NEOPLASIA
www.neoplasia.com

Volume 12 Number 12 December 2010 pp. 1031–1040 1031

ERG Cooperates with Androgen Receptor in Regulating Trefoil Factor 3 in Prostate Cancer Disease Progression[1,2,3]

David S. Rickman*,[4], Ying-bei Chen*,[4], Samprit Banerjee†, Yihang Pan*, Jindan Yu‡, Terry Vuong*, Sven Perner*, Christopher J. Lafargue*, Kirsten D. Mertz§, Sunita R. Setlur¶, Kanishka Sircar#, Arul M. Chinnaiyan‡, Tarek A. Bismar, Mark A. Rubin*,[5] and Francesca Demichelis*,††,[5]**

*Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY, USA; †Department of Public Health, Weill Cornell Medical College, New York, NY, USA; ‡Department of Pathology, University of Michigan, Ann Arbor, MI, USA; §Department of Pathology, Institute of Surgical Pathology, University Hospital Zurich, Zurich, Switzerland; ¶Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; #Department of Pathology, MD Anderson Cancer Center, Houston, TX, USA; **Departments of Pathology and Laboratory Medicine & Oncology, University of Calgary, Calgary, Canada; ††Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY, USA

Abstract

To elucidate the role of ETS gene fusions in castration-resistant prostate cancer (CRPC), we characterized the transcriptome of 54 CRPC tumor samples from men with locally advanced or metastatic disease. Trefoil factor 3 (*TFF3*) emerged as the most highly differentially regulated gene with respect to *ERG* rearrangement status and resistance to hormone ablation therapy. Conventional chromatin immunoprecipitation (ChIP)–polymerase chain reaction and ChIP followed by DNA sequencing (ChIP-seq) revealed direct binding of ERG to ETS binding sites in the *TFF3* promoter in *ERG*-rearranged prostate cancer cell lines. These results were confirmed in *ERG*-rearranged hormone-naive prostate cancer (HNPC) and CRPC tissue samples. Functional studies demonstrated that ERG has an inhibitory effect on TFF3 expression in hormone-naive cancer but not in the castration-resistant state. In addition, we provide evidence suggesting an effect of androgen receptor signaling on ERG-regulated TFF3 expression. Furthermore, TFF3 overexpression enhances ERG-mediated cell invasion in CRPC prostate cancer cells. Taken together, our findings reveal a novel mechanism for enhanced tumor cell aggressiveness resulting from *ERG* rearrangement in the castration-resistant setting through TFF3 gene expression.

Neoplasia (2010) 12, 1031–1040

Address all correspondence to: Francesca Demichelis, PhD, Department of Pathology and Laboratory Medicine, Institute for Computational Biomedicine, Weill Cornell Medical College, 1305 York Ave, Y 1307 (or Box 140), New York, NY 10065. E-mail: frd2004@med.cornell.edu

[1]This study was funded by the Department of Defense New Investigator Award PC081337 (D.S.R.), the National Cancer Institute R01 CA125612-01 (F.D. and M.A.R.), the Prostate Cancer Canada, and The Young Investigator Award of the Prostate Cancer Foundation, USA (T.A.B.).

[2]S. Perner, A. M. Chinnaiyan, M. A. Rubin, and F. Demichelis are coinventors on a patent filed by The University of Michigan and The Brigham and Women's Hospital covering the diagnostic and therapeutic fields for ETS fusions in prostate cancer. The diagnostic field has been licensed to Gen-Probe, Inc.

[3]This article refers to supplementary materials, which are designated by Tables W1 and W2 and Figures W1 to W3 and are available online at www.neoplasia.com.

[4]These authors equally contributed to this study.

[5]These authors share the senior authorship.

Received 18 June 2010; Revised 17 August 2010; Accepted 24 August 2010

DOI 10.1593/neo.10866

Introduction

Most patients with locally advanced or metastatic prostate cancer receiving androgen deprivation therapy (ADT) progress to a castration-resistant state that leads to death. The systemic treatment options for castration-resistant prostate cancer (CRPC) are limited. Docetaxel-based chemotherapy is the only therapy proven by randomized trials to provide a modest 2- to 3-month survival benefit [1,2]. The lack of effective treatment reflects a poor understanding of the molecular underpinnings of CRPC.

Most prostate cancers harbor a recurrent ETS gene fusion characterized by a rearrangement between the 5′ regulatory elements of an androgen-regulated gene and the coding region of a member of the ETS gene family of transcription factors, resulting in an androgen-driven expression of the ETS transcription factor [3]. The most common rearrangements involve the ETS transcription factor *ERG* and one of three known 5′ androgen-regulated promoters: *TMPRSS2*, *SLC45A3*, and *NDRG1* [4]. Fifty percent of all prostate cancers harbor one of these rearrangements. A recent study [5] demonstrated that transcription regulation differs based on the androgen context. In support of this, Yu et al. [6] have recently characterized the genome-wide location of AR, ERG, and epigenetic marks. In their study, they found that AR-bound genes were enriched with a large number of ERG-associated genes and, similarly, that ERG is bound to more than 90% of AR-bound genes, showing a high degree of co-occupancy. Moreover, their study showed that ERG physically interacts with AR and can inhibit its binding and its activity leading to a disruption of AR-mediated differentiation of the prostate.

It has been shown that ERG is expressed at comparable levels in *ERG*-rearranged hormone-naive prostate cancer (HNPC) and CRPC [7]. To gain insight into how *ERG*-rearranged prostate cancers differ molecularly from nonrearranged cancers, we initially interrogated the transcriptome of 354 HNPC tumor samples, leading to the key observation that estrogenic signaling through an estrogen-binding site on the *TMPRSS2* promoter might explain continued expression of the fusion transcript in a CRPC state [8].

Encouraged by these results, and to gain insight into how *ERG* rearrangement may lead to a distinct perturbation of regulatory pathways, we investigated the expression profile of CRPC. Few expression array studies have explored CRPC owing to the difficulty of obtaining frozen tissue, and they were limited to a small number of samples [9–12]. To overcome this challenge, we exploited the newly developed complementary DNA–mediated annealing, selection, extension, and ligation (DASL) expression profiling platform for formalin-fixed paraffin-embedded (FFPE) tissue, which allowed us to interrogate the expression of more than 6000 transcriptionally informative genes [8] in transurethral resection biopsies of CRPC patients.

Materials and Methods

Patients and Samples

The cohort initially included 83 transurethral resection of prostate samples from 59 patients who had been treated with one or multiple ADT protocols at McGill University Hospitals (Montreal, Canada). The clinical characteristics of 35 patients with complete information are summarized in Table W1. The castration-resistant status was determined clinically based on the prostate-specific antigen (PSA) levels and disease progression under treatment. Archived FFPE samples were retrieved with approval from the local institutional review boards. Using a recently described protocol [8], tissue cores at diameters of 0.6 and 1.5 mm were obtained from areas containing high-density tumor and were subjected to tissue microarray (TMA) construction and RNA extraction, respectively.

Fresh-frozen metastatic prostate cancer tissue samples were obtained from the University of Michigan Rapid Autopsy Program [13]. Frozen HNPC samples were obtained from men with localized and locally advanced prostate cancer who underwent radical prostatectomy as a monotherapy and were processed as previously described [4]. Publicly available expression profiling data from the tumor samples of 354 men with HNPC (clinical stage T1-T2, N0, Mx) followed on a Watchful Waiting protocol were also used in the analysis of current study (GEO series accession number GSE8402) [8].

Fluorescent In Situ *Hybridization*

ERG rearrangement was evaluated using a break-apart probe assay on TMAs as previously described [14]. Trefoil factor 3 (*TFF3*) rearrangement was assessed by applying a break-apart probe set consisting of the Biotin-14-dCTP-labeled BAC clone RP11-891L10 (eventually conjugated to produce a red signal) and the digoxigenin-dUTP–labeled BAC clone RP11-53A21 (eventually conjugated to produce a green signal), both spanning the neighboring centromeric and telomeric region of the *TFF3* locus, respectively. BAC clones were selected from the March 2006 build of the human Genome using the UCSC Genome Browser and were obtained from the BACPAC Resource Center, Children's Hospital Oakland Research Institute (Oakland, CA).

Morphologic Evaluation

Histologic features of prostate cancer after ADT were assessed using a previously described scoring system that was shown to be highly reliable for determining the effects of hormonal treatment [15]. Certain morphologic features such as small cell carcinoma (neuroendocrine) differentiation were annotated for the cases.

Expression Profiling

Complementary DNA–mediated annealing selection and ligation assay (DASL) (Illumina, San Diego, CA) was used to profile CRPC tumor samples available in FFPE blocks. The DASL method exponentially amplifies transcripts of interest and has high sensitivity for genes with low expression levels. The platform used for this study consists of four DASL assay panels (DAP) covering approximately 6000 genes selected based on the evaluation of more than 250 microarray data sets (database available at http://www.broad.mit.edu/cancer/pub/HCC). This 6K DASL platform has been recently used to evaluate more than 350 clinically localized prostate cancer [8] and 307 cases of hepatocellular carcinoma [16]. We jointly analyzed the CRPC cohort from Montreal and the HNPC Swedish Watchful Waiting cohort [8]. To resolve the potential study design confounder (expression data for the CRPC and HNPC Swedish cohort were generated at different time in different centers), we used the gene expression data from nine clinically localized prostate cancer (experiment control samples), which were profiled together with the CRPC cohort in the same experiment. All statistical analyses were performed with the R statistical software (www.r-project.org). For a more detailed description of data processing and analyses, please see Supplemental Materials and Methods. At the completion of the morphologic evaluation and of the gene expression profiling data quality assessment, interpatient and intrapatient sample correlations were assessed. In this process, when more than one tumor focus was present for a given patient from a single biopsy or multiple biopsies, we retained data only from foci with intrapatient differences

in gene expression (correlation coefficient < 0.7). Fifty-four tumor samples from 52 individuals passed these quality control measures.

Immunohistochemistry and Quantification

Serial sections of TMAs containing the CRPC cases and additional localized HNPC cases were prepared. Immunohistochemical staining of TFF3 and AR were performed on a Bond Max Autostainer (Leica Microsystems, Bannockburn, IL) using the following antibodies: anti-TFF3 (clone 15C6, 1:250; Calbiochem, Gibbstown, NJ) and anti-AR (clone F39.4.1, 1:800; Biogenex, San Ramon, CA). Briefly, sections were deparaffinized, and endogenous peroxidase was inactivated. Antigen retrieval was accomplished using the Bond Epitope Retrieval Solutions (Leica Microsystems) at 99 to 100°C for 20 minutes. The sections were then incubated sequentially with the primary antibody for 25 minutes after the primary for 15 minutes and polymer for 25 minutes followed by colorimetric development with diaminobenzidine for 10 minutes. The Ariol imaging system (Applied Imaging, San Jose, CA) was used for quantification of protein expression in TMAs. Immunohistochemistry data were analyzed separately for the two arrays containing HNPC and CRPC, respectively.

Cell Lines

The prostate cell lines including benign epithelial cells RWPE-1, cancer cell lines (VCaP, LNCaP, 22Rv1, PC3, and DU145), and prostate stromal cells PrSc, as well as the embryonic kidney epithelial cells HEK-293 were purchased from ATCC (Manassas, VA) and maintained according to the manufacturer's protocols. We used pLenti6.3/V5-TOPO plenti6.3/V5-GW/lacZ vectors (Invitrogen, Carlsbad, CA) to generate stable HEK-293 cells expressing either the FLAG (DYKDDDDK)–tagged truncated ERG (HEK-293tERG) or LacZ (HEK-293LacZ), respectively. To generate stable tERG-expressing RWPE-1 and DU145 cells, the tERG-FLAG open reading frame was subcloned into a retroviral virus expression vector (pBABE; a kind gift from Dr. William Hahn from the Broad Institute and the Dana Farber Cancer Center), which is equipped with an IRES-driving green fluorescent protein (GFP) expression. For transient overexpression of TFF3, we obtained the pcAGGS-TFF3 expression vector and pcAGGS control vector (kind gifts from Dr Steven Itzkowitz, Mount Sinai School of Medicine, New York, NY). These vectors were transfected into DU145-tERG or DU145-GFP cells using *Trans*IT-Prostate Transfection Kit (Mirus Bio, Madison, WI) following the manufacturer's instructions. TFF3 expression messenger RNA (mRNA) was assessed, and invasion assay was performed 48 hours after transfection as described below. VCaP cells were treated with 50 nM small interfering RNA (siRNA; scrambled or against the TFF3 mRNA) using Mirus TransIT-TKO and OPTI-MEMI (Mirus Bio LLC, Madison, WI) serum-free medium.

Real-time Quantification of ERG and TFF3 mRNA

After RNA extraction, the quality of RNA was assessed using the Bioanalyzer 2100 (Agilent Technologies, Inc, Santa Clara CA). Quantitative reverse transcription–polymerase chain reaction (RT-PCR) was performed using Power SYBR Green RNA-to-Ct (Applied Biosystems, Foster City, CA) following the manufacturer's protocol. Each sample was run in triplicate. The amounts of target genes were calculated relative to the reference gene *HMBS* using the comparative C_t method (ABI Bulletin 2; Applied Biosystems). For this, raw C_t values for *ERG* or *TFF3* were first normalized using the average C_t values obtained for *HMBS* and then calibrated using normalized C_t values obtained from one of the control cell lines (HEK-293LacZ). The primer sequences for *ERG*, *TFF3*, and *HMBS* are as follows: 5′-CGCAGAGTTATCGTGCCAGCAGAT-3′ and 5′-CCATATTCTTTCACCGCCCACTCC-3′ (*ERG*); 5′-TGTGCCGTGCCAGCCAAG-3′ and 5′-CTGGAGGTGCCTCAGAAGGTG-3′ (*TFF3*); 5′-CCATCATCCTGGCAACAGCT-3′ and 5′-GCATTCCTCAGGGTGCAGG-3′ (*HMBS*).

Chromatin Immunoprecipitation and Quantitative PCR

We used MatInspector (version 8.0; Genomatix Software GmbH, Munich, Germany) [17] to identify ETS family transcription factor binding sites in the 878 bp of genomic sequence near the transcription start site for *TFF3 in silico*. This region included 626 bp of upstream and 252 bp of downstream sequence. Briefly, 75×10^6 HEK-293tERG or HEK-293LacZ cells were washed in PBS twice and then fixed using 1% formaldehyde for 10 minutes at room temperature and quenched using 125 mM glycine. The cells were centrifuged, and the cell pellet was resuspended in 2 ml of dilution buffer (165 mM NaCl, 0.01% SDS. 1.1% Triton X-100, 1.2 mM EDTA pH 8.0, 16.7 mM Tris-HCl pH 8.0, 1 mM phenylmethylsulfonyl fluoride). Protein-bound chromatin was fragmented by sonication for 10 minutes (cycles of 30-second pulses of sonication followed by 30 seconds of rest). Equal volumes of chromatin were immunoprecipitated with either mouse anti-FLAG conjugated agarose (Sigma, St Louis, MO) or rabbit anti-ERG (GeneTex, Inc, Irvine, CA), rabbit anti-actin (Cell Signaling, Danvers, MA), or mouse immunoglobulin G (IgG) antibodies/protein A agarose (Sigma). After extensive washing, the DNA was eluted using 100 mM $NaHCO_3$ and 1% SDS, and the cross-links were reversed using 300 mM NaCl at 65°C for 16 hours. The eluted DNA was purified using Qiagen PCR Qiaquick kit following the manufacturer's protocol. For quantitative PCR amplification, we used the ABI 7500fast system (Applied Biosystems Life Technologies Corporation, Carlsbad, CA) and the relative standard curve method in a 96-well format. For this, we designed primer sets that target each of the three ETS bindings sites identified in the *TFF3* promoter (Table W2). Two microliters of either eluted DNA or a 1:10 dilution of the input chromatin preparation from each cell line was assayed to calculate the percentage of enrichment. Primers targeting a copy number–stable chromosomal region in *ARHGEF11* (chr1:155205397-155205600, hg18) were used as a negative control as previously described [18]. Input DNA was also analyzed at five concentrations (0.004-40 ng) to generate the standard curve per primer pair and per 96-well plate. All reactions were run in triplicates. For chromatin immunoprecipitation (ChIP) using the tissue samples, we followed the protocol described below for ChIP-seq experiments with one modification. Frozen tissue cores or eight 20-μm sections were chopped into small pieces, fixed in 1% formaldehyde, quenched with glycine as described below, and homogenized to disaggregate the tissue. Otherwise, ChIP was performed using either mouse anti-FLAG–conjugated agarose (Sigma) or rabbit anti-ERG (GeneTex, Inc), rabbit anti-actin (Cell Signaling), or mouse IgG antibodies/protein A agarose (Sigma) essentially as previously described [19].

ChIP-seq

VCaP or LNCaP cells were grown in RPMI supplemented with 10% FBS and Glutamax to 90% confluence. Cells were harvested for ChIP using an anti-ERG antibody (sc-543x; Santa Cruz Biotechnology, Santa Cruz, CA) or IgG control as previously described [20]. ChIP-seq was performed using Illumina Genome Analyzer according to standard manufacturer's procedures. The raw sequencing image data were analyzed by the Illumina analysis pipeline, aligned to the unmasked human reference genome (NCBI v36, hg18) using the ELAND software (Illumina) to generate sequence reads of 25-35 bps. ChIP-enriched

binding peaks were defined by HPeak (http://www.sph.umich.edu/csg/qin/HPeak/).

Cell Invasion and Migration Assays

Cells were harvested with trypsin/EDTA and resuspended in complete medium with 10% fetal calf serum. After centrifugation, the cells were washed with serum-free medium and centrifuged. Cell pellets were resuspended in serum-free medium, and cell counts were performed in a hematocytomer chamber to dilute the cells to a concentration of 2×10^5 cells/ml. Boyden chambers coated with Matrigel (BD Biosciences, San Jose, CA) or with 8-μm pore inserts (BD Biosciences) were used for the invasion assay or the migration assay, respectively. This was performed by placing 5×10^4 cells into the upper chamber and 0.75 ml of medium containing 10% fetal calf serum into the lower chamber. After a 24- or 48-hour incubation at 37°C, the cells attached to the upper side of the membrane were removed with cotton swabs and cells that migrated into the lower side of the membrane were fixed with 10% formaldehyde, stained with Coomassie blue and washed with PBS. Micrographs were taken with a Retica 1300i CCD camera on a Nikon eclipse 80i microscope (Nikon, Melville, NY) using a 4× objective lens. For a given experiment, each assay was performed in triplicate. Cell number per field was counted and analyzed using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA). For the VCaP siRNA knock-down/invasion assay, 1.5×10^5 cells were plated into the Boyden chambers, and the lower chamber contained NIH-3T3 medium supplemented with 10 ng/ml of epidermal growth factor and 5 ng/ml of insulin-like growth factor and incubated for 72 hours. The membranes were processed as described above.

Results

Histologic Characterization of CRPC

The castration-resistant status of prostate cancer patients was determined clinically based on rising PSA and/or other evidence of disease progression while receiving ADT treatment. A previous work from Mostaghel et al. [21] showed that despite 9 months of ADT treatment, complete intratumoral blockade of androgen-regulated gene expression was not achieved. To ensure that our study would focus on established molecular alterations due to ADT, not only did we select CRPC samples from patients that had been treated for more than 8 months (range = 8-92 months, median = 32 months; Table W1), we also evaluated the hormonal treatment effect of each case using previously described histologic criteria [15]. The complete range of histologic features is represented in Figure 1*A*. We identified 47 samples with unequivocal ADT effect (Figure 1*A*, *first three panels*) and 7 samples of small cell neuroendocrine carcinoma of the prostate, which commonly displays minimal morphologic change with ADT (Figure 1*A*, *far right panel*), a total of 54 samples from 52 patients, to be included in the transcriptome profiling analysis.

We next determined the *ERG* rearrangement status of these CRPC using an *ERG* break-apart fluorescent *in situ* hybridization assay. Of 54 samples, 18 (33%) demonstrated the *ERG* rearrangement either through interstitial deletion ($n = 12$) or through insertion ($n = 8$). Four of the seven small cell carcinoma samples demonstrated *ERG* rearrangement (Figure 1*B*), supporting their prostatic origin [22]. It has been suggested that *ERG* rearrangement may not be associated with an active fusion transcript with disease progression [23]. We investigated the levels of three main *TMPRSS2-ERG* fusion transcripts (type I, III,

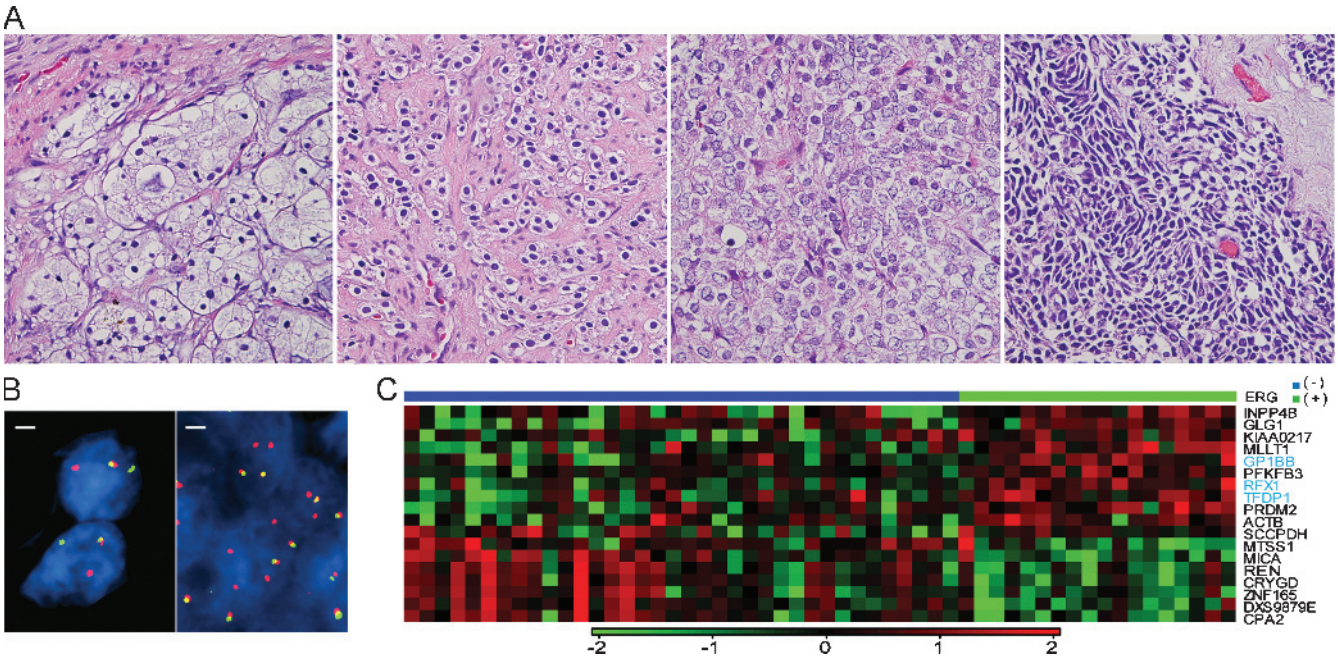

Figure 1. CRPC is morphologically and genetically heterogeneous and harbors *ERG* rearrangement in a subset of cases. (A) Distinct morphologic types of CRPC including adenocarcinoma with unequivocal hormone treatment effect (A, first three panels) and small cell carcinoma minimal hormone treatment effect (A, far right panel). (B) Small cell prostate carcinomas harbor *ERG* rearrangement occurring through insertion (left panel) or deletion with accompanying polysomy (right panel). Scale bar, 1 μm. (C) Heat map of the supervised gene expression differentiation analysis results between *ERG*-rearranged and nonrearranged CRPC. The 18 genes (FDR < 20%) are ranked according to the levels of expression alterations. Three genes (in light blue font) overlap with a recently described 87 gene signature from *ERG*-rearranged HNPC.

and VI) [24] in CRPC cases (Figure W1). Among 19 cases of *ERG*-rearranged CRPC, 14 (73%) showed detectable ERG fusion transcripts.

Expression Profiling Using DASL

Having identified and characterized the CRPC cases in the cohort, we undertook expression profiling of the FFPE tissue using DASL [8,16] where we applied previously described criteria for high-quality gene expression data [8] and interpatient and intrapatient correlation analysis (see Materials and Methods).

Using a supervised approach to investigate genes differentially expressed in *ERG*-rearranged CRPC, 18 differentially expressed genes (false discovery rate [FDR] < 20%, Wilcoxon rank sum test) were identified based on *ERG* rearrangement status (Figure 1*C*).

To explore the nature of the molecular signature due to ADT, we compared these 18 genes to a previously identified 87 gene signature of clinically localized *ERG*-rearranged HNPC cases from a Watchful Waiting Cohort [8]. Of 18 genes, 3 overlapped (Figure 1*C*, probability of detecting three common genes by chance is equal to 0.0045), namely *GP1BB*, *RFX1*, and *TFDP1*, suggesting that, although many molecular mechanisms crucial for *ERG*-rearranged prostate cancer are altered in CRPC, some ETS rearrangement–related molecular regulations are maintained independent of hormonal state. Among the top three differentially regulated genes in *ERG*-rearranged CRPC was *PFKFB3*, a unique member of the enzyme family 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase that regulates the cellular concentration of fructose-2,6-biphosphate (F2,6BP) (see Supplemental Discussion).

ERG-Rearranged HNPC and CRPC Show Inverse Patterns of Differential Regulation of TFF3

To examine the molecular signaling of *ERG*-rearranged prostate cancer in the context of androgen-enriched or androgen-deplete environments, we performed statistical interaction analysis on the expression profiling data from 408 tumor samples (354 HNPC and 54 CRPC) modeling *ERG* status and hormonal treatment status. We identified 25 genes with significantly altered expression patterns in CRPC compared with HNPC that were related to the *ERG* rearrangement status. These genes may be related to the role of *ERG* rearrangement during the progression from HNPC to CRPC (Table 1).

TFF3, previously identified as a putative prostate cancer biomarker with overexpression observed in only a subset of cases [25,26], was the top-ranked gene. Our analysis identified a significant expression pattern change between CRPC and HNPC in relation to *ERG* rearrangement ($P = 0$, permutation test; see Combined Analysis of CRPC and HNPC in Supplemental Materials and Methods). In HNPC, TFF3 was significantly downregulated in *ERG* rearrangement–positive compared with *ERG*-negative tumors ($P = 3.1e-09$, for a test of differential expression by Wilcoxon rank sum test), but this pattern was altered in an opposite direction in CRPC with *ERG*-rearranged cases showing a tendency for up-regulation ($P = .055$, for a test of differential expression by Wilcoxon rank sum test). We confirmed the two patterns at the protein level by immunohistochemistry quantified using an automated imaging system (Figure 2, *A* and *B*; $P = .03$ for the test of interaction through permutations as described in the Materials and Methods section; $P = .07$ for Wilcoxon rank sum test comparing ERG rearrangement in HNPC and $P = .09$ for the same in CRPC). These data suggest that *TFF3* expression is downregulated during the transition from HNPC to CRPC for the *ERG* rearrangement–negative tumors. Interestingly, we found that the highest levels of TFF3 protein in the *ERG*-rearranged CRPC were in cases with undetectable or low AR levels (0%-25% quartile of AR expression; Figure 2*C*). This relationship was not observed in HNPC (data not shown).

Table 1. Genes That Were Found to Be Highly Significant ($P = 0$, Permutation Test) Related to *ERG* Rearrangement and Progression to CRPC under ADT Identified by the Interaction Model.

Gene Symbol	P^* ($F^\dagger$)	P^* ($H^\dagger$)	Estimate‡ (F)	Estimate‡ (F)	Estimate‡ ($F \times H$)
TFF3	0	.049791	−0.5166	−1.08341	1.012905
INPP4B	.055974	.89646	−0.45506	−0.39481	1.355811
BSG	.060216	.546415	0.007968	−0.03601	0.334732
HNRPH2	.088557	.74128	0.058647	−0.37028	0.629867
AZGP1	.16011	.951432	−0.3513	−0.40521	0.919148
PCBP1	.173117	.680504	0.012675	−0.17463	0.250093
SLC4A4	.186616	.147655	−0.42278	−1.25905	1.35686
PHKG2	.202698	.062971	0.014572	0.948933	−0.73296
CSRP1	.230435	.071665	0.007243	−0.83526	0.624356
CAPZB	.303962	.49861	−0.00394	−0.57613	0.763955
LMAN1	.360658	.264247	−0.04036	−0.49036	0.617975
CAV1	.425037	.240416	−0.0102	0.210558	0.654177
SEMA3C	.436939	.737742	−0.20679	−0.56852	0.844928
GLG1	.493471	.763193	−0.08542	−0.21319	0.899588
NAP1L4	.535457	.125271	0.111557	−0.15817	−0.4481
RBBP7	.675381	.090465	−0.19193	−1.01176	0.857275
ENO1	.689337	.84047	−0.02496	−0.05213	0.230198
RNASE4	.698196	.8256	−0.20992	−0.62507	1.060066
RPL13A	.768596	.63122	−0.0518	−0.05548	0.380193
LDHA	.785075	.814296	0.11248	0.283165	−0.47686
GNE	.81904	.442845	−0.07524	−0.50476	0.56862
CA9	.847723	.763193	−0.1406	−0.25086	0.96787
MLLT1	.887444	.143074	−0.04602	0.099475	0.323797
ARD1A	.931244	.502358	−0.0521	−0.00137	0.27497
PRPF4B	.994072	.507388	0.077961	0.005466	−0.4533

*The P value (computed using the permutation test) corresponds to the test of its corresponding coefficient being equal to 0, representing the significance of each factor's contribution to the interaction model (see Materials and Methods). Specifically, the F P value and the H P value correspond to test of significance of the marginal effect of the *ERG* rearrangement status and hormone treatment status. Testing the marginal effect means comparing the differential expression across two levels of a factor irrespective of the levels of the other factor. For example, the marginal effect of *ERG* rearrangement status would compare *ERG* rearranged cases to those not rearranged in both CRPC and HNPC.

†F, H, and $F \times H$ correspond to the effects of *ERG* rearrangement (F), ADT (H), and the interaction between the two factors ($F \times H$), respectively.

‡The estimates correspond to the least square estimates for α, β, and γ in the interaction model (see Supplemental Materials and Methods).

Similar to the expression of *TFF3* in *ERG*-rearranged CRPC, we also observed increased *TFF3* expression in a prostate cancer cell line derived from therapy-resistant metastatic prostate cancers harboring *ERG* rearrangement and expressing AR (i.e., VCaP) but not in cell lines without the rearrangement (i.e., LNCaP, PC3, and 22Rv1) or in prostate stromal cells (i.e., PrSc; Figure 2*D*).

TFF3 *Expression Is Directly Regulated by* ERG

The three members of the TFF gene family are located near the *TMPRSS2* gene on chromosome 21q22.3, with *TFF3* being the closest at roughly 800 kb from the *TMPRSS2* transcription start site (Figure 3*A*). To exclude the possibility that the down-regulation of TFF3 in *ERG*-rearranged HNPC is driven by a genomic alteration, we interrogated a cohort of 18 HNPC cases on a test TMA using a dual-color *TFF3* break-apart fluorescent *in situ* hybridization assay and did not detect *TFF3* deletion or rearrangement (data not shown). By *in silico* analysis, we identified three putative ETS family transcription factor–binding sites within 878 bp of genomic sequence surrounding the *TFF3* transcription start site (Figure 3*A* and Table W2). On the basis of results from a ChIP assay, we detected a direct interaction of ERG to two of these three ETS binding sites (EBS1-3) in HEK-293 cell lines overexpressing the truncated open reading frame of ERG (HEK-293tERG) encoded by the most prevalent *TMPRSS2-ERG* fusion transcript type III (Figure 3*B*). Parallel ChIP using species-specific control

antibodies was performed to assess the specificity of the ChIP assays (Figure W2*A*). The nonbiased approach of ChIP followed by deep sequencing (ChIP-seq) performed in VCaP cells (androgen receptor–positive and harboring a *TMPRSS2-ERG* fusion) revealed specific binding of ERG to a region that covers EBS2 and 3 in the proximal *TFF3* promoter (Figure 3*C*). As a negative control, ChIP-seq of ERG performed in LNCaP cells that do not harbor *TMPRSS2-ERG* gene fusions detected background level enrichment with less than 700 peaks (data not shown), in contrast to more than 40,000 ERG binding peaks in VCaP cells. ERG binding was also observed at a distant site roughly 25 kb upstream of the transcription start site (Figure W2*B*). To determine whether ERG binds to these ETS sites in human samples, we performed ChIP with HNPC or metastatic prostate cancer tissue samples that harbor *TMPRSS2-ERG* gene fusions. Because of the limited amount of material, we focused on the two ETS sites showing the highest level of binding *in vitro*. As with the cells lines, we found specific binding of ERG to the *TFF3* promoter compared with benign prostate tissue (Figures 3*D* and W2*C*).

We next examined the effect of ERG binding on the expression of *TFF3*. In HEK-293tERG cells, the level of TFF3 is low compared with the control cells (Figure 4*A*, *left*), suggesting an inhibitory effect of ERG on *TFF3* expression. However, knocking down the expression of *ERG* fusion transcripts using short hairpin RNA (shRNA) molecules in VCaP cells resulted in a decreased level of TFF3 expression, suggesting an opposite effect of ERG in this cell context. VCaP cells are known to not only harbor *ERG* rearrangement but also maintain androgen receptor expression [3]. We therefore wanted to examine the effect of androgen signaling on TFF3 expression and test the hypothesis that the regulation of *TFF3* by ERG differs as a function of androgen signaling. To this end, we generated stable RWPE-1 cells overexpressing

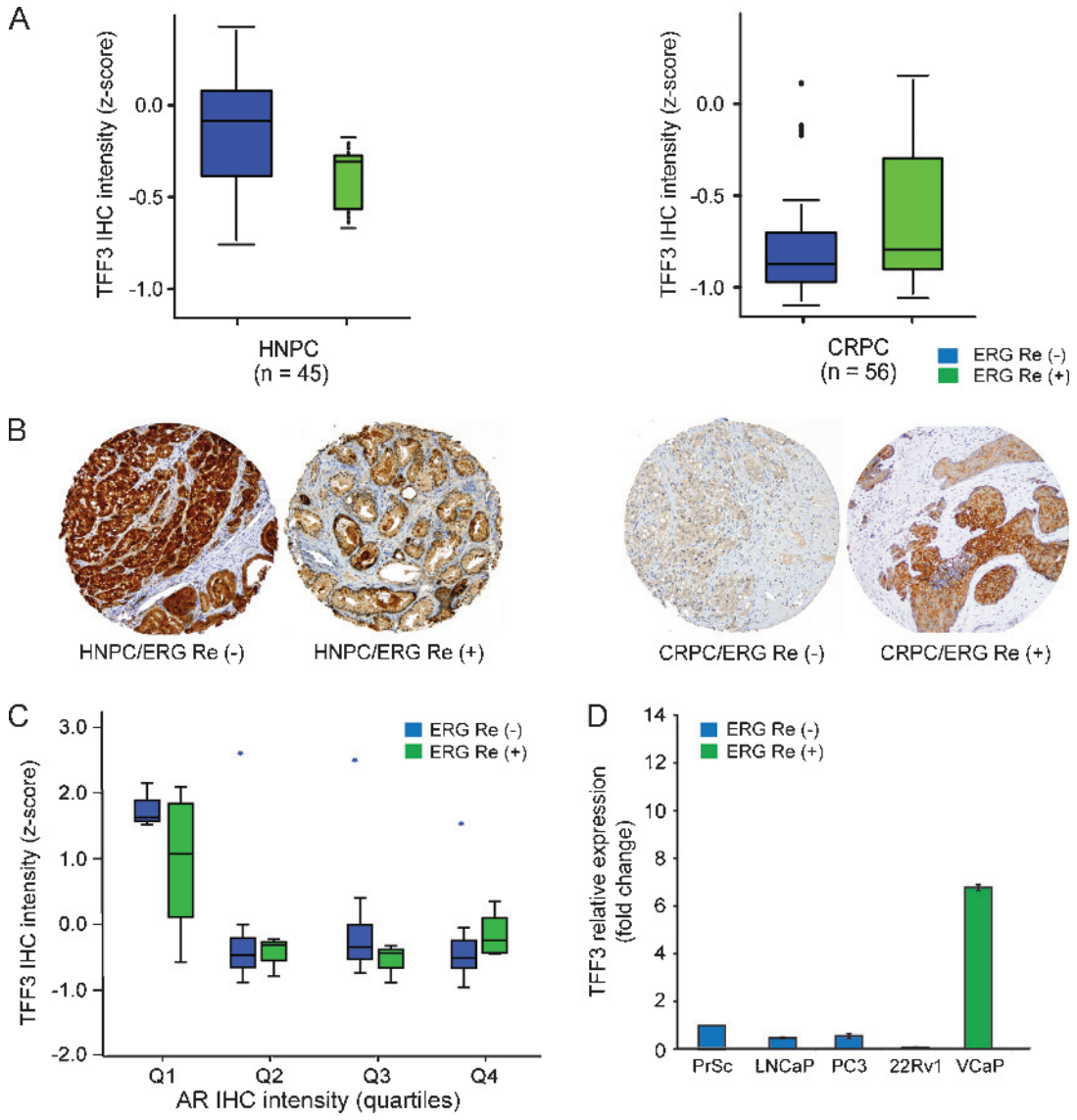

Figure 2. *ERG*-rearranged HNPC and CRPC show inverse patterns of differential regulation of TFF3 protein. (A) TFF3 protein expression as determined by immunohistochemistry (IHC) shows an inverse pattern in HNPC (left panel) and CRPC (right panel) with respect to *ERG* rearrangement, consistent with the mRNA behavior. (B) Representative images of TFF3 protein expression in HNPC and CRPC samples in the absence or presence of *ERG* rearrangement. (C) The expression of TFF3 in CRPC is higher in the cases with low AR expression as determined by IHC. Data are represented based on AR levels quartiles (Q1-Q4). Dot, outliers (1.5-3 interquartile range); asterisk, extreme values (>3 interquartile range). (D) Quantification of TFF3 mRNA in prostate cell lines with (green) or without (blue) *ERG* rearrangement.

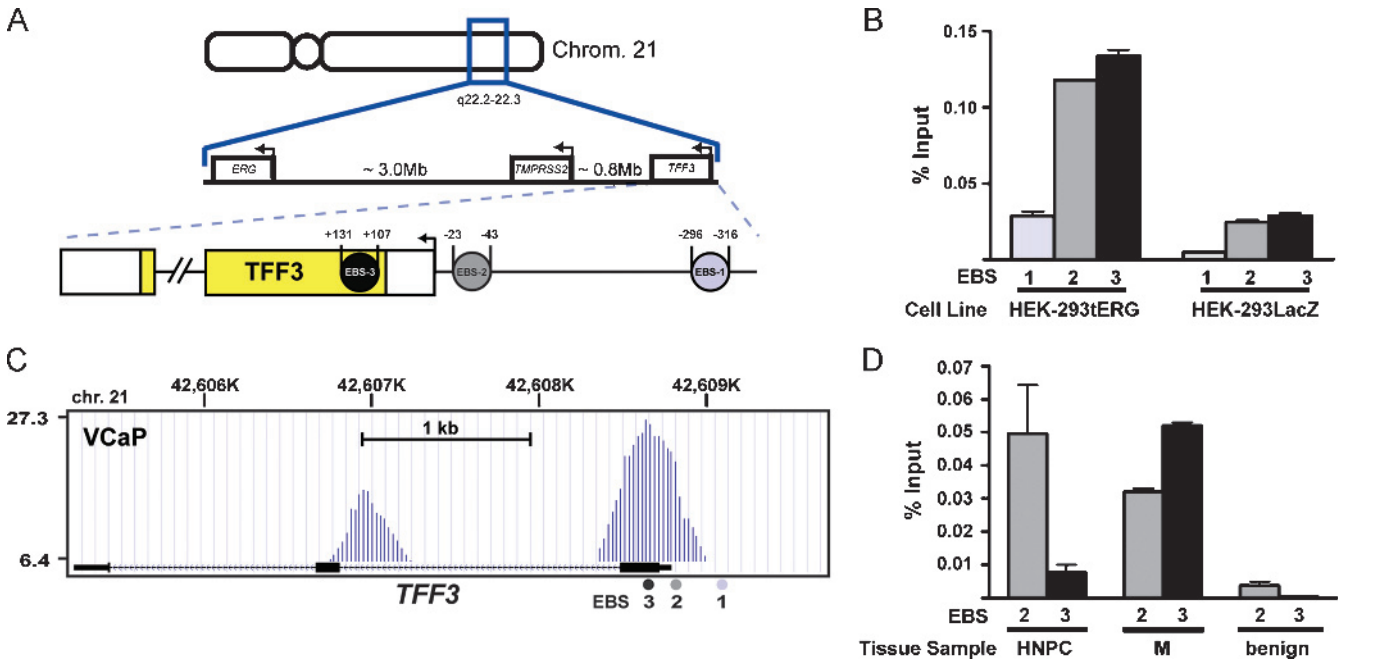

Figure 3. ERG directly binds to the *TFF3* promoter *in vitro* and in human prostate cancer samples and regulates its expression. (A) Schematic representation of *TFF3*, *TMPRSS2*, and *ERG* positions on chromosome 21. The transcription start site for each gene is indicated with an arrow. Below shows a scheme (not to scale) of the *TFF3* gene (coding sequence in yellow) and the position of the three ETS binding sites (EBS1-3, circles) relative to the transcription start site (arrow). (B) Bar graphs showing the amount of enriched DNA (relative to input chromatin preparation) for each EBS in the indicated cell lines after ChIP using either anti-FLAG (right) or anti-ERG (left) antibodies. (C) Chip-seq data obtained from VCaP cells of the region of chromosome 21 that covers the *TFF3* gene and ~800 bp of upstream sequence. The lines indicate the relative number of significant reads that mapped to the indicated region. The chromosome region is indicated above. The normalized peak reads, 27.3, and the 6.4 cutoff reads (above which the peak signal is defined as significant) are shown on the left. The circles below indicate the positions of the EBS1-3. (D) Bar graphs showing the amount of enriched DNA for EBS2 and three in the indicated tissue sample after ChIP using anti-ERG (left) antibodies.

GFP (control) or tERG (RWPE1-tERG; Figure W3). Treatment of RWPE1-GFP cells with R1881, a synthetic androgen, led to an up-regulation of TFF3 compared with cells grown in androgen-free medium, which could be abrogated with bicalutamide. RWPE1-tERG cells had much lower level of *TFF3* expression than control cells likely because of the inhibitory effect of ERG. Treatment with R1881 led to a further down-regulation of *TFF3* compared with cells grown in an androgen-free medium. Treatment with R1881 together with the AR antagonist bicalutamide restored *TFF3* levels (Figure 4*B*).

TFF3 Promotes Invasion in CRPC Cells

To explore the potential function of TFF3 in CRPC, we overexpressed *ERG* fusion transcripts with or without *TFF3* coexpression in another androgen-insensitive cell line DU145 that is negative for *ERG* rearrangement (Figure 4*C*). DU145 represents another CRPC cell line but differs from VCaP in that these cells are androgen-insensitive. Overexpression of ERG lead to an increase of TFF3 mRNA levels consistent with the results presented in Figure 4*B*. Consistent with previous studies of other cell types [27,28], overexpression of ERG fusion transcripts alone resulted in a heightened level of cell invasion (Figure 4*D*) but had no effect on cell proliferation (data not shown). Interestingly, increasing the levels of *TFF3* in ERG-negative DU145 cells resulted in similar levels of cell invasion. Combining *ERG* and *TFF3* overexpression led to a further enhanced level of cell invasion. In the hormone-insensitive VCaP cell line, which harbors the *TMRPSS2-ERG* gene fusion, *TFF3* mRNA is significantly decreased after treatment with *TFF3* siRNA compared with scrambled control (Figure 4*E*). This decrease in *TFF3* mRNA after siRNA treatment is associated with a significant decrease in cell invasion (Figure 4*F*).

Discussion

The current study is the first, to our knowledge, to systematically characterize the molecular changes associated with *ERG* rearrangement in CRPC. A major finding is the demonstration of distinct molecular signaling in *ERG*-rearranged prostate cancer in the context of the androgen environment. It has been shown that ADT leads to an incomplete inhibition of androgen-regulated genes [21] while maintaining AR expression at the transcript level [29]. On the basis of our profiling data, we also found that *AR* mRNA is higher in CRPC *versus* HNPC ($P = 1.4e-05$, Wilcoxon rank sum test, data not shown). Our data now extend this observation suggesting that this transcriptome deregulation differentially occurs based on *ERG* rearrangement status. A previous study by Hermans et al. [23] suggested that the ETS fusion transcript is no longer produced after loss of the AR, where they found four xenografts negative for AR did not express ERG, whereas other members of the ETS family (namely, ETV4 and FLI1) were overexpressed through an androgen-independent mechanism. Saramaki et al. [30] also reported two AR-negative xenografts generated from *ERG*-rearranged small cell carcinomas of the prostate that did not express ERG. However, other hormonally sensitive xenografts and clinical tumor specimens they tested indeed showed elevated ERG expression levels correlated with *ERG* rearrangement. In line with this finding,

most studies including the initial index case [3] and the current study have found that gene fusion expression continues into the castration-resistant state perhaps due to the incomplete inhibition. Works by Nelson et al. [31] and Lin et al. [32] characterizing androgen response genes place the 5′ fusion partners of *ERG* (i.e., *TMPRSS2*, *SCL45A3*, and *NDRG1*) among the most androgen-responsive genes. One important implication is that ETS-rearranged prostate cancer may respond more dramatically than rearrangement-negative tumors due to a higher

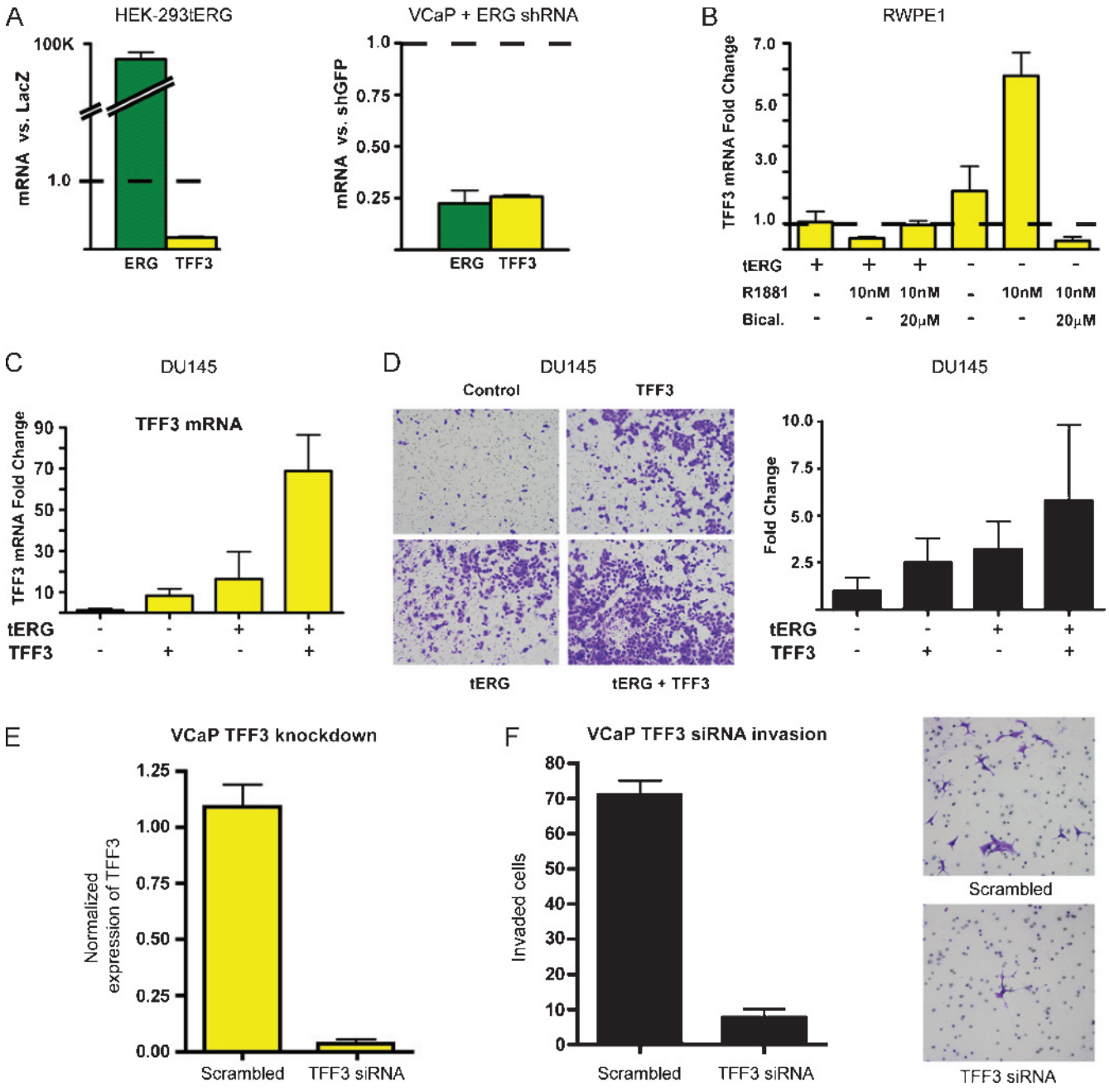

Figure 4. (A) Relative levels of ERG (green) and TFF3 (yellow) mRNA in HEK-293tERG cells (normalized to HEK-293LacZ cells; left) or in VCaP cells (normalized to VCaP cells treated with shRNA against GFP; right) after *ERG* shRNA expression. The dotted line represents the level corresponding to the expression in control cells. (B) Relative average levels of *TFF3* mRNA in RWPE1-tERG (tERG +) or RWPE1-GFP (tERG −) cells grown in androgen-free (charcoal-stripped, phenol red–free) medium and treated 24 hours with vehicle or 10 nM of the synthetic androgen R1881 in the presence or absence of the AR antagonist 20 μM bicalutamide (Bical.). Expression levels were normalized to RWPE1-tERG cells grown in androgen-free vehicle-treated conditions (dotted line). (C) Relative levels *TFF3* (yellow) mRNA in DU145 cells stably overexpressing tERG (ERG-positive) or GFP (ERG-negative) with or without transient overexpression of TFF3 (transient). The data was normalized to DU145-GFP control cells. (D) Representative images of cells (left) and bar chart depicting the corresponding relative cell counts (right) obtained from cells described in (C) at the completion of the invasion assay. (E) Normalized mRNA expression of *TFF3* in VCaP cells is effectively reduced after exposure to *TFF3* siRNA compared with control (scrambled siRNA molecule). (F) Reduction of *TFF3* using siRNA significantly decreases VCaP cell invasion compared with control. Representative areas of invasion are demonstrated (right side). For all experiments, error bars demonstrate SD across triplicate experiments.

addiction to hormonal regulation at any level. Reports have shown that *ERG* rearrangement does not predict response to traditional hormonal therapy [33,34]. However, this is not the case when a more complete AR block is achieved as suggested by the results of a recent clinical trial testing the ability of the potent CYP17 inhibitor, abiraterone, to decrease PSA levels greater than 90% in men with CRPC [35]. In this phase 2 clinical trial, men with *ERG*-rearranged CRPC, as determined by the evaluation of circulating tumor cells, responded significantly more frequently than men without detectable *ERG* rearrangements. Taken together, these observations support the hypothesis that ETS rearrangements continue to play an important role in CRPC. Interestingly, unlike *AR* gene amplification and *PTEN* loss, which were heterogeneously altered within the circulating tumor cells isolated from individual cases, all tumor cells in *ERG*-rearranged cases showed the gene fusion, supporting the notion that the *ERG* rearrangement occurs early in tumorigenesis and persists in CRPC [14,36].

Understanding how ERG directly regulates downstream targets is an important step in developing targets to key signaling pathways. Several recent studies have explored this using tumor cell lines [27,28,37]. We report for the first time that ERG directly regulates *TFF3*, a gene located within 0.8 MB from *TMPRSS2* on 21q22.3. Although TFF3 has previously been reported as a potential prostate cancer biomarker overexpressed in a subset of primary and metastatic cases [25,26,38], the underlying regulatory mechanism was entirely unknown. TFF3 has also been shown to regulate cell migration, invasion, and angiogenesis in other cancer types (see Supplemental Discussion). Here we show that in HNPC TFF3 is significantly downregulated in *ERG*-rearranged cases. After exposure to ADT (CRPC cases), we observe a strong reversal of this relationship; TFF3 is expressed at a higher level in *ERG*-rearranged cases. We further demonstrate that in an *ERG*-rearranged cell line (VCaP), *TFF3* is expressed at a higher level than in the *ERG*-nonrearranged cell line (LNCaP). Moreover, we show that ERG binds to *TFF3* gene promoter region ETS binding sites in *ERG*-rearranged HNPC and CRPC tumor samples. Interestingly, the work of Hollenhorst and Graves has shown that the different members of the ETS family of transcription factors bind to the same ETS binding elements in gene promoters [39,40]. Therefore, because LNCaP cells harbor a cryptic *ETV1* rearrangement [41], it is possible that ETV1 could bind to the same ETS binding sites located in the *TFF3* promoter that was found to be bound by ERG in VCaP cells, but this has yet to be tested. Whereas ERG exerts an inhibitory effect on *TFF3* expression in HEK-293tERG cells, we found that ERG stimulates *TFF3* expression in VCaP cell lines, as well as in human CRPC samples, supporting a potential cross-talk between *ERG* rearrangement and AR signaling in CRPC. However, we cannot rule out the influence of other key factors that, by orchestrated interactions, may also contribute to these differences observed between the different stages of tumor progression. Because increased level of TFF3 might enhance ERG-induced cell invasion as shown by our *in vitro* study, the different levels of TFF3 expressed in CRPC under the influence of ERG and AR signaling might thus alter the spread of tumor cells. The mechanisms underlying the observation that ERG directly regulates TFF3 in HNPC and CRPC and how TFF3 affects the pathogenesis of CRPC warrant investigation.

Therefore, based on ChIP-PCR, ChIP-seq, knock-in, and knockdown shRNA *in vitro* data, we find that ERG regulates TFF3 expression by directly binding to ETS binding sites in the *TFF3* promoter and that ERG exerts dual activities on TFF3 regulation depending on the level of androgen signaling. Our results further support the idea first suggested by Yu et al. [6] that ERG is hijacking AR-mediated prostate cell maintenance as we show that, in the context of low AR activity, ERG is inducing *TFF3* expression, which, in turn, results in an increased invasive potential of CRPC cells. In summary, *TFF3* is a direct molecular target of ERG and is regulated in a context related to hormonal status and increases the invasive potential of CRPC cells. These observations suggest that understanding ETS rearrangement-related molecular alterations will be important in the interpretation of ongoing ADT clinical trials and in the development of novel therapeutics for CRPC.

Acknowledgments

The authors thank John Day and Jack Groskopf from Gen-Probe (San Diego, CA) for performing the TMPRSS2-ERG transcript expression assays on the archival samples.

References

[1] Petrylak DP, Tangen CM, Hussain MH, Lara PN Jr, Jones JA, Taplin ME, Burch PA, Berry D, Moinpour C, Kohli M, et al. (2004). Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer. *N Engl J Med* **351**, 1513–1520.

[2] Tannock IF, de Wit R, Berry WR, Horti J, Pluzanska A, Chi KN, Oudard S, Theodore C, James ND, Turesson I, et al. (2004). Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N Engl J Med* **351**, 1502–1512.

[3] 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.

[4] Pflueger D, Rickman DS, Sboner A, Perner S, LaFargue CJ, Svensson MA, Moss BJ, Kitabayashi N, Pan Y, de la Taille A, et al. (2009). N-*myc* downstream regulated gene 1 (*NDRG1*) is fused to *ERG* in prostate cancer. *Neoplasia* **11**, 804–811.

[5] Wang Q, Li W, Zhang Y, Yuan X, Xu K, Yu J, Chen Z, Beroukhim R, Wang H, Lupien M, et al. (2009). Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. *Cell* **138**, 245–256.

[6] Yu J, Yu J, Mani R-H, Cao Q, Brenner CJ, Cao X, Wang GX, Wu L, Li J, Hu M, et al. (2010). An integrated network of androgen receptor, polycomb, and *TMPRSS2-ERG* gene fusions in prostate cancer progression. *Cancer Cell* **17**(5), 443–454.

[7] Cai C, Wang H, Xu Y, Chen S, and Balk SP (2009). Reactivation of androgen receptor–regulated *TMPRSS2:ERG* gene expression in castration-resistant prostate cancer. *Cancer Res* **69**(15), 6027–6032.

[8] Setlur SR, Mertz KD, Hoshida Y, Demichelis F, Lupien M, Perner S, Sboner A, Pawitan Y, Andren O, Johnson LA, et al. (2008). Estrogen-dependent signaling in a molecularly distinct subclass of aggressive prostate cancer. *J Natl Cancer Inst* **100**, 815–825.

[9] Stanbrough M, Bubley GJ, Ross K, Golub TR, Rubin MA, Penning TM, Febbo PG, and Balk SP (2006). Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. *Cancer Res* **66**, 2815–2825.

[10] Holzbeierlein J, Lal P, LaTulippe E, Smith A, Satagopan J, Zhang L, Ryan C, Smith S, Scher H, Scardino P, et al. (2004). Gene expression analysis of human prostate carcinoma during hormonal therapy identifies androgen-responsive genes and mechanisms of therapy resistance. *Am J Pathol* **164**, 217–227.

[11] Shah RB, Mehra R, Chinnaiyan AM, Shen R, Ghosh D, Zhou M, Macvicar GR, Varambally S, Harwood J, Bismar TA, et al. (2004). Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program. *Cancer Res* **64**, 9209–9216.

[12] Tamura K, Furihata M, Tsunoda T, Ashida S, Takata R, Obara W, Yoshioka H, Daigo Y, Nasu Y, Kumon H, et al. (2007). Molecular features of hormone-refractory prostate cancer cells by genome-wide gene expression profiles. *Cancer Res* **67**, 5117–5125.

[13] Rubin MA, Putzi M, Mucci N, Smith DC, Wojno K, Korenchuk S, and Pienta KJ (2000). Rapid ("warm") autopsy study for procurement of metastatic prostate cancer. *Clin Cancer Res* **6**, 1038–1045.

[14] 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.

[15] Rubin MA, Allory Y, Molinie V, Leroy X, Faucon H, Vacherot F, Huang W, Kuten A, Salomon L, Rebillard X, et al. (2005). Effects of long-term finasteride treatment on prostate cancer morphology and clinical outcome. *Urology* **66**, 930–934.

[16] Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, Gupta S, Moore J, Wrobel MJ, Lerner J, et al. (2008). Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* **359**, 1995–2004.

[17] Quandt K, Frech K, Karas H, Wingender E, and Werner T (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* **23**, 4878–4884.

[18] Rickman DS, Pflueger D, Moss B, VanDoren VE, Chen CX, de la Taille A, Kuefer R, Tewari AK, Setlur SR, Demichelis F, et al. (2009). SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. *Cancer Res* **69**, 2734–2738.

[19] Sandoval J, Rodriguez JL, Tur G, Serviddio G, Pereda J, Boukaba A, Sastre J, Torres L, Franco L, and Lopez-Rodas G (2004). RNAPol-ChIP: a novel application of chromatin immunoprecipitation to the analysis of real-time gene transcription. *Nucleic Acids Res* **32**, e88.

[20] Yu J, Cao Q, Mehra R, Laxman B, Yu J, Tomlins SA, Creighton CJ, Dhanasekaran SM, Shen R, Chen G, et al. (2007). Integrative genomics analysis reveals silencing of beta-adrenergic signaling by polycomb in prostate cancer. *Cancer Cell* **12**, 419–431.

[21] Mostaghel EA, Page ST, Lin DW, Fazli L, Coleman IM, True LD, Knudsen B, Hess DL, Nelson CC, Matsumoto AM, et al. (2007). Intraprostatic androgens and androgen-regulated gene expression persist after testosterone suppression: therapeutic implications for castration-resistant prostate cancer. *Cancer Res* **67**, 5033–5041.

[22] Han B, Mehra R, Suleman K, Tomlins SA, Wang L, Singhal N, Linetzky KA, Palanisamy N, Zhou M, Chinnaiyan AM, et al. (2009). Characterization of *ETS* gene aberrations in select histologic variants of prostate carcinoma. *Mod Pathol* **22**(9), 1176–1185.

[23] 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.

[24] 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**(17), 8347–8351.

[25] Faith DA, Isaacs WB, Morgan JD, Fedor HL, Hicks JL, Mangold LA, Walsh PC, Partin AW, Platz EA, Luo J, et al. (2004). Trefoil factor 3 overexpression in prostatic carcinoma: prognostic importance using tissue microarrays. *Prostate* **61**, 215–227.

[26] Garraway IP, Seligson D, Said J, Horvath S, and Reiter RE (2004). Trefoil factor 3 is overexpressed in human prostate cancer. *Prostate* **61**, 209–214.

[27] Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, Cao Q, Prensner JR, Rubin MA, Shah RB, et al. (2008). Role of the *TMPRSS2-ERG* gene fusion in prostate cancer. *Neoplasia* **10**, 177–188.

[28] Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A, Alimonti A, Nardella C, Varmeh S, Scardino PT, et al. (2009). Aberrant *ERG* expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat Genet* **41**, 619–624.

[29] Chen CD, Welsbie DS, Tran C, Baek SH, Chen R, Vessella R, Rosenfeld MG, and Sawyers CL (2004). Molecular determinants of resistance to antiandrogen therapy. *Nat Med* **10**, 33–39.

[30] Saramaki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, and Visakorpi T (2008). *TMPRSS2:ERG* fusion identifies a subgroup of prostate cancers with a favorable prognosis. *Clin Cancer Res* **14**, 3395–3400.

[31] Nelson PS, Clegg N, Arnold H, Ferguson C, Bonham M, White J, Hood L, and Lin B (2002). The program of androgen-responsive genes in neoplastic prostate epithelium. *Proc Natl Acad Sci USA* **99**, 11890–11895.

[32] 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.

[33] Leinonen KA, Tolonen TT, Bracken H, Stenman UH, Tammela TL, Saramaki OR, and Visakorpi T (2010). Association of *SPINK1* expression and *TMPRSS2:ERG* fusion with prognosis in endocrine-treated prostate cancer. *Clin Cancer Res* **16**, 2845–2851.

[34] Boormans JL, Hermans KG, Made AC, van Leenders GJ, Wildhagen MF, Collette L, Schroder FH, Trapman J, and Verhagen PC (2009). Expression of the androgen-regulated fusion gene TMPRSS2-ERG does not predict response to endocrine treatment in hormone-naive, node-positive prostate cancer. *Eur Urol*.

[35] Attard G, Swennenhuis JF, Olmos D, Reid AH, Vickers E, A'Hern R, Levink R, Coumans F, Moreira J, Riisnaes R, et al. (2009). Characterization of *ERG*, *AR* and *PTEN* gene status in circulating tumor cells from patients with castration-resistant prostate cancer. *Cancer Res* **69**, 2912–2918.

[36] Mosquera JM, Perner S, Genega EM, Sanda M, Hofer MD, Mertz KD, Paris PL, Simko J, Bismar TA, Ayala G, et al. (2008). Characterization of *TMPRSS2-ERG* fusion high-grade prostatic intraepithelial neoplasia and potential clinical implications. *Clin Cancer Res* **14**, 3380–3385.

[37] Klezovitch O, Risk M, Coleman I, Lucas JM, Null M, True LD, Nelson PS, and Vasioukhin V (2008). A causal role for *ERG* in neoplastic transformation of prostate epithelium. *Proc Natl Acad Sci USA* **105**, 2105–2110.

[38] Vestergaard EM, Borre M, Poulsen SS, Nexo E, and Torring N (2006). Plasma levels of trefoil factors are increased in patients with advanced prostate cancer. *Clin Cancer Res* **12**, 807–812.

[39] Hollenhorst PC, Chandler KJ, Poulsen RL, Johnson WE, Speck NA, and Graves BJ (2009). DNA specificity determinants associate with distinct transcription factor functions. *PLoS Genet* **5**, e1000778.

[40] Hollenhorst PC, Shah AA, Hopkins C, and Graves BJ (2007). Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the *ETS* gene family. *Genes Dev* **21**, 1882–1894.

[41] Tomlins SA, Bjartell A, Chinnaiyan AM, Jenster G, Nam RK, Rubin MA, and Schalken JA (2009). *ETS* gene fusions in prostate cancer: from discovery to daily clinical practice. *Eur Urol* **56**, 275–286.

Supplemental Materials and Methods

Computational and Statistical Data Analysis of Expression Profiling Data

Normalization and data filtering/quality control. The expression values generated from the DASL 6K platform are corrected for background noise using the Beadstudio software (Illumina, San Diego, CA). A detection *P* value is calculated for each probe by comparing its intensity to the rank of the negative controls. Lower *P* values pertain to reliable expression data. We considered as missing all the gene expression values that did not meet a *P* value threshold of .05. We then excluded from further analysis all the samples with more than 55% of missing values (six samples excluded). The same criterion was applied for the selection of genes (1603 genes excluded). Data were then *logarithm* (base 2)–transformed. To correct for an array effect, each sample for each DAP was median-centered. To correct for a DAP effect, all samples belonging to a DAP was median-centered. When multiple samples were profiled from the same patient (see cohort description), the similarity of the expression profile was evaluated pairwise using the concordance correlation coefficient [1]. If an intrapatient sample pair had a concordance correlation coefficient greater than 0.7, the sample with more reliably expressed genes was retained for further analysis. This resulted in a total of 54 samples from patients and 4541 genes.

Differential expression analysis. Wilcoxon rank test with a two-sided alternative was used to evaluate the differential expression of genes with respect to *ERG* rearrangement. FDR was calculated using the procedure of Benjamini and Hochberg [2], and genes with FDR less than 20% are reported.

Combined analysis of CRPC and HNPC. Adding the localized samples to the CRPC samples during the gene expression profiling experiment is critical for the CRPC and HNPC joint analysis as it allows us to estimate the confounding effect. The CRPC and control samples expression data were q-splined together with the median array as the target. Similar q-spline normalization was performed for the Swedish HNPC cohort. We considered the following model for the analysis of variance for each gene: $\log_2(S_{gi}) = \mu_g + \delta_g E_{gi} + \alpha_g F_{gi} + \beta_g H_{gi} + \gamma_g H_{gi} F_{gi}$ (*E*, *F*, and *H* correspond to the effects of experiment, *ERG* rearrangement status, and ADT, respectively. *HF* represents the interaction between *ERG* rearrangement status and ADT. The subscripts *g* and *i* correspond to gene and sample, respectively). To account for deviation from normality of the combined data, we performed permutation tests to assess the significance of the variance components. Permutations were performed on the raw data labels. Although permuting residuals can yield more power, the gain in power is marginal [3].

TMPRSS2-ERG *Fusion mRNA Assay*

TMPRSS2-ERG gene fusion mRNA and *PSA* mRNA were quantified using magnetic target capture (purification of target mRNA by hybridization to magnetic particles through target-specific oligonucleotides), transcription-mediated amplification (amplification of target RNA sequences), and a hybridization protection assay (specific detection of amplification products by use of target-specific acridinium ester (AE)–labeled probes) as described by Groskopf et al. [4]. Three *TMPRSS2-ERG* splice variants were targeted: *TMPRSS2-ERGa*, b, and c [5], also known as types III, I, and VI, respectively [6]. Amplification primers for *TMPRSS2-ERGa* mRNA were located in *TMPRSS2* exon 1 and *ERG* exon 4 yielding an amplification product of 117 nucleotides. The *TMPRSS2-ERGb* assay used amplification primers in *TMPRSS2* exon 1 and *ERG* exon 2 yielding an amplification product of 109 nucleotides. The *TMPRSS2-ERGc* assay used primers in *TMPRSS2* exon 2 and ERG exon 4, yielding an amplification product of 84 nucleotides. The AE probes for each *TMPRSS2-ERG* splice variant spanned the junction between the two exons. Primers for *PSA* targeted exons 2 and 3 of the *PSA* mRNA, and the AE probe spanned the exon 2/3 junction. Calibrators and controls consisted of *TMPRSS2-ERG* or *PSA in vitro* transcripts (IVTs) in detergent solution. The *TMPRSS2-ERGa*, b, and c IVTs were prepared from plasmids provided by A. Chinnaiyan of the University of Michigan [5]. IVT copy levels were determined by A_{260}. Assays were performed at Gen-Probe Incorporated (San Diego, CA) using DTS 400 Systems; the assay protocol uses the reagent addition volumes and incubation times and temperatures specified in the APTIMA COMBO2 package insert and as described in Groskopf et al. [4]. *PSA* mRNA was used to normalize for the total amount of prostate-specific mRNA in each sample. Samples with less than 500 copies of *PSA* mRNA per 10 ng of total RNA were considered invalid because of insufficient mRNA for analysis.

Supplemental Discussion

Among the top three differentially regulated genes in *ERG*-rearranged CRPC (Figure 1*F*) is *PFKFB3*, a unique member of the enzyme family 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase that regulates the cellular concentration of fructose-2,6-biphosphate (F2,6BP). The steady-state concentration of F2,6BP controls the overall rate of glycolysis by allosterically activating phosphofructokinases, enzymes that mediate a key rate-limiting step of glycolysis. In contrast to other family members, *PFKFB3* has the highest kinase-phosphatase activity ratio, and up-regulation significantly promotes glycolysis [7]. The fact that *PFKFB3* was identified as one of the most significantly upregulated genes in *ERG*-rearranged CRPC indicates that increased glycolysis might be a feature of this subset of CRPC. It has long been established that some tumors have elevated glycolytic rates at even normal oxygen tension and are thus more reliant on this pathway for cell growth, a phenomenon referred to as the Warburg effect. Enhanced glycolysis is also an important mechanism regulated by hypoxia-inducible factor 1 signaling in fast-growing tumors to adapt to hypoxia. The presence of a significantly higher level of PFKFB3 in concert with many enzymes mediating glycolysis, as described in this current study, suggests that *ERG*-rearranged CRPCs are reliant on this pathway and may have a selective growth advantage in an oxygen poor state, but further work is needed to validate this hypothesis.

The function of TFF3 in prostate cancer has not been well studied. Trefoil factors constitute a family of extracellular peptides (TFF1, 2, and 3) initially identified in the gastrointestinal tract that contain distinctive cystine-rich "trefoil" domains and are resistant to proteolytic degradation. In colon cancer cell lines, TFF3 is known to activate the nuclear factor κB or STAT3 signaling pathways to regulate cell migration, invasion and angiogenesis [8,9]. Previous studies on the transcriptional regulation of TFF3 have revealed that hypoxia-associated hypoxia-inducible factor 1 activation can induce TFF3 expression by directly binding to a responsive element in the TFF3 promoter [10]. In addition, phosphatidylinositol 3-kinase and Akt activation can promote the expression of TFF3 [11]. TFF3 was also identified as a steroid-responsive gene in breast cancer [12], where its expression was related not only to ER but also to AR [13]. On the basis of *in vitro* studies, we

show that restoring the level of TFF3 in cells perturbs ERG-induced cell invasion.

One can also speculate that ETS rearrangement status alone or in the context of other genetic alterations, such as *PTEN* mutations [14,15], begin to define molecularly distinct subclasses of prostate cancer. We previously suggested that *ERG*-rearranged HNPC can activate estrogenic signaling for maintenance of tumor growth [16], which was recently supported by a study from Jhavar et al. [17]. The finding that *TMPRSS2* harbors an ER binding site and is regulated by ER-α suggests a gene fusion specific growth advantage in certain settings. We imagined that ER signaling might play a more important role in the setting of ADT. The current study would suggest that context-specific differences based on gene fusion status could provide important insight into therapeutic targeting, analogous to defining breast cancers based on ER, PR, and HER-2/*neu* receptor status.

References

[1] Lin LI (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**, 255–268.

[2] Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc B* **57**, 289–300.

[3] Anderson MJ and Braak CJFT (2003). Permutation tests for multi-factorial analysis of variance. *J Stat Comput Simul* **23**, 85–113.

[4] Groskopf J, Aubin SM, Deras IL, Blase A, Bodrug S, Clark C, Brentano S, Mathis J, Pham J, Meyer T, et al. (2006). APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. *Clin Chem* **52**, 1089–1095.

[5] 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.

[6] 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.

[7] Chesney J (2006). 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase and tumor cell glycolysis. *Curr Opin Clin Nutr Metab Care* **9**, 535–539.

[8] Zhu YQ and Tan XD (2005). TFF3 modulates NF-κB and a novel negative regulatory molecule of NF-κB in intestinal epithelial cells via a mechanism distinct from TNF-α. *Am J Physiol Cell Physiol* **289**, C1085–C1093.

[9] Rivat C, Rodrigues S, Bruyneel E, Pietu G, Robert A, Redeuilh G, Bracke M, Gespach C, and Attoub S (2005). Implication of STAT3 signaling in human colonic cancer cells during intestinal trefoil factor 3 (TFF3)– and vascular endothelial growth factor–mediated cellular invasion and tumor growth. *Cancer Res* **65**, 195–202.

[10] Furuta GT, Turner JR, Taylor CT, Hershberg RM, Comerford K, Narravula S, Podolsky DK, and Colgan SP (2001). Hypoxia-inducible factor 1–dependent induction of intestinal trefoil factor protects barrier function during hypoxia. *J Exp Med* **193**, 1027–1034.

[11] Durual S, Blanchard C, Estienne M, Jacquier MF, Cuber JC, Perrot V, Laboisse C, and Cuber JC (2005). Expression of human TFF3 in relation to growth of HT-29 cell subpopulations: involvement of PI3-K but not STAT6. *Differentiation* **73**, 36–44.

[12] Doane AS, Danso M, Lal P, Donaton M, Zhang L, Hudis C, and Gerald WL (2006). An estrogen receptor–negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen. *Oncogene* **25**, 3994–4008.

[13] Naderi A and Hughes-Davies L (2008). A functionally significant cross-talk between androgen receptor and ErbB2 pathways in estrogen receptor negative breast cancer. *Neoplasia* **10**, 542–548.

[14] Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A, Alimonti A, Nardella C, Varmeh S, Scardino PT, et al. (2009). Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat Genet* **41**, 619–624.

[15] King JC, Xu J, Wongvipat J, Hieronymus H, Carver BS, Leung DH, Taylor BS, Sander C, Cardiff RD, Couto SS, et al. (2009). Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. *Nat Genet* **41**, 524–526.

[16] Setlur SR, Mertz KD, Hoshida Y, Demichelis F, Lupien M, Perner S, Sboner A, Pawitan Y, Andren O, Johnson LA, et al. (2008). Estrogen-dependent signaling in a molecularly distinct subclass of aggressive prostate cancer. *J Natl Cancer Inst* **100**, 815–825.

[17] Jhavar S, Brewer D, Edwards S, Kote-Jarai Z, Attard G, Clark J, Flohr P, Christmas T, Thompson A, Parker M, et al. (2009). Integration of ERG gene mapping and gene-expression profiling identifies distinct categories of human prostate cancer. *BJU Int* **103**, 1256–1269.

Table W1. Clinical Characteristics of 35 Patients with CRPC in the Current Study.

Age (years)	
Range	56-94
Median	73
Gleason score at initial diagnosis	
≤6	4 (11%)
7-8	13 (37%)
9-10	18 (52%)
Primary treatment	
Radical prostatectomy	2 (5%)
Radiation + ADT	5 (14%)
ADT	28 (80%)
ADT protocols	
Orchiectomy	6 (17%)
LHRH	4 (11%)
LHRH + antiandrogens	25 (72%)
Time to development of resistance to ADT (months)	
Range	8-92
Median	32
PSA level at TURP (ng/ml)	
Range*	0.1-4537
Median	34

TURP indicates transurethral resection of prostate.
*Patients without biochemical recurrence (PSA level elevation) were clinically diagnosed as CRPC based on clinical evidences of disease progression (i.e. newly developed metastatic foci).

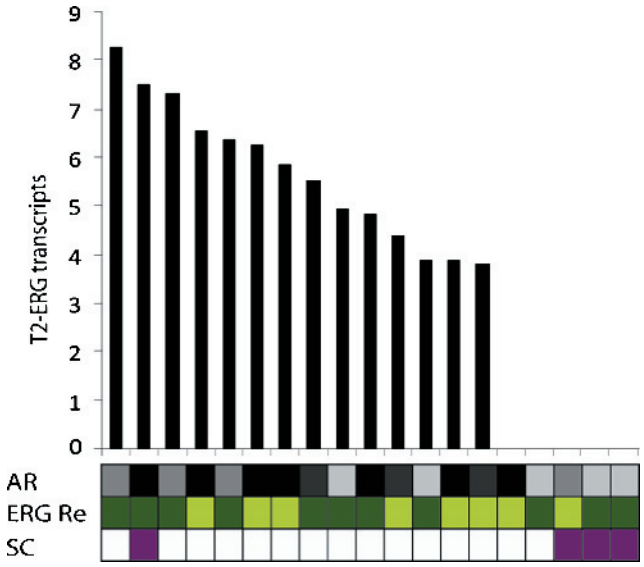

Figure W1. Expression of ERG fusion transcripts in CRPC. Bar graph of quantified transcripts. The normalized level of *ERG* fusion transcript with higher expression in each sample is plotted after natural logarithm transformation. The level of AR evaluated by immunohistochemistry, the types of *ERG* rearrangement, and the presence of small cell carcinoma (SC) features are annotated below. AR: gray shades indicate quartiles of expression (lightest gray corresponds to the lowest quartile). *ERG* rearrangement: light green, rearrangement through translocation; dark green, rearrangement through deletion. SC: purple indicates SC carcinoma.

Table W2. Primer Sets Targeting Each of the Three ETS Binding Sites (EBS) Identified in the *TFF3* Promoter and Control Primer Set Targeting *ARHGEF*.

Name	Gene	Location to TSS	Predicted ETS Site	Sequence
TFF3 EBS-1 f	*TFF3*	−375	−296 to −316	AAACTGTCTGGGAGCTTGAC
TFF3 EBS-1 r	*TFF3*	−252		CTCAGGACTCGCTTCATGG
TFF3 EBS-2 f	*TFF3*	−61	−23 to −43	CACATCCGCTCCCCAGTAG
TFF3 EBS-2 r	*TFF3*	78		GGCTGCTCCTGTTCTCTCC
TFF3 EBS-3 f	*TFF3*	47	+107 to +131	TTTCAATCAATTCATTTCATCCAC
TFF3 EBS-3 r	*TFF3*	179		GAGACAGGCAGCTCTGAG
AREGH f	*AREGH*	NA	NA	TCTCTGCTCCCTCACTCTCAA
AREGH r	*AREGH*	NA	NA	TGTGCCTCTTCCATCGTTCT

NA indicates not applicable.

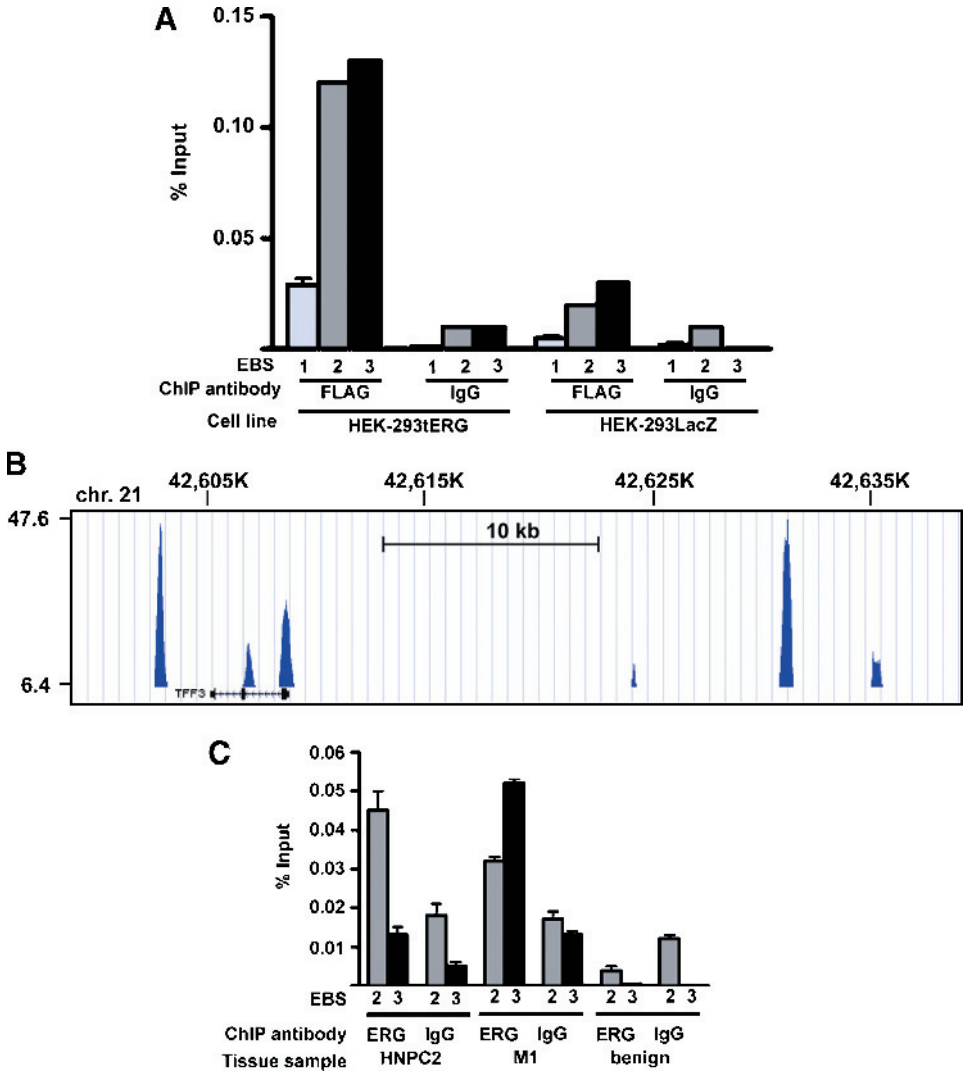

Figure W2. ERG directly regulates the expression of TFF3. (A) Bar graphs showing the amount of enriched DNA (relative to input chromatin preparation) for each EBS in HEK-293tERG and HEK-293LacZ after ChIP using mouse anti-FLAG or mouse IgG antibodies. (B) Chip-seq data obtained from VCaP cells of the region of chromosome 21 that covers the *TFF3* gene and ~30 kb of upstream sequence and ~6.5 kb of sequence downstream from the last exon. The lines indicate the relative number of significant reads that mapped to the indicated region. The chromosome region is indicated above. The normalized peak reading (47.6) and the cutoff reading (6.4; above which is defined as significant) are shown on the left. (C) Bar graphs showing the amount of enriched DNA (relative to input chromatin preparation) that correspond to the indicated EBS in a HNPC, metastatic (M) prostate cancer, or benign tissue samples after ChIP using rabbit anti-ERG or rabbit anti-ERG or rabbit IgG antibodies. Error bars, SD across triplicate experiments.

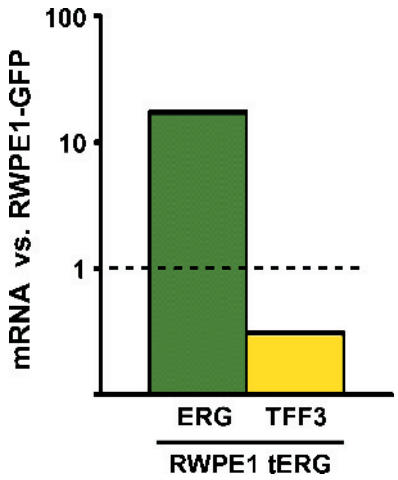

Figure W3. Stable overexpression of tERG in RWPE-1 cells (RWPE1-tERG). This graph shows the relative levels of tERG or TFF3 mRNA in stable RWPE1-tERG cells compared with RWPE1-GFP control cells. These cells correspond to the cells treated with androgen shown in Figure 4*B*.