ERG induces androgen receptor-mediated regulation of SOX9 in prostate cancer

Changmeng Cai,1 Hongyun Wang,1 Housheng Hansen He,2,3 Sen Chen,1 Lingfeng He,1 Fen Ma,1 Lorelei Mucci,4 Qianben Wang,2 Christopher Fiore,2 Adam G. Sowalsky,1 Massimo Loda,2 X. Shirley Liu,3 Myles Brown,2 Steven P. Balk,1 and Xin Yuan1

1Hematology-Oncology Division, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA. 2Department of Medical Oncology and 3Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA. 4Departments of Epidemiology and Nutrition, Dana Farber Cancer Institute and Harvard School of Public Health, Boston, Massachusetts, USA.

Fusion of the androgen receptor-regulated (AR-regulated) TMPRSS2 gene with ERG in prostate cancer (PCa) causes androgen-stimulated overexpression of ERG, an ETS transcription factor, but critical downstream effectors of ERG-mediating PCa development remain to be established. Expression of the SOX9 transcription factor correlated with TMPRSS2:ERG fusion in 3 independent PCa cohorts, and ERG-dependent expression of SOX9 was confirmed by RNAi in the fusion-positive VCaP cell line. SOX9 has been shown to mediate ductal morphogenesis in fetal prostate and maintain stem/progenitor cell pools in multiple adult tissues, and has also been linked to PCa and other cancers. SOX9 overexpression resulted in neoplasia in murine prostate and stimulated tumor invasion, similarly to ERG. Moreover, SOX9 depletion in VCaP cells markedly impaired invasion and growth in vitro and in vivo, establishing SOX9 as a critical downstream effector of ERG. Finally, we found that ERG regulated SOX9 indirectly by opening a cryptic AR-regulated enhancer in the SOX9 gene. Together, these results demonstrate that ERG redirects AR to a set of genes including SOX9 that are not normally androgen stimulated, and identify SOX9 as a critical downstream effector of ERG in TMPRSS2:ERG fusion–positive PCa.

Introduction

The androgen receptor (AR) plays a central role in prostate cancer (PCa) development, and its transcriptional functions are partially or fully restored in the tumors that relapse after androgen deprivation therapy (castration-resistant prostate cancer, CRPC) (1). The role of AR in PCa was further strengthened by the discovery of recurrent genomic rearrangements that result in AR-driven overexpression of ETS family transcription factor proto-oncogenes, and in particular the v-ets erythroblastosis virus E26 oncogene homolog, ERG. In approximately half of primary PCa cases, a gene fusion between the 5′ untranslated region of the androgen-regulated TMPRSS2 gene and an exon in the ERG gene results in androgen-regulated high-level expression of a transcriptionally active, N-terminal–truncated ERG protein (amino acids 1–44 being deleted in the most common fusion) (2, 3). This fusion is an early event, as it is found in precursor prostatic intraepithelial neoplasia (PIN) lesions located adjacent to TMPRSS2:ERG fusion–positive cancers (4). The fusion gene is also highly expressed in CRPC, indicating that overexpressed ERG contributes to PCa development and progression (5).

The importance of ERG is further supported by studies in the TMPRSS2:ERG fusion–positive VCaP cell line, where RNAi-mediated downregulation of ERG impairs cell growth and invasion (6, 7). Moreover, transgenic overexpression of ERG in mouse prostate causes increased proliferation, and in combination with the loss of one PTEN allele, results in PIN or invasive PCa (8–11). Several ERG-regulated genes in TMPRSS2:ERG fusion–positive PCa cases have been identified, with EZH2 being one such candidate ERG target gene (6, 7, 11–15). Additionally, a recent study found that there was marked overlap between ERG and AR binding sites, and that ERG repressed both AR expression and AR activity on these sites, suggesting that downregulation of AR activity on subsets of genes may be a mechanism of ERG action (12). However, the downstream effectors and functions of ERG that drive PCa remain to be firmly established.

SOX9 belongs to the SOX (SRY-related HMG box) family of transcription factors and regulates many developmental processes (16, 17). SOX9 mutations cause the disease campomelic dysplasia, which is characterized by extreme cartilage and bone malformation, frequent XY sex reversal, and multiple defects in other organs, consistent with an important role in tissue development (18, 19). In the adult, SOX9 contributes to maintenance of stem/progenitor cells in tissues including intestine, liver, pancreas, and hair follicle (20–24), and dysregulated SOX9 expression has been implicated in the pathogenesis of several cancers (25–31). Similar to AR, SOX9 is required for prostate development, and accumulating evidence indicates that it contributes to the development of PCa (32–39).

SOX9 in developing prostate is expressed by epithelial cells invading into urogenital sinus mesenchyme, and loss of SOX9 causes a profound defect in prostate ductal morphogenesis (33, 35, 36, 39). In adult human prostate, SOX9 is normally expressed primarily by the basal cells, which surround the strongly AR-positive luminal epithelium and express only low levels of AR (34). In contrast, SOX9 and AR are coexpressed in PCa cells, with an increased fraction of SOX9-positive cells being found in CRPC.
and correlating with more aggressive behavior (32, 34). Moreover, a 17q24.3 SNP associated with PCa has been functionally mapped to an enhancer of the SOX9 gene, and allele-specific recruitment of transcription factors to this enhancer can increase SOX9 and may mediate the increased PCa risk (38). Silencing endogenous SOX9 expression in PCa xenografts by shRNA causes reduced tumor growth, while SOX9 overexpression enhances PCa xenograft growth and invasion (35). Finally, similarly to ERG, transgenic overexpression of SOX9 in mouse prostate epithelium causes hyperplasia, and on a Pten-deficient background, can enhance PIN development (32). We report here that SOX9 is a critical down-stream effector of ERG in TMPRSS2:ERG fusion–positive PCa, and that ERG stimulates SOX9 expression by redirecting AR to a cryptic AR-regulated enhancer in the SOX9 gene.

**Results**

SOX9 expression is correlated with TMPRSS2:ERG fusion in primary and metastatic CRPC. We reported previously that expression of the AR-regulated TMPRSS2:ERG fusion gene is restored in CRPC, consistent with the reactivation of AR transcriptional activity and a continued important role for ERG in CRPC (5). Therefore, to identify genes that may be regulated by ERG in vivo, we

Figure 1

SOX9 expression is increased in TMPRSS2:ERG-positive PCa. (A) Heat maps (green, lower expression; red, higher expression) of Affymetrix microarray data showing mRNA expression for genes with increased levels in TMPRSS2:ERG fusion–positive versus TMPRSS2:ERG fusion–negative tumor subsets of both primary PCa and CRPC bone metastases (fold change of mean expression >1.5 and a 2-tailed Student's t test P value < 0.05). (B) Quartile plot (mean and each quartile) of ERG and SOX9 expression. The differences between the fusion-positive and fusion-negative samples were *P < 0.05 or **P < 0.01 (2-tailed Student's t test). (C) MSKCC dataset showing that SOX9 expression was significantly higher in ERG-positive tumors (based on the outlier expression of ERG, 47/103) versus ERG-negative tumors. Heat maps for ERG and SOX9 expression in individual tumors are shown in Supplemental Figure 1A. (D) SOX9 expression index based on SOX9 immunostaining intensity and fraction of positive cells (mean SOX9 addscore) are plotted for TMPRSS2:ERG fusion–positive versus –negative radical prostatectomy samples from the Physicians' Health and Health Professionals Follow-up Studies.
examined our previously reported PCA gene expression dataset (40) for genes that were increased in TMPRSS2:ERG-positive primary PCa and metastatic CRPC (Figure 1A). SOX9 was among the genes that were most significantly correlated with TMPRSS2:ERG fusion in both the primary and metastatic CRPC tumors (Figure 1, A and B). To confirm this correlation between SOX9 and TMPRSS2:ERG fusion, we next examined a recently reported large primary PCa dataset from MSKCC (41), and again found that SOX9 expression was significantly correlated with TMPRSS2:ERG fusion (Figure 1C; heat map shown in Supplemental Figure 1B). This analysis again showed a significant correlation between TMPRSS2:ERG fusion and SOX9 expression (Figure 1D). Together, these results indicated that SOX9 may be directly or indirectly regulated by ERG in TMPRSS2:ERG-positive PCa. Previous reports have established critical roles for SOX9 during prostate development and in stem/progenitor cell maintenance of multiple tissues, as well as in the pathogenesis of several cancers including PCa (20–35, 37, 38). Therefore, we focused on SOX9 as a possible effector of ERG in the pathogenesis of several cancers including PCa.

SOX9 is a critical downstream effector of ERG in TMPRSS2:ERG fusion–positive PCa. Therefore, we focused on whether ERG was regulating SOX9 expression. We generated stable lentiviral shERG-infected VCaP cells and found that basal- and DHT-stimulated SOX9 mRNA and protein expression were markedly decreased (Figure 2, G and H). We obtained similar results using another ERG shRNA (shERG-2, Supplemental Figure 2B) and by transient transfection with an independent ERG siRNA (Supplemental Figure 6B). Although additional mechanisms may further modulate the expression of SOX9 in fusion-negative or -positive tumors (34), these data establish that SOX9 is increased by ERG in TMPRSS2:ERG fusion–positive PCa.

To determine whether SOX9 protein expression also correlated with TMPRSS2:ERG fusion, we carried out SOX9 immunohistochemical (IHC) assessment on a primary PCa tissue microarray from the Physicians’ Health and Health Professionals Follow-up Studies, which had been assessed previously for TMPRSS2:ERG fusion by FISH and IHC (42, 43). Each tumor was scored based on staining intensity and fraction of cells that were positive (SOX9 addscore; representative SOX9-positive and -negative staining is shown in Supplemental Figure 1B). This analysis again showed a significant correlation between TMPRSS2:ERG fusion and SOX9 expression (Figure 1D). Together, these results indicated that SOX9 may be directly or indirectly regulated by ERG in TMPRSS2:ERG-positive PCa. Previous reports have established critical roles for SOX9 during prostate development and in stem/progenitor cell maintenance of multiple tissues, as well as in the pathogenesis of several cancers including PCa (20–35, 37, 38).

Sox9 was then highly expressed again in the relapsed castration-resistant xenograft. Consistent with the IHC assessment, this transgene in these mice is controlled by the tetracycline transactivator (tTA), which is controlled by the androgen-regulated PSA and ERG proteins in these cells, and this was blocked by an AR antagonist, bicalutamide (Bic) (Figure 2D, and data not shown). In contrast, SOX9 protein was decreased by DHT in LNCaP cells (Figure 2E).

We next examined VCaP xenografts to determine whether there was a correlation in vivo between SOX9 and ERG. We showed previously that AR and ERG were highly expressed in these xenografts prior to castration, declined markedly immediately after castration, and were again highly expressed in the relapsed castration-resistant xenografts (5). As shown in Figure 2F, SOX9 protein showed a similar marked decrease after castration and was then highly expressed again in the relapsed castration-resistant xenograft. Consistent with the IHC assessment, ERG and SOX9 mRNA levels declined immediately after castration and were restored in the relapsed castration-resistant xenografts (Supplemental Figure 2A). Finally, we used RNAi to directly examine whether ERG was regulating SOX9 expression. We generated stable lentiviral shERG-infected VCaP cells and found that basal- and DHT-stimulated SOX9 mRNA and protein expression were markedly decreased (Figure 2, G and H). We obtained similar results using another ERG shRNA (shERG-2, Supplemental Figure 2B) and by transient transfection with an independent ERG siRNA (Supplemental Figure 6B). Although additional mechanisms may further modulate the expression of SOX9 in fusion-negative or -positive tumors (34), these data establish that SOX9 is increased by ERG in TMPRSS2:ERG fusion–positive PCa.

The expression of SOX9 in prostate epithelial overexpression of Flag epitope–tagged SOX9 resulted in PIN lesions in a fraction of mice (4 of 9) analyzed at 5 to 8 months of age. Although the SOX9 transgene was expressed in all prostate lobes (Supplemental Figure 3A), the PIN lesions were primarily in the ventral lobe (Figure 3, SOX9). We also observed PIN in a fraction of mice (5 of 9) that were Pten+/−, although these lesions were more prominent in the dorsolateral and anterior lobes, and spared the ventral lobe (Figure 3, Pten+/−, and Supplemental Figure 3B). In contrast, all compound Pten+/−;SOX9 mice (19 of 19) developed PIN lesions, which were also much larger than those in Pten+/− mice and occurred in all lobes (Figure 3, Pten+/−;SOX9). As expected, these lesions had high levels of Flag-SOX9, Ki67, and phospho-AKT (Supplemental Figure 3C), and high AR expression as well (Supplemental Figure 3D). Lesions in 2 Pten+/−;SOX9 mice had a loss of basal cells (indicated by a lack of high molecular weight cytokeratin staining) and marked host inflammatory reactions, suggesting local invasion (Figure 3, bottom panels).

The SOX9 transgene in these mice is controlled by the tetracycline operon. Its expression in prostate was stimulated by the tetracycline transactivator (tTA), which is controlled by the androgen-regulated MMTV promoter. Since tTA activity can be repressed by the addition of tetracycline or doxycycline (tet-off), we also addressed whether established PIN lesions would respond to tTA repression by doxycycline. A series of approximately 8-month-old Pten+/−;SOX9 mice were treated with doxycycline (added to their food and drinking water) for 2 weeks prior to sacrifice, and PIN lesions in their prostates were compared with untreated controls. There was a decrease in the number of PIN lesions and a decrease in the proportion of larger PIN lesions (> 0.5 mm2) from 22% to 9% in the doxycycline-treated mice (Supplemental Figure 3, E and F), suggesting a continued role for SOX9 in established lesions.
Figure 2

SOX9 expression is androgen and ERG regulated in TMPRSS2:ERG-positive VCaP PCa cells. (A) VCaP; LNCaP; LNCaP-CSS3 (derived from LNCaP cells adapted to hormone-depleted conditions); C4-2; LNCaP-ABL; LAPC4-PR; LAPC4-CR (derived from androgen-dependent or castration-resistant LAPC4 xenografts); and CWR22-Rv1 cells were treated with or without DHT for 24 hours followed by qRT-PCR. (B) VCaP and LNCaP cells were treated with 0, 0.01, 0.1, 1, or 10 nM DHT for 24 hours and SOX9 mRNA was measured by qRT-PCR. (C) LNCaP and VCaP cells were stimulated with or without DHT for 4 hours, and binding of activated phospho-RNA polymerase II (phospho-Ser5 on CTD) to the SOX9 transcriptional start site (TSS) was assessed by ChIP. (D) VCaP cells were treated with 0–10 nM (0, 0.1, 1, or 10) DHT and with vehicle (DMSO) or 10 μM Bic for 24 hours and immunoblotted. (E) LNCaP cells were treated with 0, 0.01, 0.1, 1, or 10 nM DHT for 24 hours, followed by immunoblotting. (F) Expression of SOX9, ERG, or AR was examined by IHC in VCaP xenografts prior to castration (Androgen-dependent), at 4 days after castration, or at relapse (Castration-resistant). Original magnification, ×400; Scale bars: 20 μm. (G and H) VCaP cells stably infected with nontarget control (NTC) or ERG shRNA-1 were treated with or without DHT for 24 hours, and SOX9 or ERG mRNA was then measured by qRT-PCR, or SOX9, ERG, and β-tubulin protein were assessed by immunoblotting. Data in bar graphs represent means ± SD of at least 3 biological repeats. Significant differences from DHT-negative controls are indicated (*).
We next used Matrigel invasion assays to determine whether SOX9 contributes to ERG-mediated invasion. Consistent with our previous results in LNCaP xenograft cells overexpressing SOX9, which showed increased invasion in vivo (35), SOX9 induction in LNCaP cells with doxycycline-inducible SOX9 (LN-toSOX9, protein induction shown in Supplemental Figure 4A) strongly stimulated basal Matrigel invasion (Figure 4A). DHT-stimulated invasion was not significantly enhanced, which could reflect SOX9 activation of DHT-regulated genes in mediating invasion or alternative mechanisms. To determine whether SOX9 could mediate ERG-dependent invasion, we generated VCaP cells with doxycycline-inducible SOX9 (VCaP-toSOX9, protein induction shown in Supplemental Figure 4B) and used siRNA to downregulate endogenous ERG (Figure 4B). As expected, Matrigel invasion was decreased by ERG siRNA, which also decreased endogenous SOX9 (Figure 4, B and C). Importantly, invasion was restored by the addition of doxycycline to stimulate exogenous SOX9 expression. Interestingly, while SOX9 induction could restore invasion after ERG knockdown, it did not stimulate invasion in androgen-depleted VCaP cells, indicating that additional DHT-regulated genes in these cells are required to support invasion (Supplemental Figure 4C).

To determine whether VCaP invasion was dependent on endogenous SOX9, we next used a SOX9 shRNA lentivirus (shSOX9-1) to downregulate SOX9 (Figure 4D), which resulted in a marked decrease in DHT-stimulated Matrigel invasion (Figure 4E). This result was confirmed using an independent SOX9 shRNA lentivirus (shSOX9-2) (Supplemental Figure 4, D and E). Finally, a previous study had found that tissue-type plasminogen activator (PLAT), which may contribute to invasion, was the major gene downregulated in response to ERG RNAi in VCaP cells (6). Therefore, we addressed whether PLAT expression was SOX9 regulated. As shown in Figure 4F, SOX9 siRNA markedly decreased the expression of PLAT mRNA in VCaP cells. Moreover, in VCaP cells with doxycycline-regulated SOX9, we found that SOX9 induction with doxycycline could substantially restore PLAT mRNA in cells treated with ERG siRNA (Figure 4G). Together, these results indicate that SOX9 is an effector of ERG-mediated PCa invasion.

To determine whether SOX9 contributes to tumor growth, we next assessed the effects of SOX9 RNAi on growth in vitro and in vivo. In LNCaP cells expressing low levels of SOX9, growth in vitro was not impaired by SOX9 siRNA (Supplemental Figure 4F). In contrast, infection of VCaP cells with the shSOX9-1 lentivirus, which markedly downregulates SOX9 protein (Figure 4D), substantially decreased in vitro growth (Figure 4H). Indeed, we have not been able to obtain adequate numbers of these cells for xenograft studies. Infection with the shSOX9-2 lentivirus did not as effectively deplete SOX9 (Supplemental Figure 4D), and in vitro growth was not substantially decreased (not shown). Nonetheless, SOX9 downregulation by shSOX-2 markedly impaired the ability to develop xenografts (Figure 4I). Moreover, in the two shSOX9-2 xenografts that developed, the rate of proliferation (assessed by Ki67 immunostaining) was decreased compared with the control tumor on the opposite flank of the same mouse (Figure 4J).

While these results show that VCaP growth in vitro can be decreased by SOX9 shRNA, we and others have not found marked effects of ERG RNAi on VCaP growth in vitro (5). This may reflect a need to more dramatically deplete SOX9 in order to affect growth versus invasion in vitro, as VCaP growth in vitro was impaired only by the more efficient shSOX9-1. Consistent with this hypothesis, doxycycline induction of SOX9 can partially rescue VCaP-toSOX9 cells from the growth-suppressive effects of androgen-deprivation VCaP-toSOX9 (which reduces ERG and markedly reduces endogenous SOX9) (Supplemental Figure 4G). As expected, DHT is more effective, as it clearly stimulates the
research article

A

LN-toSOX9

DMSO

Dox

EtOH

DHT

B

VCP-toSOX9

ERG

FLAG

SOX9

β-actin

siNTC  siERG

siERG + dox

C

siNTC

siERG

siERG + dox

F

VCP

PLAT

SOX9

Relative mRNA expression

G

VCP-toSOX9

Relative mRNA expression

H

VCP-toSOX9

Relative cell density

I

% Tumor volume

J

VCP-shNTC

VCP-shSOX9-2

VCP-shNTC

VCP-shSOX9-2

K67: 53 ± 6%

K67: 36 ± 4%

K67: 48 ± 7%

K67: 23 ± 6%
expression of multiple other genes in addition to ERG and SOX9.

Taken together, these in vitro and in vivo results further support the conclusion that SOX9 is a critical mediator of ERG effects in TMPRSS2:ERG fusion–positive PCa.

**AR directly stimulates SOX9 expression in TMPRSS2:ERG fusion–positive cells.** Since ERG expression in TMPRSS2:ERG fusion–positive PCa cells is stimulated by androgen, we initially presumed that the DHT-stimulated increase in SOX9 in VCaP cells was secondary to the increased expression of ERG (Figure 2). However, the rapid induction of SOX9 mRNA and SOX9 protein in response to DHT suggested that AR may be directly inducing SOX9 (Figure 5, A and B). To determine whether the DHT-stimulated increase in SOX9 mRNA was mediated by increased ERG protein, we used cycloheximide (CHX) to block new protein synthesis. Importantly, the addition of CHX did not prevent the DHT-stimulated increase in SOX9 mRNA (Figure 5C). As a control for the inhibition of protein synthesis, we confirmed that CHX caused a rapid decrease in SOX9 protein in DHT-stimulated and control cells (Supplemental Figure 5A). Moreover, we found that overexpression of full-length ERG or ERG(del1-44), which is expressed in VCaP cells, did not increase SOX9 in AR-negative PC3 cells (Supplemental Figure 5B). These results indicate that AR may be directly stimulating SOX9 expression.

Significantly, examination of available AR ChIP massively parallel DNA sequencing (ChIP-seq) data (12, 44, 45) for AR binding sites within 50 kb of the SOX9 gene transcriptional start site revealed a major peak located at approximately 35 kb 3' of the SOX9 gene in VCaP cells that was not present in LNCaP cells (Figure 5D and Supplemental Figure 5C). Conversely, the major SOX9 gene–associated AR binding peak in LNCaP cells (S1 site, located at ~40 kb 5' of the SOX9 start site) was a very minor peak in VCaP cells. Moreover, the S1 peak in LNCaP cells was much weaker than the S2 peak in VCaP cells (Figure 5D). Interestingly, a portion of the S1 site is highly conserved across 17 vertebrate species (Supplemental Figure 5C), although we do not know if this is related to AR function at this site. Using AR ChIP and quantitative PCR (qPCR) with S2 site–specific primers, we confirmed the marked DHT-stimulated recruitment of AR to this site in VCaP cells (Figure 5E). This could be blocked by the AR antagonist Bic, which, as reported previously, can more weakly stimulate AR binding in the absence of androgen. In contrast, there was minimal DHT-stimulated AR binding to the S2a and S2b sites flanking the S2 site, consistent with the major sharp AR peak observed by ChIP-seq (Supplemental Figure 5, C and D).

AR binding to the S1 and S2 sites in VCaP versus LNCaP cells is compared in detail in Figure 5, F and G, respectively. Weak AR binding to the S1 site, which could be modestly increased by DHT, was observed in both VCaP and LNCaP cells. Anti-FOXA1 ChIP showed that FOXA1, which has been previously identified as a pioneer transcription factor that is required to initially open loci for subsequent AR binding (46), was bound to the S1 sites in both cells. Low levels of GATA2, which also associates with AR on many AR-regulated genes (46), were present in the LNCaP cells, but not in the VCaP cells. Most importantly, DHT did not stimulate recruitment of the p300 coactivator to the S1 site in either cell, and modestly decreased p300 binding to the S1 site in LNCaP cells. These findings indicate that the weak AR binding to the S1 site does not stimulate SOX9 expression, but may contribute to DHT-mediated SOX9 repression in LNCaP cells.

In contrast to the S1 site, AR binding to the S2 site was dramatically increased by DHT in VCaP cells, but not in LNCaP cells (Figure 5, F and G, middle panels). Since the AR in LNCaP cells has a point mutation in the ligand binding domain (T877A), we also examined a stable LNCaP cell line expressing a 3xFlag-tagged wild-type AR and found that it was similarly recruited to the S1 site by DHT, but not to the S2 site (data not shown). Consistent with the S2 site functioning as an AR-regulated enhancer in VCaP cells and interacting with the promoter, DHT stimulated an association between the S2 site and phospho-RNA polymerase II (Figure 5F, middle panel). FOXA1 binding and DHT-stimulated GATA2 binding to the S2 site were also observed in the VCaP cells, but not in the LNCaP cells. Moreover, DHT stimulated the recruitment of p300 to the S2 site in VCaP cells. The DHT-stimulated binding of AR, the association with phospho-RNA polymerase II, and the recruitment of p300 at the S2 site in VCaP cells were comparable to those observed for the major AR-regulated enhancer in the PSA gene in both VCaP and LNCaP cells (Figure 5, F and G, bottom panels). Finally, using ChIP we found that levels of monomethylated and dimethylated H3K4 (H3K4me1 and H3K4me2), which are associated with active enhancers, were higher at the S2 site in VCaP cells compared with the S1 site in VCaP cells, and were higher than the levels at the S1 and S2 sites in LNCaP cells (Supplemental Figure 5E). Together, these findings indicate that the S2 site is an AR-regulated enhancer for the SOX9 gene in VCaP cells, but not in LNCaP cells.

To assess AR binding to the S2 site in vivo, we performed AR ChIP on VCaP xenografts prior to castration and at 4 days after castration. Comparable levels of AR were bound to the S2 site and to the PSA gene enhancer prior to castration, and AR binding to both sites was markedly decreased after castration (Supplemental Figure 5F). To evaluate the contribution of the S2 site to SOX9 expression in CRPC, we also examined a cell line (VCS2) generated from a castration-resistant VCaP xenograft (45). Basal levels (in steroid-depleted medium) of both PSA and
SOX9 mRNA were higher in the VCS2 cells compared with the parental VCaP cells, and could be further elevated by DHT and suppressed by Bic (Supplemental Figure 5H). Basal p300 and GATA2 binding on the PSA enhancer in VCS2 cells was very high, consistent with the substantial basal PSA expression in these cells (Supplemental Figure 5H, bottom panels). Similarly to VCaP cells, DHT in the VCS2 cells markedly induced AR, GATA2, and p300 binding to the S2 site, but not to the S1 site (Supplemental Figure 5H). These findings further support the conclusion that the S2 site is an AR-regulated enhancer for the SOX9 gene in TMPRSS2:ERG fusion–positive PCa.

ERG activates the AR-regulated S2 enhancer in TMPRSS2:ERG fusion–positive PCa. We next examined available ERG ChIP-seq data in VCaP cells to determine whether ERG may be regulating AR binding to the S2 site (12). Significantly, we found a broad peak of ERG binding over approximately 2 kb that overlapped...
the S2 site (Figure 6A). Using anti-ERG ChIP and qPCR with a series of primers targeting S2 and its neighboring sites (S3–S5) that cover this approximately 2-kb span, we confirmed ERG binding to this region (Figure 6B). We next used ERG siRNA to determine whether ERG downregulation altered AR binding to the S2 site in VCaP cells. Importantly, although we could greatly reduce cellular ERG protein levels, ChIP studies showed that ERG binding over the S2–S5 region was reduced but not eliminated (data not shown). Nonetheless, ERG siRNA reduced DHT-stimulated AR binding to the S2 site (Supplemental Figure 6A), and similarly reduced DHT-stimulated expression of SOX9 (Supplemental Figure 6B).

As a complimentary approach, we determined whether the ERG(del1-44) expressed from the TMPRSS2:ERG fusion in VCaP cells is sufficient to confer DHT stimulation of SOX9 expression in fusion-negative LNCaP cells. In 2 independent LNCaP cell lines expressing a doxycycline-inducible ERG protein (LN-to-ERG-1 and -2), we found that DHT could stimulate SOX9 mRNA expression (Figure 6C) and protein expression (Supplemental Figure 6C), which could be blocked by Bic (Supplemental Figure 6D). Importantly, this exogenous ERG in LNCaP cells was bound to the S2–S5 region identified in the VCaP cells (Figure 6D). Moreover, the S2 site in the ERG-expressing LNCaP cells had increased basal FOXA1 binding prior to DHT stimulation, consistent with ERG binding to this site and making it accessible to transcription factors (Figure 6E). Finally, DHT strongly stimulated recruitment of AR and p300 to the previously unavailable S2 site in the ERG-expressing LNCaP cells (Figure 6E). It should be noted that LNCaP cells have a distinct gene fusion that results in androgen-stimulated overexpression of another ETS factor, ETV1, indicating that not all ETS factors can stimulate SOX9 expression. Together, these findings demonstrate that ERG increases SOX9 expression in TMPRSS2:ERG fusion–positive PCa cells by opening an AR-regulated enhancer in the SOX9 gene.

Identification of additional ERG-dependent AR-regulated genes in TMPRSS2:ERG fusion–positive PCa. Since these data indicated that ERG was functioning as a pioneer factor for the AR-regulated S2 site SOX9 gene enhancer, we hypothesized that ERG expression may open cryptic AR-regulated enhancers associated with additional genes and thereby further reprogram AR transcription in TMPRSS2:ERG fusion–positive PCa. As outlined below, to test this hypothesis we first sought to identify additional genes (exemplified by SOX9) that were androgen stimulated and contained unique AR binding sites in VCaP cells compared with LNCaP cells. We then assessed whether the unique AR binding sites in these genes were associated with ERG binding sites. For this analysis, we used expression microarray data from androgen-stimulated VCaP cells and LNCaP cells (44, 45). In addition, we used
Figure 7
Identification of additional direct AR-regulated genes associated with TMPRSS2:ERG fusion. (A) Identification of genes that are DHT induced with AR binding sites in VCaP cells, but not in LNCaP cells, and have associated ERG binding sites. (B) Quartile plots from Affymetrix arrays showing expression of 9 genes from this subset in fusion-positive versus fusion-negative primary PCa (PR) or CRPC bone metastases (B-Met). Significant differences are indicated. (C) Five genes were further selected for qRT-PCR analysis in DHT-treated (24 hours) VCaP cells stably expressing an NTC shRNA (V-shNTC), ERG shRNA-1 (V-shERG), or LNCaP cells. Androgen stimulation of the control FKBP5 gene, which is strongly androgen regulated in LNCaP and VCaP cells, was not significantly altered in the V-shERG versus control V-shNTC VCaP cells. (D and E) AR, FOXA1, or ERG binding on the CSGALNACT1-ARE (Supplemental Figure 7C) in VCaP, LN-toERG-2 (with dox), or LNCaP-Ctrl (with dox) cells was measured by ChiP-qPCR.
published ChIP microarray technology (ChIP-chip) and ChIP-seq data identifying AR binding sites in DHT-treated LNCaP and VCaP cells, and ERG binding sites in VCaP cells (7A). Among these 471 genes, the VCaP AR ChIP-seq data showed that 274 (58%) had AR binding sites within 50 kb of their transcription start sites (versus a background of 37% of all genes in these cells having an AR binding site within 50 kb), consistent with them being directly regulated by AR (Supplemental Figure 7A). Of these 274 genes that appeared to be AR regulated in VCaP cells, analysis of the microarray expression data in LNCaP cells showed that 75 genes had less than a 50% expression level in LNCaP cells versus VCaP cells, and were not DHT induced in LNCaP cells (less than 1.5-fold induction). Comparison of the AR ChIP-chip and ChIP-seq data in LNCaP cells versus VCaP cells then showed that 66 of these 75 genes had AR binding sites in VCaP cells that were not present in LNCaP cells.

Significantly, using the VCaP ERG ChIP-seq data, we then found that ERG binding sites were located within 20 kb of the AR site in 56 of these 66 genes (Figure 7A). Forty of these 66 genes (61%) have ERG binding sites within 1 kb of the AR binding sites. By comparison, among the total set of 274 AR-regulated genes in VCaP cells, only 97 genes (35%) have ERG binding sites within 1 kb of the AR binding sites, indicating that these VCaP-specific AR binding sites are further enriched for ERG binding. The VCaP-specific AR-regulated genes with closely associated ERG and AR binding sites (within 4 kb) are listed in Table 1.

Gene ontology analysis of these genes revealed only a weak overall association with secretion (Supplemental Figure 7B). To identify genes in this group that may be clinically significant, we assessed their expression in TMPRSS2:ERG fusion–positive versus TMPRSS2:ERG fusion–negative primary PCa and CRPC bone metastases (40). In addition to SOX9 (Figure 1), 9 of these genes were increased in TMPRSS2:ERG fusion–positive PCa primary tumors (PR) or bone marrow metastases (B-Met) (Figure 7B). AR and ERG binding sites for these 9 genes are shown in Supplemental Figure 7C. However, these genes were also not strongly associated with specific pathways (Supplemental Figure 7B, right panel), and none were significantly increased in both the primary and metastatic CRPC tumors.

Table 1

<table>
<thead>
<tr>
<th>List of ERG-dependent AR-activated genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGR2</td>
</tr>
<tr>
<td>ANXA2</td>
</tr>
<tr>
<td>ASRG1</td>
</tr>
<tr>
<td>ATP7B</td>
</tr>
<tr>
<td>CDK2AP2</td>
</tr>
<tr>
<td>CMTM8</td>
</tr>
<tr>
<td>CORO2A</td>
</tr>
<tr>
<td>CSGALNACT1</td>
</tr>
<tr>
<td>DPP8</td>
</tr>
<tr>
<td>EPDR1</td>
</tr>
<tr>
<td>FAM13C</td>
</tr>
<tr>
<td>FKB1P</td>
</tr>
<tr>
<td>FXYD3</td>
</tr>
</tbody>
</table>

VCaP-specific AR-regulated genes with ERG binding site located within 4 kb of the AR binding site.

Three of these 9 genes (CSGALNACT1, ATP7B, and SLC26A2) were increased in the fusion-positive primary tumors from the MSKCC dataset (Supplemental Figure 7D) (41).

To confirm the ERG-dependent androgen regulation of these genes, we selected 5 of the 9 and assessed the effects of ERG shRNA on their DHT-stimulated expression in VCaP cells. In all cases, qRT-PCR showed low expression in LNCaP cells and confirmed that basal and DHT-stimulated expression in VCaP cells was markedly diminished by ERG shRNA, while expression of a gene that is strongly androgen stimulated in VCaP and LNCaP cells (FKBP5) was not markedly altered (Figure 7C). Since CSGALNACT1 expression was strongly correlated with TMPRSS2:ERG fusion in both clinical datasets, we further assessed DHT-stimulated recruitment of transcription factors to its AR binding site in VCaP and LNCaP cells. As observed for SOX9, FOXA1 was constitutively associated with the CSGALNACT1 AR binding site in VCaP cells (Figure 7D) and in LNCaP cells stably expressing ERG (LN-to-ERG-2, Figure 7E), as well as DHT-stimulated AR binding to this site, while neither FOXA1 nor AR were bound to this site in control LNCaP cells (Figure 7E). Finally, DHT strongly stimulated CSGALNACT1 mRNA expression in LNCaP cells expressing ERG, but not in control LNCaP cells, and this expression was blocked by Bic (Supplemental Figure 7E). Overall, this genome-wide study shows that ERG allows AR to bind and transactivate a series of genes in TMPRSS2:ERG fusion–positive PCa that are not AR regulated in fusion-negative tumors.

Discussion

Approximately half of PCa cases express the AR-driven TMPRSS2:ERG fusion gene, but the downstream targets of ERG that contribute to PCa remain to be established. We found that increased SOX9 expression was correlated with TMPRSS2:ERG fusion in 3 independent PCa cohorts, and confirmed the ERG-dependent expression of SOX9 in TMPRSS2:ERG-positive VCaP cells. We further demonstrated that ERG regulates SOX9 indirectly by functioning as a pioneer factor to open a cryptic AR-regulated enhancer in the SOX9 gene, and that ERG can similarly redirect AR to a series of other genes. SOX9 expression in fetal prostate is required for ductal morphogenesis, and its aberrant expression has been linked to a number of cancers including PCa. We confirmed that SOX9, similarly to ERG, stimulates tumor invasion and can drive neoplasia when overexpressed in mouse prostate. Moreover, we showed that SOX9 RNAi impairs ERG-mediated tumor invasion and suppresses tumor growth in vitro and in vivo. Taken together, these studies reveal a molecular mechanism through which ERG functions to redirect AR to a series of genes including SOX9, and show that SOX9 is a major downstream effector of both AR and ERG in TMPRSS2:ERG fusion–positive PCa.

Significantly, primary PCa is characterized by increased formation of glands that lack basal cells and may have well-formed lumens (Gleason pattern 3) or ill-defined lumens (Gleason pattern 4), which is very reminiscent of developing prostate. Moreover, we showed that SOX9 RNAi impairs ERG-mediated tumor invasion and suppresses tumor growth in vivo. Taken together, these studies reveal a molecular mechanism through which ERG functions to redirect AR to a series of genes including SOX9, and show that SOX9 is a major downstream effector of both AR and ERG in TMPRSS2:ERG fusion–positive PCa.
of paracrine factors that support the luminal epithelium. In this case, ERG and AR-mediated SOX9 expression in PIN lesions may stimulate the autocrine production of these factors, and thereby permit the loss of basal cells that is a defining feature of PCa. Further studies are clearly needed to define the precise sets of SOX9-regulated genes that contribute to fetal prostate development, adult basal cell functions, and PCa.

Although these data reveal that ERG opens a cryptic AR-regulated enhancer in the SOX9 gene, the precise mechanisms through which this occurs and the roles of other proteins including FOXA1 remain to be precisely defined. A straightforward model is that ERG binding directly or indirectly increases local H3K4 methylation and subsequent binding of FOXA1, which then facilitates androgen-stimulated binding of AR. However, while FOXA1 can interact with AR and functions as a pioneer factor for AR binding to other enhancers, further RNAi studies are needed to determine whether AR binding to this site is dependent on FOXA1 or is facilitated by interactions with other proteins including ERG. Further studies are also needed to characterize the interaction between the S2 site and the SOX9 promoter, and specifically to determine whether the interaction is mediated by an AR-stimulated chromatin looping mechanism, as has been shown for other AR-regulated enhancers.

Substantial overlap between AR and ERG binding sites was also observed in a recent study. However, that study found that ERG was functioning to repress AR transcriptional activity at these sites, as well as to directly decrease expression of the AR gene (12). Our data are not inconsistent with this previous study, as we have not examined the effects of ERG on genes that are normally strongly stimulated by androgens. Moreover, they are not inconsistent with the hypothesis that ERG may stimulate tumor growth in part by repressing AR activity on a subset of AR-regulated genes. However, while this mechanism may contribute to tumor progression, we suggest that the ERG-mediated redirection of AR, rather than the suppression of AR, makes a more significant contribution to tumor initiation in TMPRSS2:ERG fusion–positive PCa.

In contrast to PCa cells expressing ERG, androgen repressed the low basal expression of SOX9 in LNCaP cells. Consistent with this result, a previous study found that castration increased SOX9 expression in murine prostate epithelium (36). Our data indicate that AR binding to the S1 site may mediate this repression, but more studies are needed to determine the role of this site in the response to androgen. In any case, the evolutionary conservation of the S1 site suggests a significant role for this region in regulating SOX9 expression. Indeed, it should be emphasized that additional mechanisms may drive SOX9 expression in TMPRSS2:ERG fusion–negative PCa, and may further modulate its expression in fusion-positive PCa. Interestingly, among these other mechanisms are the Wnt/β-catenin and receptor tyrosine kinase/MAP kinase pathways (24, 29, 47).

Therefore, higher levels of SOX9 expression (above the AR- and ERG-stimulated levels in fusion-positive tumors or above the lower basal levels in fusion-negative tumors) may be a biomarker for the activation of additional pathways that can drive PCa progression through SOX9-dependent and -independent mechanisms. Further studies of mechanisms controlling SOX9 and identification of the downstream genes it regulates should provide new therapeutic targets, which may be particularly relevant to TMPRSS2:ERG fusion–positive PCa.

Methods

Cell culture and xenografts. LNCaP cells were cultured in RPMI-1640 with 10% FBS. VCaP cells were cultured in DMEM with 10% FBS, and VCS2 cells were cultured in DMEM with 8% charcoal-dextran–stripped serum (CSS) and 2% FBS (HyClone). The Ln-to-ERG cell lines were generated by transfecting the pTET-Splice/ERG vector into a LNCaP clone carrying the pcDNA6/TR, followed by selection (34). pTET-Splice/ERG vector was made by inserting the HA-del 1-44 ERG coding fragment between the SalI-Sph sites of the pTET-Splice vector. The VCaP-toSOX9 cell line was generated by injection of lentivirus generated with the plLVX-Tet-On Advanced system (Clontech, in which a 3xFlag-SOX9 fragment was cloned between the BamHI and EcoRI sites of plLVX-Tight-Puro vector). The tet-inducible cells were maintained in RPMI-1640 with 10% tetracycline-free FBS. For androgen stimulation assays, cells were grown to 50%–60% confluency in medium with 5% CSS for 3 days prior to treatment. Cells were stimulated using 10 nM DHT unless otherwise indicated. VCaP xenografts were established in the flanks of male SCID mice by injecting approximately 2 million cells in 50% Matrigel.

RT-PCR, immunoblotting, and IHC. Quantitative real-time RT-PCR amplification was performed with Taqman 1-step RT-PCR reagents and results were normalized to coamplified 18S RNA or GAPDH. Data shown are representative of at least 3 experiments. Primers and probes are listed in Supplemental Materials. Proteins were extracted by boiling for 15 minutes in 2% SDS and detected by blotting with anti-ERG (Santa Cruz Biotechnology), anti-PSA (BioDesign), anti-AR (Upstate USA), anti-SOX9 (34), anti-β-actin (Abcam), or anti-β-tubulin (Upstate USA). Gels shown are representative of at least 3 independent experiments. Paraffin sections underwent antigen retrieval and were then blocked using 5% goat serum and avidin-blocking solution (Vector Laboratories). Anti-AR, anti-ERG (Santa Cruz Biotechnology), anti-SOX9, anti-FLAG (Sigma-Aldrich), or anti-Ki67 (Dako) were then added overnight at 4°C, followed by biotinylated secondary antibody and streptavidin-HRP (Vector Laboratories). Sections compared in each figure were stained at the same time and photographed under identical conditions.

SOX9 IHC and ERG fusion status in PCa tissue microarrays. IHC was performed on tissue microarrays (TMAs) created from prostatectomy samples among men with prostate cancer in the Physicians’ Health and Health Professionals Follow-up Studies tumor cohort as described in detail in 2 recent publications (42, 48). IHC was performed on 5-micron sections of 6 TMAs using the SOX9 antibody as described above. The immunostained TMAs were then scanned and quantitatively scored using the Ario SL-50 system (Genetix, Molecular Devices), after the areas of tumor were identified and electronically circled by a study pathologist (X. Yuan). Scores were generated using the MultiStain assay, providing areas of staining per cell (area_score), intensity of staining (intensity_score), and a combination of the two to produce a logarithmic score with an approximately normal distribution (addscore). The TMPRSS2:ERG fusion status of the TMA samples was determined by IHC using the ERG antibody (43). A 2-tailed Student’s t test was used to compare mean SOX9 expression in fusion-positive and -negative PCa.

CNP. Dispensed cells or xenograft tissues were formalin fixed, lyzed, and sonicated to break the chromatin into 500–800 bp fragments. Anti-AR, anti-p300, anti-oCT1, anti-GATA2, anti-RNA polymerase II, rabbit or mouse IgG (Santa Cruz Biotechnology), anti-RNA polymerase II CTD repeat (phospho Ser5), or anti-FOXA1 (Abcam) were used to precipitate chromatin fragments from cell extracts. Real-time qPCR was used to analyze binding to the SOX9-TSS, S1-S5 sites, and androgen-responsive elements (AREs) in the PSA enhancer (ARE3) or TMPRSS2 enhancer (14 kb upstream). The primers are listed in the Supplemental Information. We used real-time qPCR (SYBR Green) to amplify the DNA fragment in the
antibody-precipitated DNA and the unprecipitated input DNA to calculate ACT values. The relative quantity (RQ) value (RQ = 2^{−ΔΔCt}) represented the precipitated DNA relative to input. Results are represented as mean ± SD for replicate samples. Data are representative of at least 3 experiments.

CHIP-seq. CHIP-seq assay was performed in VCaP cells treated with DHT for 4 hours. CHIP-seq library construction for AR was performed as previously described (44), and the libraries were sequenced to 35 bp with the Illumina Genome Analyzer. Significantly (P < 1 × 10^-15) enriched regions were detected with MACS software using default parameters.

Cell proliferation and invasion assays. Cell growth was examined using the MTT assay (Sigma-Aldrich). Cell invasion was measured using the Cell Invasive Assay Kit from Chemicon following the manufacturer’s protocol. Briefly, cell suspensions in an upper chamber containing 800,000 cells/ml in serum-free medium were treated with or without 10 nM DHT and then monitored for invasion into a lower chamber containing RPMI-1640 medium with 10% FBS. After 72 hours of incubation at 37°C, cells attached to the membrane of the lower chamber were stained.

Transgenic mice. A 3xFlag-SOX9 cDNA fragment was cloned into the Hind III-EcoRV site of the pTet-splice vector, which contains 7 copies of the tet operator (TetO; Invitrogen). To generate Tet-STOP transgenic mice, the Xho I-NLS fragment containing the TetO-3xFlag-SOX9 and SV40 intron/polyadenylation signal was gel purified and microinjected into the pronuclei of fertilized mouse eggs at the Beth Israel Deaconess Transgenic Core Facility. The MMTV-tTA mice (49) were provided by M. Kelliher (University of Massachusetts Medical School, Worcester, Massachusetts, USA). The Pten+/− mice (C57Bl/6 background) were provided by P. Pandolfi (50). The combined Pten+/− transgenic mice overexpressing SOX9 were on a mixed C57Bl/6 and FVB background.

Data deposits. The following have been deposited in the Gene Expression Omnibus database: PCa tissue Affymetrix gene expression microarray data (Gene Expression Omnibus accession number GSE32269); VCaP cell line Affymetrix gene expression data (accession number GSE32344); and the results of the ChIP-seq analysis for AR binding in VCaP cells (accession number GSE32345).

**Statistics.** Data in bar graphs represent means ± SD of at least 3 biological repeats. Statistical analysis was performed by a 2-tailed Student’s t test. *P < 0.05, **P < 0.01 versus control values. P < 0.05 was considered statistically significant.

**Study approval.** The analyses of deidentified human tissue samples for gene expression were approved by the Beth Israel Deaconess Medical Center’s IRB. The development and analyses of genetically modified mice described in the manuscript were approved by the IACUC of Beth Israel Deaconess Medical Center.

Acknowledgments

This work was supported by grants from the NIH (R01 DK079962 to X. Yuan, and K99 CA166507 to C. Cai); SPORE in Prostate Cancer P50 CA090381; postdoctoral awards from the Department of Defense (W81XWH-08-1-0160 to H. Wang, and W81XWH-09-1-0159 to C. Cai); Idea Development Award from the Department of Defense (W81XWH-09-1L-0435 to S.P. Balk); and by a Challenge Award from the Prostate Cancer Foundation.

Received for publication August 31, 2012, and accepted in revised form December 17, 2012.

Address correspondence to: Steven P. Balk or Xin Yuan, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, Massachusetts 02215, USA. Phone: 617.735.2065; Fax: 617.735.2050; E-mail: sbalk@bidmc.harvard.edu (S.P. Balk). Phone: 617.735.2066; Fax: xyan@bidmc.harvard.edu (X. Yuan).