6	Colon	3
7	Esophagus	3
8	Fallopian tube	4
9	Heart	6
10	Kidney	5
11	Liver	5
12	Lung	4
13	Lymph node	5
14	Muscle	2
15	Ovary	5
16	Pancreas	2
17	Parathyroid	3
18	Salivary Gland	4
19	Seminal Vesicle	3
20	Small Bowel	3
21	Spleen	3
22	Stomach	4
23	Testes	3
24	Thymus	2
25	Thyroid	6
26	Tonsil	4
27	Uterus	5
	Prostate	5

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.

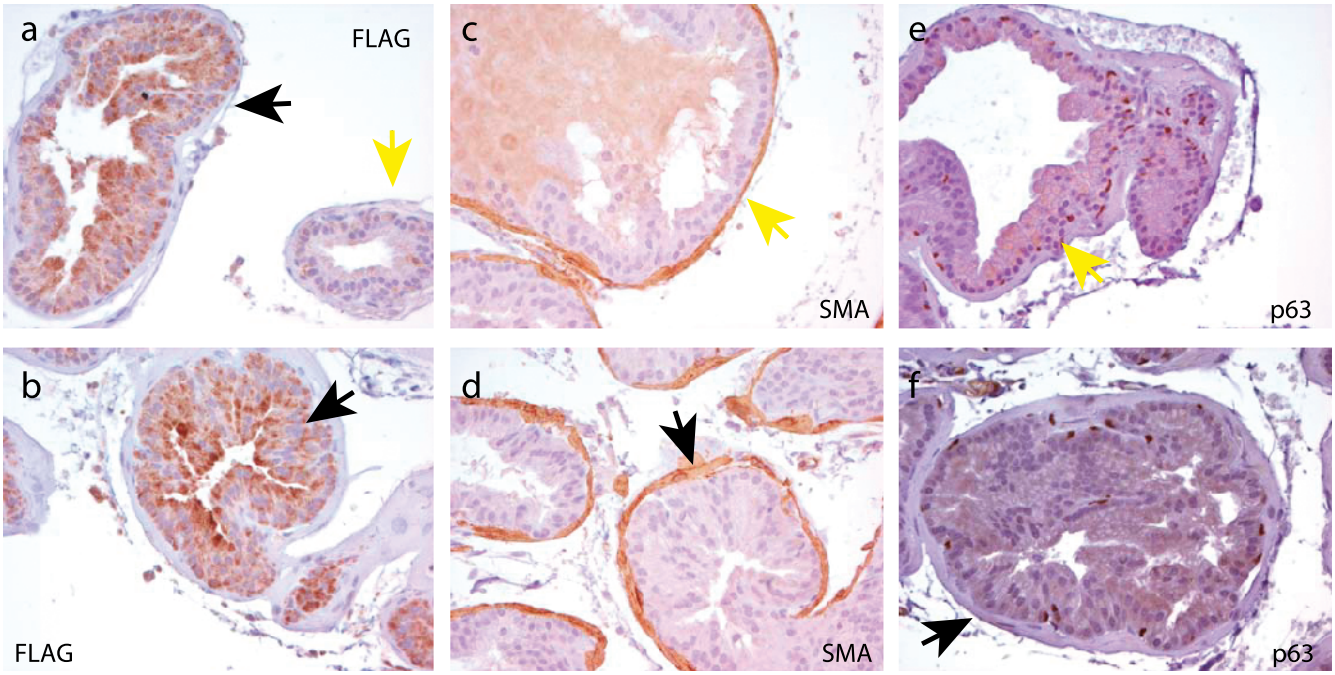

Figure W1. Development of mPIN in *TMPRSS2–ERG* transgenic mice. (a and b) Immunohistochemistry confirmed ERG-FLAG expression exclusively in areas of mPIN and not benign glands in ARR2Pb-*ERG* mice. Benign epithelia and areas of mPIN are indicated by yellow and black arrows, respectively. (c and d) Immunohistochemistry 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.

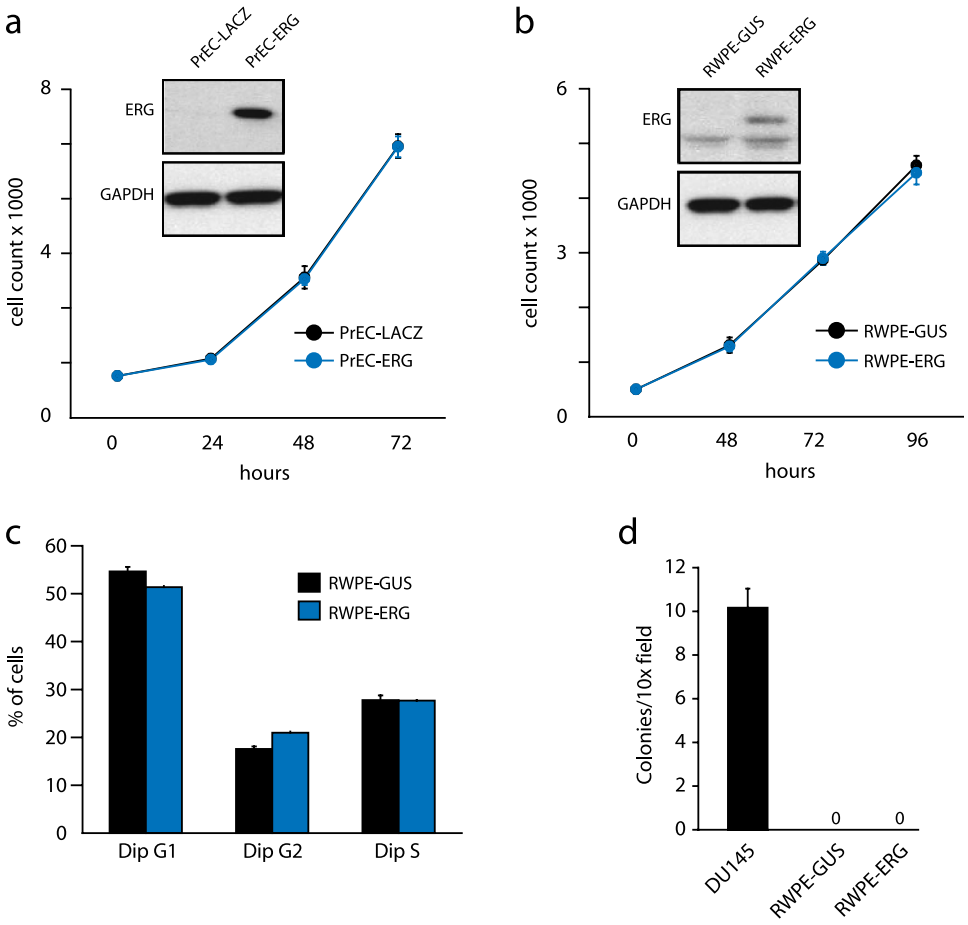

Figure W2. Over-expression of *ERG* does not affect proliferation or transform benign prostatic epithelial cells. (a) Primary prostatic epithelial cells (PrEC) were infected with *ERG* or *LACZ* adenovirus as indicated and assayed for proliferation. Mean ($n = 3$) ± SEM are shown. Results are representative 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.

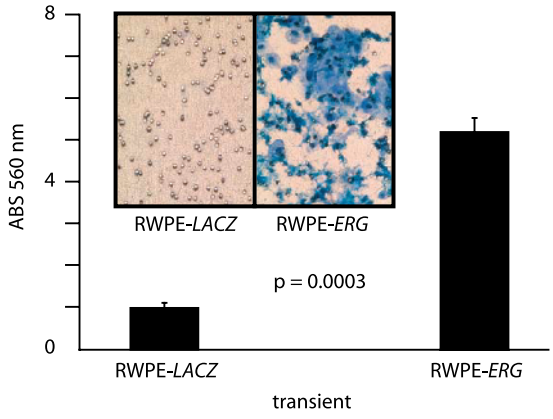

Figure W3. Transient over-expression of *ERG* increases invasion in RWPE cells. We infected the benign immortalized prostate cell line RWPE with *ERG* or *LACZ* adenovirus and assayed for invasion through a modified basement membrane, mean ($n = 3$) ± SEM. Inset shows photomicrograph of invaded cells.

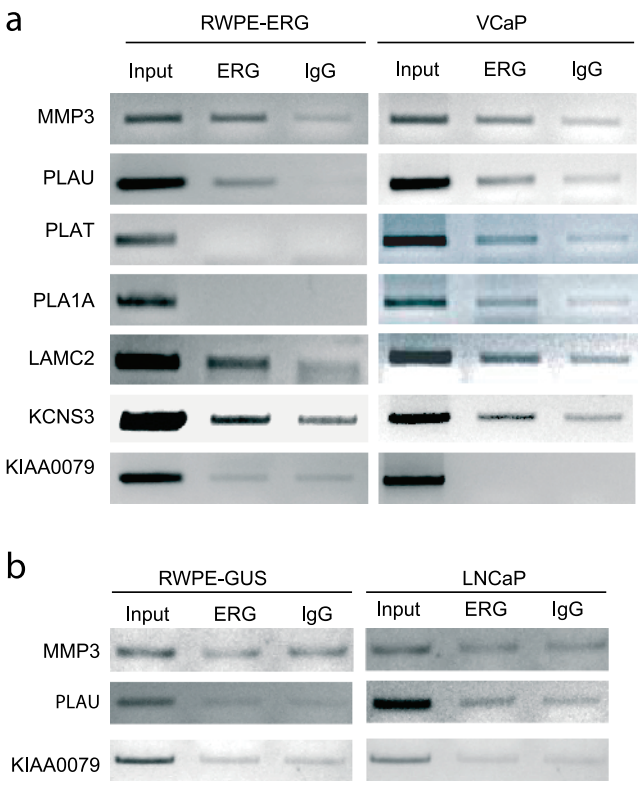

Figure W4. Chromatin immunoprecipitation across *TMPRSS2–ERG* model systems. (a) Chromatin immunoprecipitation to detect enrichment of ERG binding to the proximal promoters of indicated genes compared to IgG control in RWPE-*ERG* and VCaP cells. The promoter of KIAA0089 was used as a negative control. (b) RWPE-*GUS* and LNCaP failed to show any enrichment of ERG binding to assayed promoters.

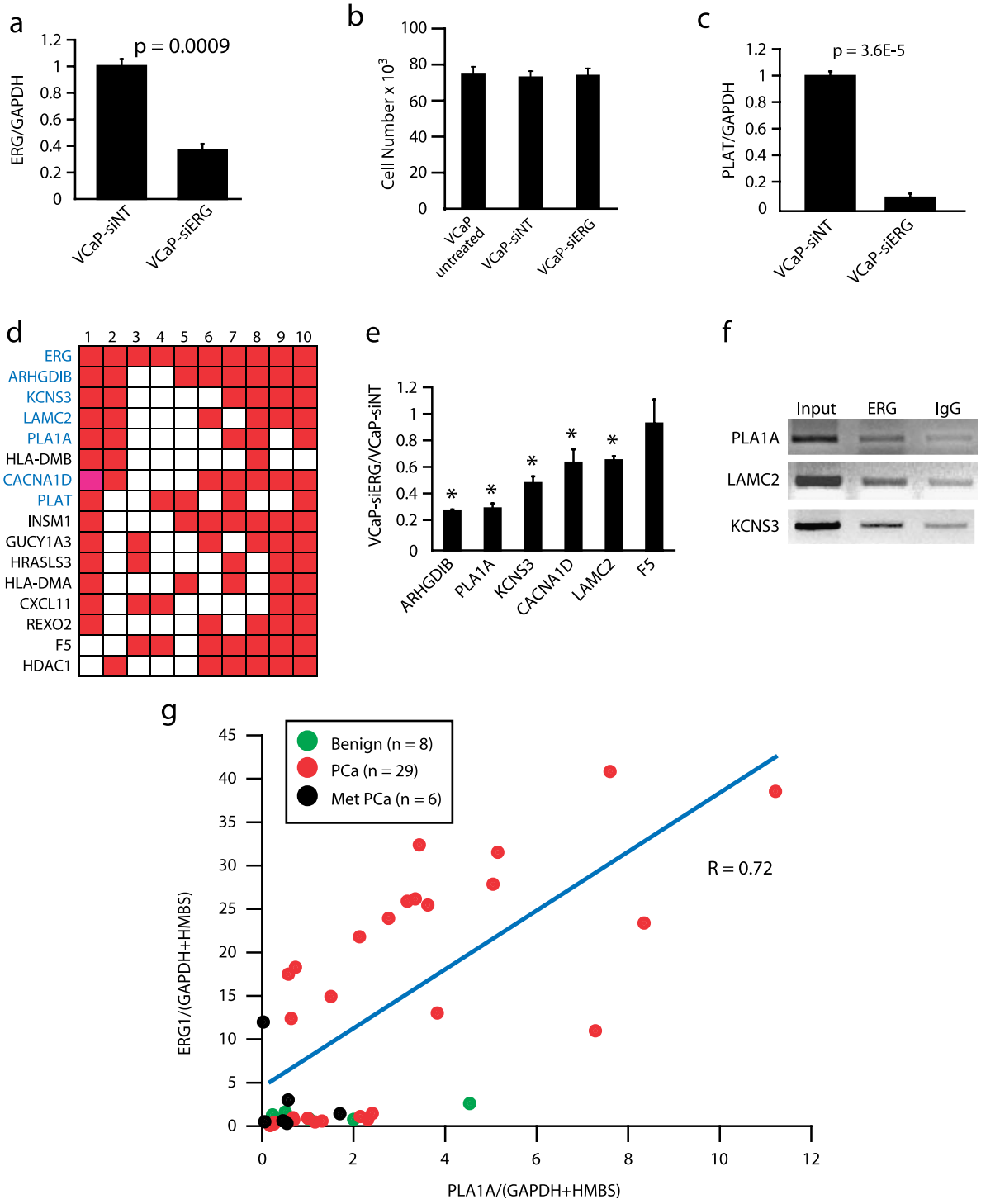

Figure W5. *ERG* knockdown in VCaP attenuates a transcriptional program over-expressed in *TMPRSS2:ETS*–positive tumors. (a) siRNA knockdown of *ERG* in the *TMPRSS2–ERG*–positive prostate cancer cell line VCaP. VCaP cells were either treated with transfection reagent alone (untreated) or transfected 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.

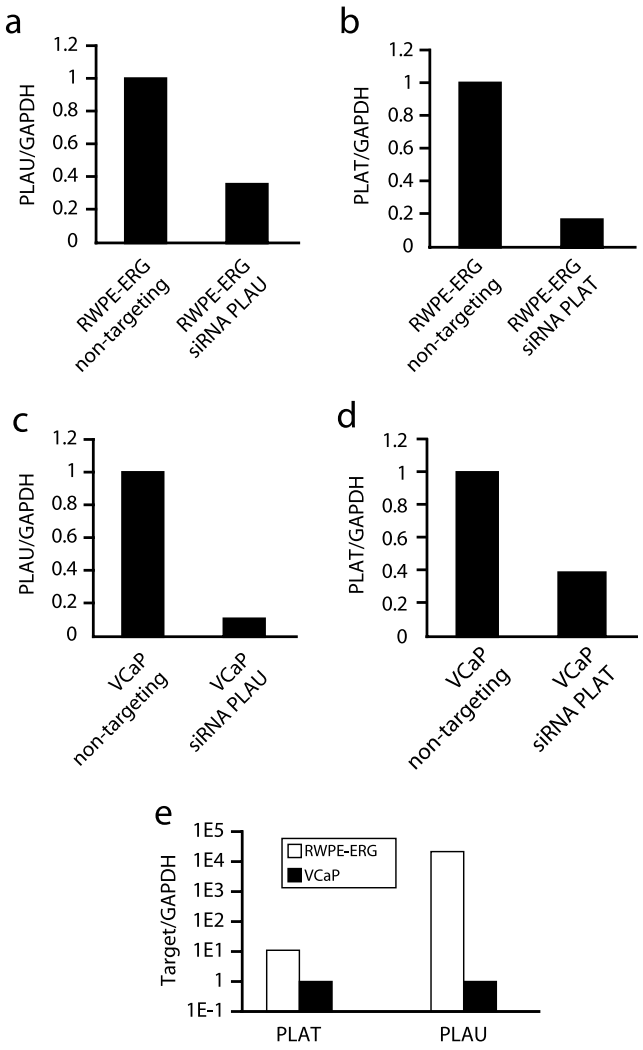

Figure W6. qPCR confirmation of *PLAU* and *PLAT* knockdown in RWPE-*ERG* and VCaP cells. (a and b) RWPE-*ERG* cells were treated with non-targeting siRNA or siRNA against (a) *PLAU* or (b) *PLAT*, and knockdown was confirmed by qPCR. (c and d) VCaP cells were treated with nontargeting siRNA or siRNA 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.

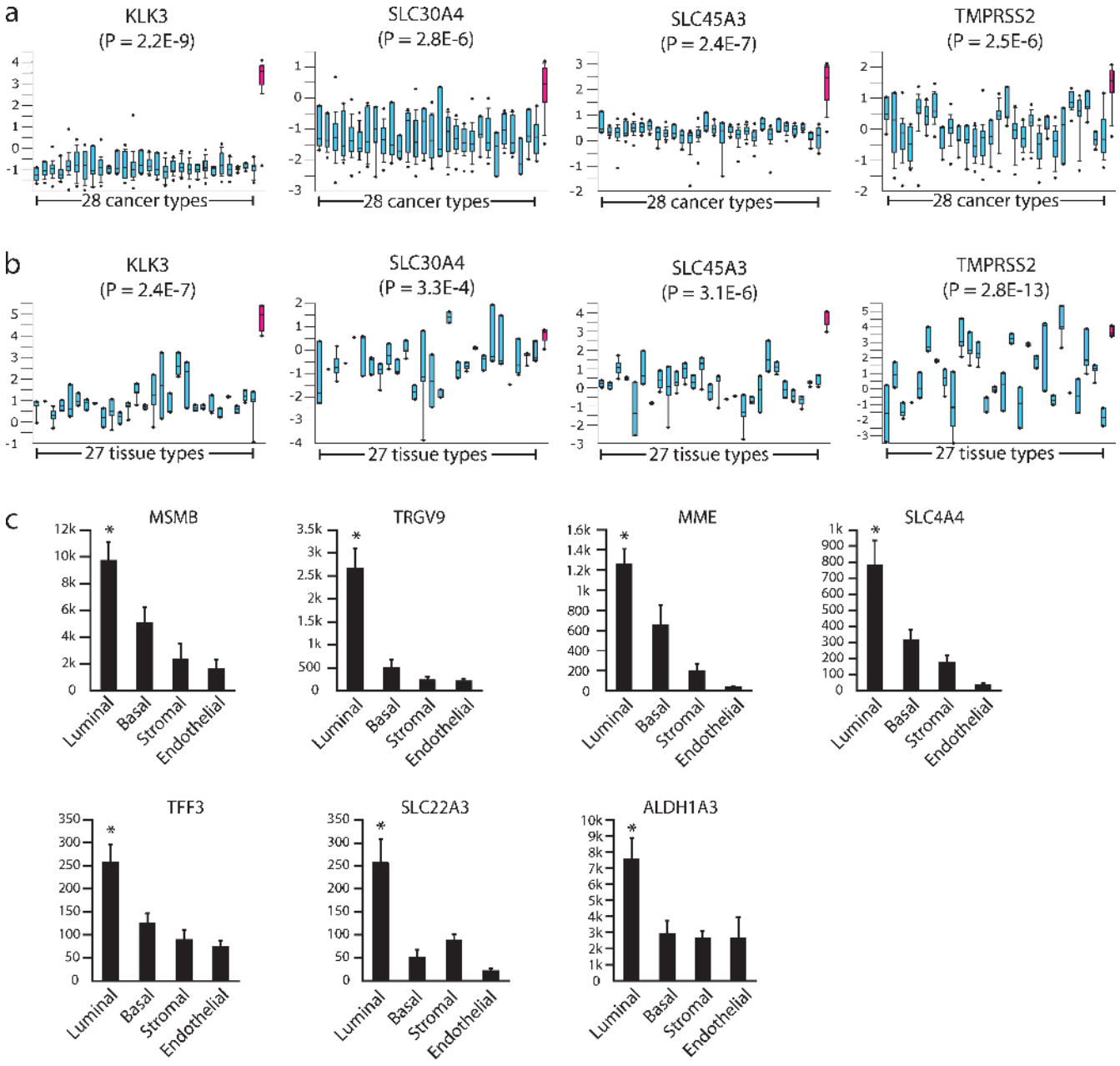

Figure W8. Prostate epithelial specificity of genes induced in VCaP on ERG knockdown. (a) Genes confirmed by qPCR to be over-expressed in VCaP cells treated with *ERG* siRNA were interrogated in the expO multicancer data set, containing expression profiles from 28 cancer types (blue) and prostate cancer (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 ($n = 5 \pm$ SEM) are shown. $^*P < .05$, for all pairwise *t* tests involving luminal cells.