

ERG orchestrates chromatin interactions to drive prostate cell fate reprogramming

Fei Li, Qiuyue Yuan, Wei Di, Xinyi Xia, Zhuang Liu, Ninghui Mao, Lin Li, Chunfeng Li, Juan He, Yunguang Li, Wangxin Guo, Xiaoyu Zhang, Yiqin Zhu, Rebiguli Aji, Shangqian Wang, Xinyuan Tong, Hongbin Ji, Ping Chi, Brett Carver, Yong Wang, Yu Chen, Dong Gao

J Clin Invest. 2020. <https://doi.org/10.1172/JCI137967>.

Research Article Cell biology Oncology

Although cancer is commonly perceived as a disease of dedifferentiation, the hallmark of early-stage prostate cancer is paradoxically the loss of more plastic basal cells and the abnormal proliferation of more differentiated secretory luminal cells. However, the mechanism of prostate cancer proluminal differentiation is largely unknown. Through integrating analysis of the transcription factors (TFs) from 806 human prostate cancers, we found that ERG was highly correlated with prostate cancer luminal subtyping. ERG overexpression in luminal epithelial cells inhibited those cells' normal plasticity to transdifferentiate into a basal lineage, and ERG superseded PTEN loss, which favored basal differentiation. ERG KO disrupted prostate cell luminal differentiation, whereas AR KO had no such effects. Trp63 is a known master regulator of the prostate basal lineage. Through analysis of 3D chromatin architecture, we found that ERG bound and inhibited the enhancer activity and chromatin looping of a Trp63 distal enhancer, thereby silencing its gene expression. Specific deletion of the distal ERG binding site resulted in the loss of ERG-mediated inhibition of basal differentiation. Thus, ERG, in its fundamental role in lineage differentiation in prostate cancer initiation, orchestrated chromatin interactions and regulated prostate cell lineage toward a proluminal program.

Find the latest version:

<https://jci.me/137967/pdf>



ERG orchestrates chromatin interactions to drive prostate cell fate reprogramming

Fei Li,^{1,2} Qiuyue Yuan,^{3,4} Wei Di,^{1,2} Xinyi Xia,^{1,2} Zhuang Liu,^{1,2} Ninghui Mao,⁵ Lin Li,^{1,2} Chunfeng Li,^{1,2} Juan He,^{1,2} Yunguang Li,^{1,2} Wangxin Guo,^{1,2} Xiaoyu Zhang,^{1,2} Yiqin Zhu,^{1,2} Rebiguli Aji,^{1,2} Shangqian Wang,⁶ Xinyuan Tong,^{1,2} Hongbin Ji,^{1,2} Ping Chi,^{5,7,8,9} Brett Carver,^{5,10} Yong Wang,^{3,4,11,12} Yu Chen,^{5,7,8,9} and Dong Gao^{1,2,13}

¹State Key Laboratory of Cell Biology, Shanghai Key Laboratory of Molecular Andrology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai, China. ²University of Chinese Academy of Sciences, Beijing, China. ³Center for Excellence in Mathematical Sciences (CEMS), National Center for Mathematics and Interdisciplinary Sciences (NCMIS), Key Laboratory of Management, Decision and Information Systems (MDIS), Academy of Mathematics and Systems Science, National Center for Mathematics and Interdisciplinary Sciences, and ⁴School of Mathematical Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China. ⁵Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA. ⁶Department of Urology, First Affiliated Hospital, Nanjing Medical University, Nanjing, China. ⁷Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA. ⁸Department of Medicine and ⁹Department of Cell and Developmental Biology, Weill Cornell Medical College and New York-Presbyterian Hospital, New York, New York, USA. ¹⁰Division of Urology, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York, USA. ¹¹Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China. ¹²Key Laboratory of Systems Biology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China. ¹³Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing, China.

Although cancer is commonly perceived as a disease of dedifferentiation, the hallmark of early-stage prostate cancer is paradoxically the loss of more plastic basal cells and the abnormal proliferation of more differentiated secretory luminal cells. However, the mechanism of prostate cancer proluminal differentiation is largely unknown. Through integrating analysis of the transcription factors (TFs) from 806 human prostate cancers, we found that ERG was highly correlated with prostate cancer luminal subtyping. ERG overexpression in luminal epithelial cells inhibited those cells' normal plasticity to transdifferentiate into a basal lineage, and ERG superseded PTEN loss, which favored basal differentiation. ERG KO disrupted prostate cell luminal differentiation, whereas AR KO had no such effects. Trp63 is a known master regulator of the prostate basal lineage. Through analysis of 3D chromatin architecture, we found that ERG bound and inhibited the enhancer activity and chromatin looping of a Trp63 distal enhancer, thereby silencing its gene expression. Specific deletion of the distal ERG binding site resulted in the loss of ERG-mediated inhibition of basal differentiation. Thus, ERG, in its fundamental role in lineage differentiation in prostate cancer initiation, orchestrated chromatin interactions and regulated prostate cell lineage toward a proluminal program.

Introduction

Tumor initiation, progression, and therapy resistance involve epigenetic reprogramming that leads to aberrant cell lineage specification and transition (1–5). It is critical to understand the underlying mechanisms of cancer cell lineage differentiation and transition, which will provide novel insights into anti-cancer research. Master transcription factors (TFs) have been widely recognized as having a function in cell lineage transdifferentiation and cell fate reprogramming (6–8). The identification of master TFs in regulation of cancer cell lineage specification and transition would provide tremendous insights into the mechanism of lineage plasticity in cancer progression and therapy resistance.

Prostate cancer is the most common cancer and the second leading cause of cancer death in Western men (9). The normal prostate is a pseudostratified exocrine gland, and its epithelium consists of functional luminal cells that secrete proteins of the prostatic fluid, supportive basal cells that interact with stroma, and rare neuroendocrine cells. Compared with normal luminal cells, normal basal cells express higher levels of stemness genes (10) and exhibit greater stem cell properties, such as increased colony and organoid formation in vitro and graft formation in vivo (11–14). Although cancer is perceived as a disease of increased plasticity and stemness, primary and untreated prostate cancer is pathologically defined by luminal cell expansion and absence of basal cells with loss of P63 or CK5 by IHC. Primary and untreated prostate cancers that exhibit a true basal or neuroendocrine phenotype are extremely rare (15). We and others have recently found that normal single luminal cells when grown as organoids in vitro or grafted in vivo can form normal prostate glandular structures with both secretory luminal cells and basal cells that interact with stroma (11, 12, 16). The fact that cancerous luminal cells in human

Authorship note: FL, QY, WD, and XX contributed equally to this work.

Conflict of interest: YC has stock holdings in Oric Pharmaceuticals.

Copyright: © 2020, American Society for Clinical Investigation.

Submitted: March 9, 2020; **Accepted:** July 22, 2020; **Published:** October 5, 2020.

Reference information: *J Clin Invest.* <https://doi.org/10.1172/JCI137967>.

prostate cancer cannot form basal cells but directly interact with stroma suggests that prostate tumorigenesis paradoxically involves a loss of normal plasticity.

Under intense selection pressure of androgen deprivation therapy that targets the luminal lineage, progression to castration-resistant prostate cancer is associated with secondary gain of plasticity, with subsets of cancers that become neuroendocrine prostate cancer, androgen receptor (AR)/neuroendocrine double-negative prostate cancer, or AR/neuroendocrine double-positive “amphicrine” prostate cancer, some of which gain expression of basal markers (10, 17–19). These findings indicate that lineage differentiation and transition may play a pivotal role across multiple stages of prostate cancer progression.

Identification of the master TFs has provided significant insights in understanding the plasticity of prostate cancer lineages and the mechanism of therapy resistance. For example, N-Myc was identified as an oncogenic driver to promote neuroendocrine prostate cancer differentiation in the context of PI3K pathway activation in genetically engineered mouse models (20) and in a urogenital sinus mesenchyme (UGSM) tissue recombination assay of human prostate epithelial cells (21). In addition, SOX2 was recognized as a key TF to facilitate the lineage transition from prostate luminal cell lineage to neuroendocrine and basal cell lineage in TP53-deficient and RB1-deficient genetically engineered mouse models (22) and cellular models of human prostate cancer cell lines (23). Together, these findings proposed that SOX2 played a vital and context-dependent role in regulation of prostate cancer lineages. SOX11, as another member of the SOX gene family, also promoted neuroendocrine differentiation and treatment resistance to prostate cancers in the context of PTEN and TP53 inactivation (24). Given that advanced prostate cancer lineage is predominantly regulated by these known TFs, it is reasonable to question how primary prostate cancers gain their luminal differentiation features.

Several erythroblast transformation specific (ETS) family members have been demonstrated to be master TFs in the differentiation and lineage transition of several cell types (25–29). Previously, TMPRSS2-ERG fusion was reported as an early genetic alteration event occurring in 50% of prostate cancers (30, 31). Numerous previous studies revealed that ERG played an oncogenic role in promoting proliferation and invasion of prostate cancer cells (32–36). Further, we and others have shown that ERG is a master regulator that alters the chromatin enhancer landscape and AR cistrome (37–40). However, the function of ERG fusion during prostate cell lineage differentiation and transition is still largely unknown. Here, through integrating analysis of an integrative classifier (41) and the PAM50 classifier (42, 43), we have identified ERG as having the potential role of a master TF in prostate luminal lineage differentiation. In order to further consolidate functional verifications and to elucidate the detailed mechanisms of ERG’s role as a master TF, we performed multiomics analysis (RNA-Seq, ATAC-Seq, ChIP-Seq, and 3D genome analysis Bridge Linker-Hi-C [BL-Hi-C]) and preclinical model assays (in vitro organoid culture, in vivo transplantation, and genetically engineered mouse models). Through this analysis, we defined ERG as a master TF to regulate prostate luminal lineage through orchestrating chromatin interactions.

Results

Identification of the potential master TFs that regulate prostate cancer cell lineage. To identify the master TFs involved in prostate cancer lineage regulation, we developed a pipeline analysis by evaluating the correlation of TFs with prostate cancer subtyping (Figure 1A). We first chose 3 prostate cancer cohorts with available transcriptomic profiles: 158 samples in Fred Hutchinson Cancer Research Center (FHCRC), 150 samples in Memorial Sloan Kettering Cancer Center (MSKCC), and 498 samples in The Cancer Genome Atlas (TCGA); each cohort contained more than 100 human prostate cancer samples (31, 44, 45). In each cohort, we applied 2 prostate cancer subtyping methods, the PAM50 classifier categorizing prostate cancer into 3 lineage-related subtypes based on prostate lineage gene expression (42, 43); and the integrative classifier that revealed 3 distinct prostate cancer subtypes by combining several data types, including transcriptomic profiles and histone modifications (41). As expected, prostate cancer samples in the MSKCC cohort were categorized into 3 subtypes by the PAM50 classifier (Supplemental Figure 1A; supplemental material available online with this article; <https://doi.org/10.1172/JCI137967DS1>) and 3 subtypes by the integrative classifier (Supplemental Figure 1B). To estimate the relationship of each TF expression level with the above 3 cancer subtypes by the PAM50 or integrative classifier, cancer samples were further classified into 3 groups according to each TF’s expression level, termed TF-high, TF-medium, and TF-low. We next performed Pearson’s χ^2 test to identify TFs that significantly correlated with prostate cancer lineages by the PAM50 classifier and integrative classifier. For each cohort, overlapped TFs were further defined by overlapping the identified TFs by 2 subtyping methods (122 in FHCRC, 208 in MSKCC, and 399 in TCGA) (Supplemental Figure 1, C–E, and Supplemental Table 1). This combinational analysis ensured that the identified overlapped TFs would highly correlate with prostate cancer lineages and epigenetic classifiers. Taking the reproducibility and confidence into consideration, we defined the 154 master TFs from the overlapped TFs that were included in at least 2 of the 3 cohorts (Figure 1B and Supplemental Table 2). Among these TFs, ERG showed consistent and high correlation with prostate cancer subtyping in all 3 cohorts (Figure 1C). Gene Set Enrichment Analysis (GSEA) revealed that prostate luminal cell signature (10) was significantly enriched in ERG-high prostate cancer samples, validated in 2 different prostate cancer cohorts (Figure 1, D and E). As expected, AR and FOXA1 were also included in these TFs. FOXA1 was reported with a pioneering function in prostate cancer lineage differentiation and the determination programs (46). These results revealed the high efficiency of our method to identify the master TFs. Overall, these results indicated that ERG, as a master TF, highly correlates with prostate luminal cell lineage differentiation.

ERG regulates normal prostate epithelial cell lineage. To investigate the cell lineage plasticity of normal prostate epithelial cells, we isolated basal cells (Epcam⁺CD49f^{hi}YFP⁺) and luminal cells (Epcam⁺CD49f^{lo}YFP⁺) from the anterior prostate of tamoxifen-treated *Tmprss2*^{CreERT2/+} *Rosa26*^{YFP/YFP} (T2Y) mice and characterized the histology features of in vitro organoids and in vivo allografts (Supplemental Figure 2, A and B, and ref. 47). Immunofluorescence analysis of luminal- and basal cell-derived mouse prostate organoids demonstrated that both were comprised of

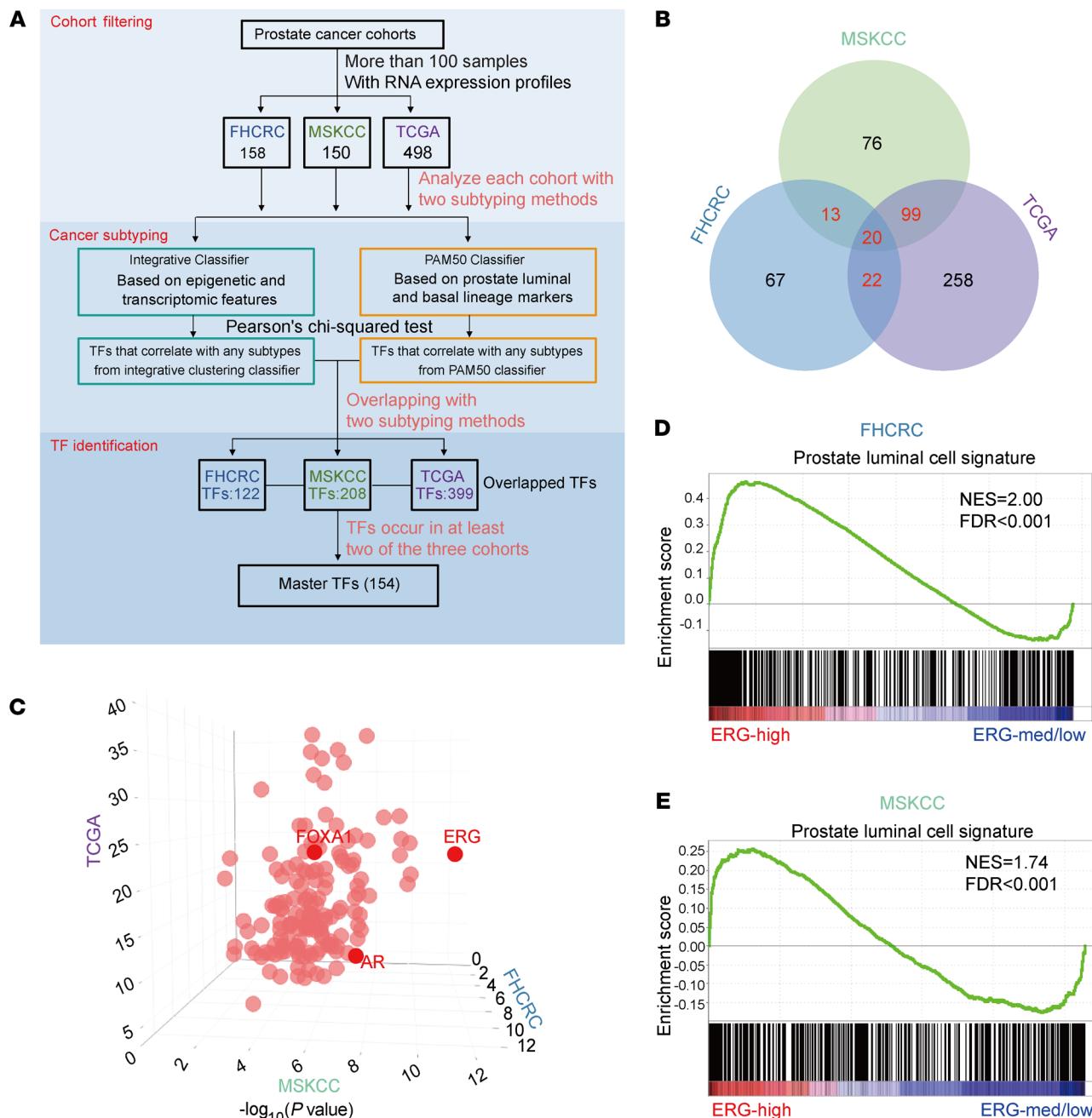


Figure 1. Identification of the master TFs that have the potential to regulate prostate cancer lineage. (A) Analysis pipeline to identify the master TFs involved in prostate cancer lineage differentiation containing 3 steps: (a) cohort filtering to select prostate cancer cohorts for downstream analysis, (b) cancer subtyping to categorize prostate cancer samples into several subtypes by 2 subtyping methods, and (c) TF identification to define master TFs with high reproducibility and confidence. (B) Venn diagram showing the number of master TFs generated from overlapped TFs that occurred in at least 2 of the 3 cohorts ($n = 13$ in MSKCC and FHCRC, $n = 99$ in MSKCC and TCGA, $n = 22$ in TCGA and FHCRC, $n = 20$ in all 3 cohorts). (C) Bubble plot of the 154 master TFs. The value for 3 axes represents $-\log_{10}(P\text{ value})$ calculated from Pearson's χ^2 test for MSKCC (x axis), FHCRC (y axis), and TCGA (z axis). (D and E) GSEA enrichment plot of ERG-high samples versus ERG-medium/low samples from FHCRC cohorts (D) (top) and MSKCC cohorts (E) (bottom) using signature genes of prostate luminal cells.

Krt8⁺ inner luminal cell layers and Krt5⁺ outer basal cell layers (Supplemental Figure 2C). Urogenital sinus mesenchyme (UGSM) tissue recombination assay is a useful in vivo method for prostate development and cancer research (48). Using a prostate UGSM tissue recombination assay, we further verified that basal and luminal prostate epithelial cells from T2Y mice could reconsti-

tute grafts with normal prostate architecture with Krt8⁺ luminal cell layers and Trp63⁺ basal cell layers in their renal grafts (Supplemental Figure 2D). Taken together, these results revealed that prostate luminal and basal cells maintained bipotential plasticity under in vitro organoid culture and in vivo renal transplantation conditions, similar to a previous study (12, 16).

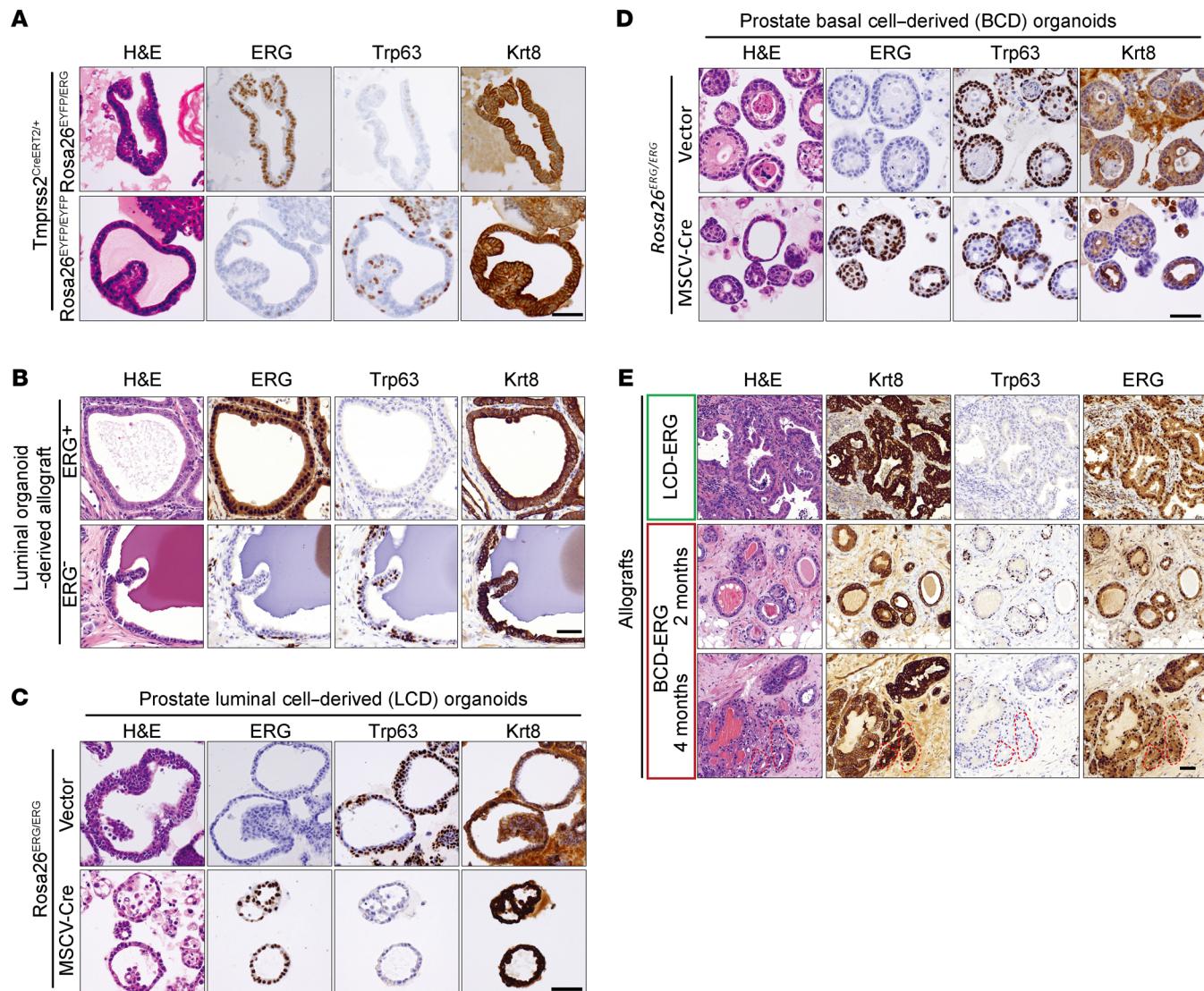


Figure 2. ERG promotes luminal lineage differentiation of normal prostate epithelial cells. (A) H&E and ERG, Trp63, and Krt8 IHC staining of luminal cell-derived organoids generated from T2YE (top) and T2Y (bottom) mice. (B) H&E and ERG, Trp63, and Krt8 IHC staining of allografts derived from luminal cell-derived organoids generated from T2YE (top) and T2Y (bottom) mice. (C and D) H&E and ERG, Trp63, and Krt8 IHC staining of luminal cell-derived (LCD) organoids (C) and basal cell-derived (BCD) organoids (D) generated from *Rosa26*^{ERT2/ERT2} mice infected with retrovirus carrying Cre recombinase (MSCV-Cre, bottom) or control backbone (MSCV-Vector, top). (E) H&E and Krt8, Trp63, and ERG IHC staining of allografts derived from LCD-ERG organoids (top) and BCD-ERG organoids (short-term for 2 months, middle; long-term for 4 months, bottom); red dashed line indicates the regions with predominant luminal phenotype. Scale bars: 50 μ m.

To explore the role of ERG expression in prostate cell lineage differentiation, we isolated luminal cells from the anterior prostates of tamoxifen-treated *Tmprss2*^{CreERT2/+} *Rosa26*^{ERT2/ERT2} (T2YE) mice and control T2Y mice to generate prostate organoids. Luminal cell-derived (LCD) YFP⁺ organoids from T2YE mice expressed ERG by IHC and were composed of a single luminal layer of Krt8⁺ cells with loss of Trp63⁺ basal cells, distinct from TY mice (Figure 2A). We further analyzed the organoids derived from prostate epithelial cells of *Pb-Cre4* *Rosa26*^{ERT2/ERT2} mice and *Tmprss2*-ERG knockin mice. We confirmed that the organoids with ERG expression from these 2 mice also maintained luminal cell features (Supplemental Figure 2E). Next, we performed UGSM tissue recombination assays with ERG⁺ and ERG⁻ LCD organoids that

were generated from T2YE and T2Y mice, respectively. The ERG⁺ allografts from T2YE mice exhibited pure luminal cell features with a single layer of Krt8⁺ luminal cells after 2 months of transplantation (Figure 2B). On the other hand, the ERG⁻ grafts from T2Y mice regenerated the normal prostate architecture composed of Krt8⁺ luminal cells and Trp63⁺ basal cells (Figure 2B). Together, these results suggest that ERG overexpression could maintain luminal cell lineage features under the conditions of in vitro 3D organoid culture and in vivo UGSM tissue recombination.

Given the potential role of ERG in lineage differentiation, we next sought to examine the differences in lineage responses in luminal cells and basal cells with ERG overexpression. Briefly, we performed FACS sorting to isolate prostate luminal cells

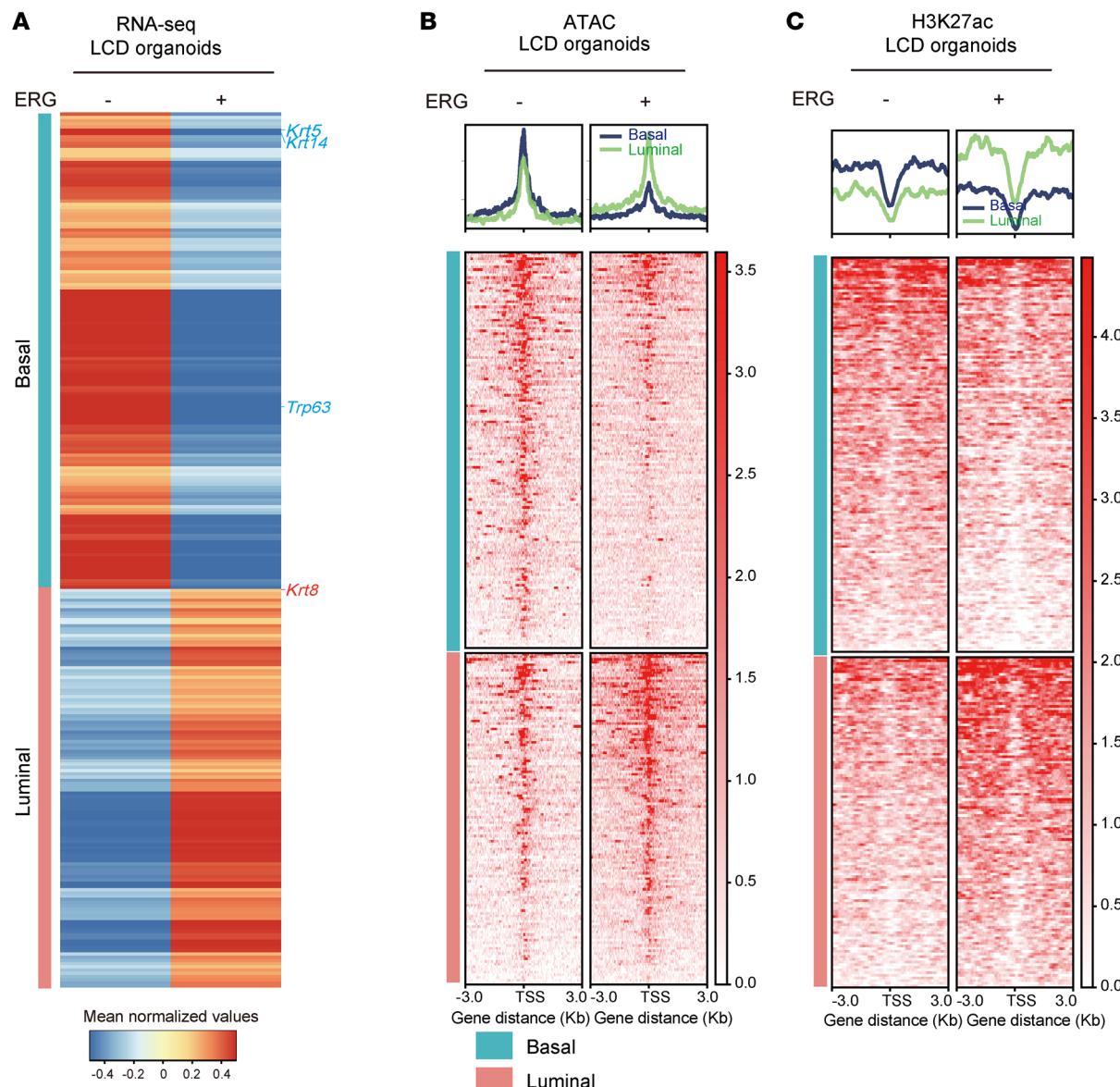


Figure 3. Multiomics data demonstrated the function of ERG in promoting luminal lineage differentiation. (A) Heatmap of RNA-Seq showing the expression of downregulated basal lineage genes and upregulated luminal lineage genes in LCD and LCD-ERG organoids. (B and C) Profile plot (top) and heatmap (bottom) of ATAC-Seq (B) and H3K27ac ChIP-Seq (C) around the transcriptional start site (TSS) of downregulated basal lineage genes and upregulated luminal lineage genes in LCD and LCD-ERG organoids.

($\text{Cd49f}^{\text{lo}}\text{Cd24}^{\text{hi}}$) and basal cells ($\text{Cd49f}^{\text{hi}}\text{Cd24}^{\text{lo}}$) from *Rosa26^{ERG/ERG}* mice (Supplemental Figure 2F). Intracellular flow cytometry for Krt5 and Krt8/Krt18 verified the cellular identities of the 2 populations. Next, we retrovirally transduced the Cre recombinase or a retrovirus control into these basal cell-derived (BCD) organoids or LCD organoids. Remarkably, ERG expression in LCD organoids (LCD-ERG) induced a single Krt8⁺ luminal cell layer with loss of the Trp63⁺ basal cell layer, strongly indicating the predominant role of ERG in prostate cell luminal lineage differentiation (Figure 2C). BCD organoids with ERG expression (BCD-ERG) still maintained a Trp63⁺ outer basal cell layer, but with an apparent decrease in the number of Trp63⁺ basal cells (Figure 2D and Supplemental Figure 2G). In addition, we performed UGSM tissue recombination assays to validate these findings *in vivo*. Allografts derived from LCD-ERG

organoids exhibited pure luminal cell features with the absence of Trp63⁺ basal cells after 2 months of transplantation (Figure 2E). On the other hand, allografts derived from BCD-ERG organoids were composed of ERG⁺Krt8⁺ luminal cells and ERG⁺Trp63⁺ basal cells after 2 months of transplantation (Figure 2E). Intriguingly, BCD-ERG allografts also exhibited predominant luminal features with the absence of Trp63⁺ basal cells after 4 months of transplantation (Figure 2E). Collectively, these results demonstrated that ERG promoted prostate luminal lineage differentiation with LCD organoids more vulnerable to ERG-induced luminal lineage differentiation when compared with BCD organoids.

To identify whether ERG expression-induced lineage changes were associated with chromatin status, we next performed integrative analyses of transcriptome (RNA-Seq) and chromatin

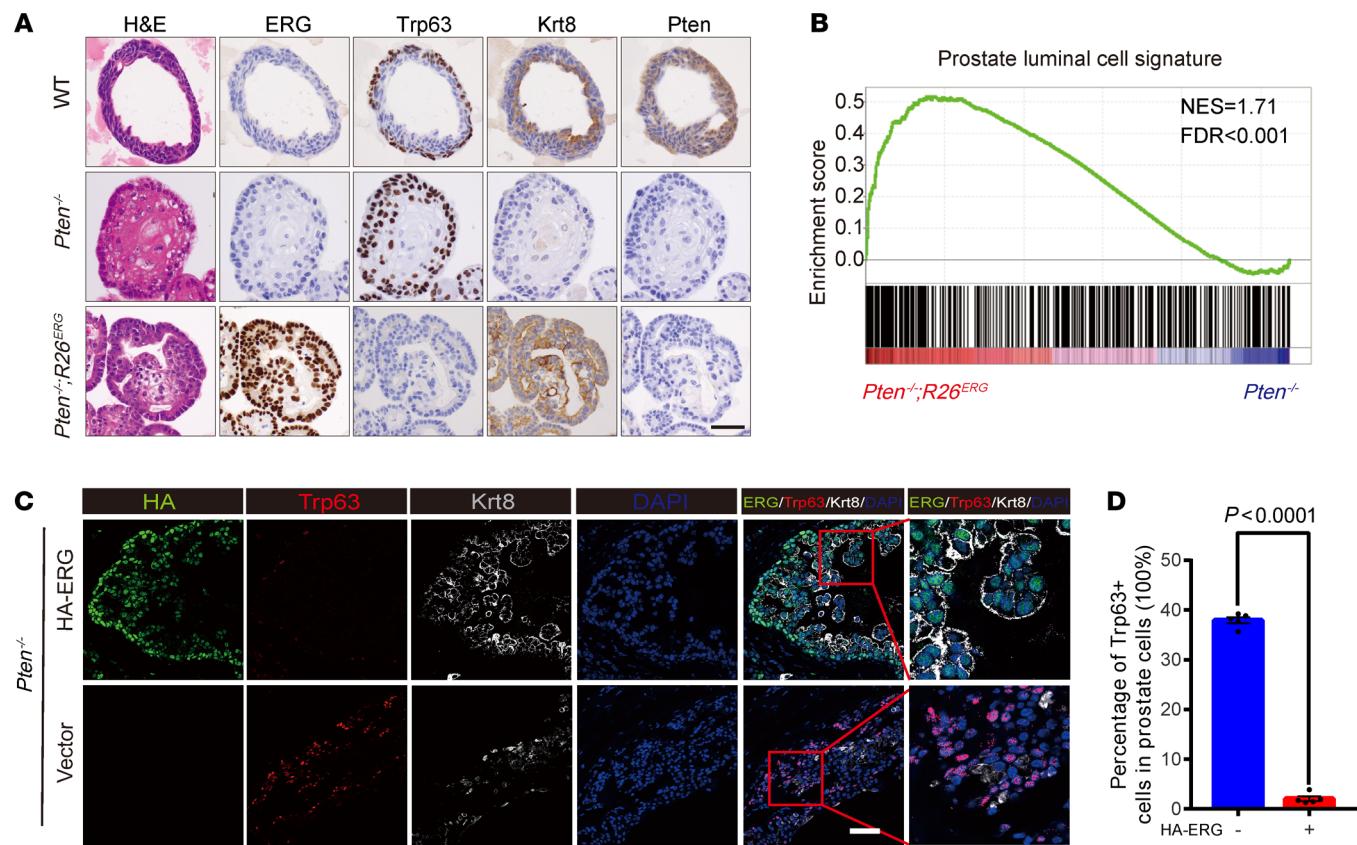


Figure 4. ERG promotes luminal differentiation of prostate cancer cells under the Pten loss condition. (A) H&E and ERG, Trp63, Krt8, and Pten IHC staining of WT (top), *Pten*^{-/-} (middle), and *Pten*^{-/-};R26^{ERG} (bottom) organoids. (B) GSEA enrichment plot of *Pten*^{-/-};R26^{ERG} organoids versus *Pten*^{-/-} organoids using prostate luminal cell signature genes. (C) HA, Trp63, Krt8, and DAPI IF staining of allografts derived from UGSM tissue recombination assay in SCID mice 8 weeks after transplantation of organoids overexpressing the TMPRSS2-ERG fusion protein with HA tag (top) or a control vector (bottom). (D) Quantification statistics for the percentage of Trp63⁺ cells in prostate cells (analyses were performed based on 2522 HA⁺ cells and 3086 vector cells, 2-tailed *t* test, mean ± SEM, *n* = 5). Scale bars: 50 μ m.

accessibility (ATAC-Seq) of LCD organoids and LCD-ERG organoids. By assessing the lineage changes in both mRNA expression and chromatin accessibility, we identified 177 downregulated basal signature genes, such as *Krt5*, *Krt14*, and *Trp63*, with decreased chromatin accessibility at their promoters in LCD-ERG organoids compared with LCD organoids. On the other hand, 86 upregulated luminal signature genes, such as *Krt8*, with increased chromatin accessibility were identified in LCD-ERG organoids compared with LCD organoids (Figure 3, A and B). We further confirmed the increased H3K27ac levels of the upregulated luminal signature genes and attenuated H3K27ac levels of the downregulated basal signature genes at their promoters in LCD-ERG organoids compared with LCD organoids (Figure 3C). Collectively, these data suggested that ERG-induced changes in the expression of lineage genes were associated with chromatin status, including chromatin accessibility and histone modifications, revealing the potential relationship mediated by ERG between lineage regulation and chromatin status.

ERG regulates prostate cancer cell lineage. Given the above finding that ERG could promote luminal lineage differentiation in normal prostate epithelial cells, we next investigated whether ERG could regulate the luminal lineage differentiation of pros-

tate cancer cells. The combination of ERG rearrangements and loss of PTEN is regarded as one of the most concurrent genetic events in human prostate cancer (36, 49, 50). We generated a *Tmprss2*^{CreERT2/+} *Pten*^{fl/fl} *Rosa26*^{ERG/ERG} (T2PE) mouse model to test ERG function in a Pten loss condition. Because of heterogeneous recombination efficiency, ERG was only expressed in a subset of Pten-deleted regions. We examined the histological features of the T2PE mice prostates at 7 months after tamoxifen injection. Remarkably, ERG⁺ prostate epithelial cells with Pten deletion exhibited pure Krt8⁺ luminal features and the mutual exclusion of *Krt5*⁺ and *Trp63*⁺ prostate basal epithelial cells (Supplemental Figure 3A). On the other hand, ERG⁻ Pten-deleted prostate epithelial cells exhibited increased levels of basal differentiation with the expansion of *Krt5*⁺ and *Trp63*⁺ prostate epithelial cells (Supplemental Figure 3A). Notably, neither the combination of ERG and *Krt5* nor that of ERG and *Trp63* showed colocalization, which was confirmed by costaining assays of immunofluorescence (Supplemental Figure 3, B and C). Next, we isolated *Pten*^{-/-} and *Pten*^{-/-};R26^{ERG} prostate cancer cells individually from the harvested prostates of 20-month-old *Pb-Cre4* *Pten*^{fl/fl} mice and *Pb-Cre4* *Pten*^{fl/fl} *Rosa26*^{ERG/ERG} mice, respectively. We found that prostate cancer cells with Pten loss differentiated toward basal

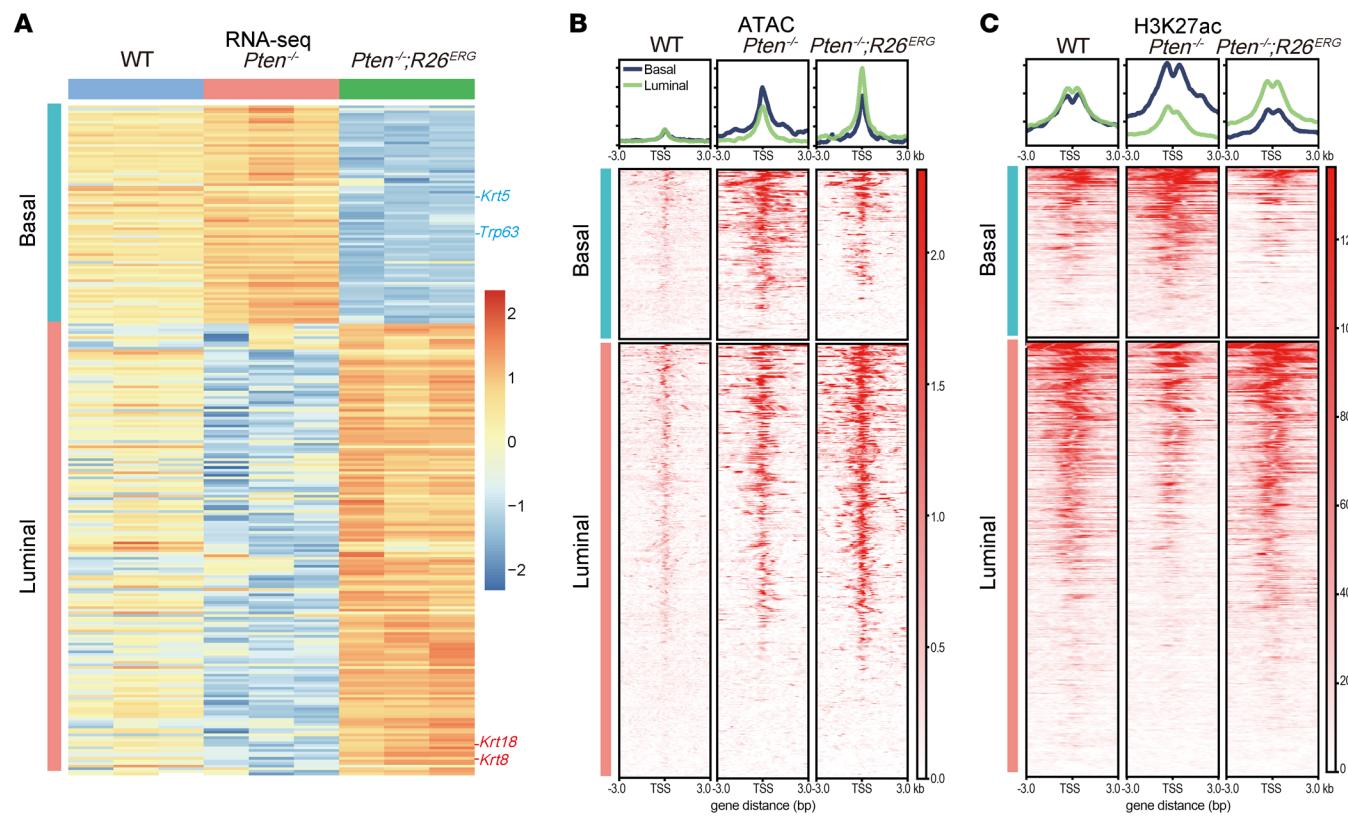


Figure 5. Chromatin status associated with ERG-promoted lineage changes. (A) Heatmap showing the expression of ERG-upregulating luminal cell signature genes and ERG-downregulating basal cell signature genes in WT, *Pten*^{-/-}, and *Pten*^{-/-};R26^{ERG} organoids from RNA-Seq. (B and C) Profile plot (top) and heatmap (bottom) of ATAC-Seq (B) and H3K27ac ChIP-Seq (C) around the transcriptional start site (TSS) of ERG-upregulating luminal cell signature genes and ERG-downregulating basal cell signature genes in WT, *Pten*^{-/-}, and *Pten*^{-/-};R26^{ERG} organoids.

lineage with the predominant expansion of Trp63⁺ basal cells, whereas ERG overexpression maintained luminal features in the context of Pten null (Figure 4A and Supplemental Figure 3D). Compared with *Pten*^{-/-} organoids, the *Pten*^{-/-};R26^{ERG} organoids exhibited increased expression of luminal cell lineage markers, such as Krt8 and Krt18, and were negative for basal cell lineage markers, such as Krt5, Krt14, and Trp63 (Supplemental Figure 3, E and F). To further validate the above results *in vivo*, we performed UGSM tissue recombination assays of *Pten*^{-/-} organoids overexpressing the TMPRSS2-ERG fusion gene with an HA tag or a control vector. Consistent with our previous work, ERG overexpression efficiently promoted tumor growth (Supplemental Figure 4E and refs. 36, 38). In addition, ERG expression promoted luminal differentiation of prostate cancer cells under the Pten loss condition (Figure 4, C and D, and Supplemental Figure 4D), further suggesting the conserved role of ERG in regulating prostate cell luminal features using multiple models.

To characterize the impact of ERG on global gene expression, we performed RNA-Seq on WT, *Pten*^{-/-}, and *Pten*^{-/-};R26^{ERG} organoids. GSEA analysis showed that the prostate luminal cell signature genes were highly enriched in *Pten*^{-/-};R26^{ERG} organoids, rather than in *Pten*^{-/-} organoids (Figure 4B). Notably, the expression of luminal cell lineage markers (Krt8 and Krt18) and basal cell lineage markers (Krt5 and Trp63) were all included in the differentially expressed genes (DEGs) (Figure 5A), consistent with the results of qRT-PCR

analyses and Western blotting (Supplemental Figure 3, E and F; see complete unedited blots in the supplemental material). Here, we defined an ERG-upregulating luminal signature (177 genes, Supplemental Table 3) by using the overlap between upregulated DEGs of *Pten*^{-/-};R26^{ERG} organoids and prostate luminal cell signature; therefore, these genes were rigorously associated with both ERG expression and prostate luminal lineage. On the other hand, the ERG-downregulating basal signature (86 genes, Supplemental Table 3) was also defined by using the overlap between downregulated DEGs of *Pten*^{-/-};R26^{ERG} organoids and prostate basal cell signature. ATAC-Seq and H3K27ac ChIP-Seq were also performed to systematically investigate the transcriptomic and epigenetic regulations associated with ERG expression. Consistently, through ATAC-Seq and H3K27ac ChIP-Seq analyses, we identified the increases of chromatin accessibility and H3K27ac levels at the promoters of luminal cell lineage markers (Krt8 and Krt18), as well as the decreases of chromatin accessibility and H3K27ac levels at the promoters of basal cell lineage markers (Krt5 and Krt14) in *Pten*^{-/-};R26^{ERG} organoids when compared with *Pten*^{-/-} organoids (Figure 5, B and C, and Supplemental Figure 4, B and C). In addition, since WT organoids were composed of luminal and basal cells, medium levels of mRNA expression and medium H3K27ac levels in the promoters of luminal and basal lineage genes were identified in WT prostate organoids compared with *Pten*^{-/-} and *Pten*^{-/-};R26^{ERG} prostate cancer organoids (Figure 5, A and C, and Supplemental Figure 4, A–C).

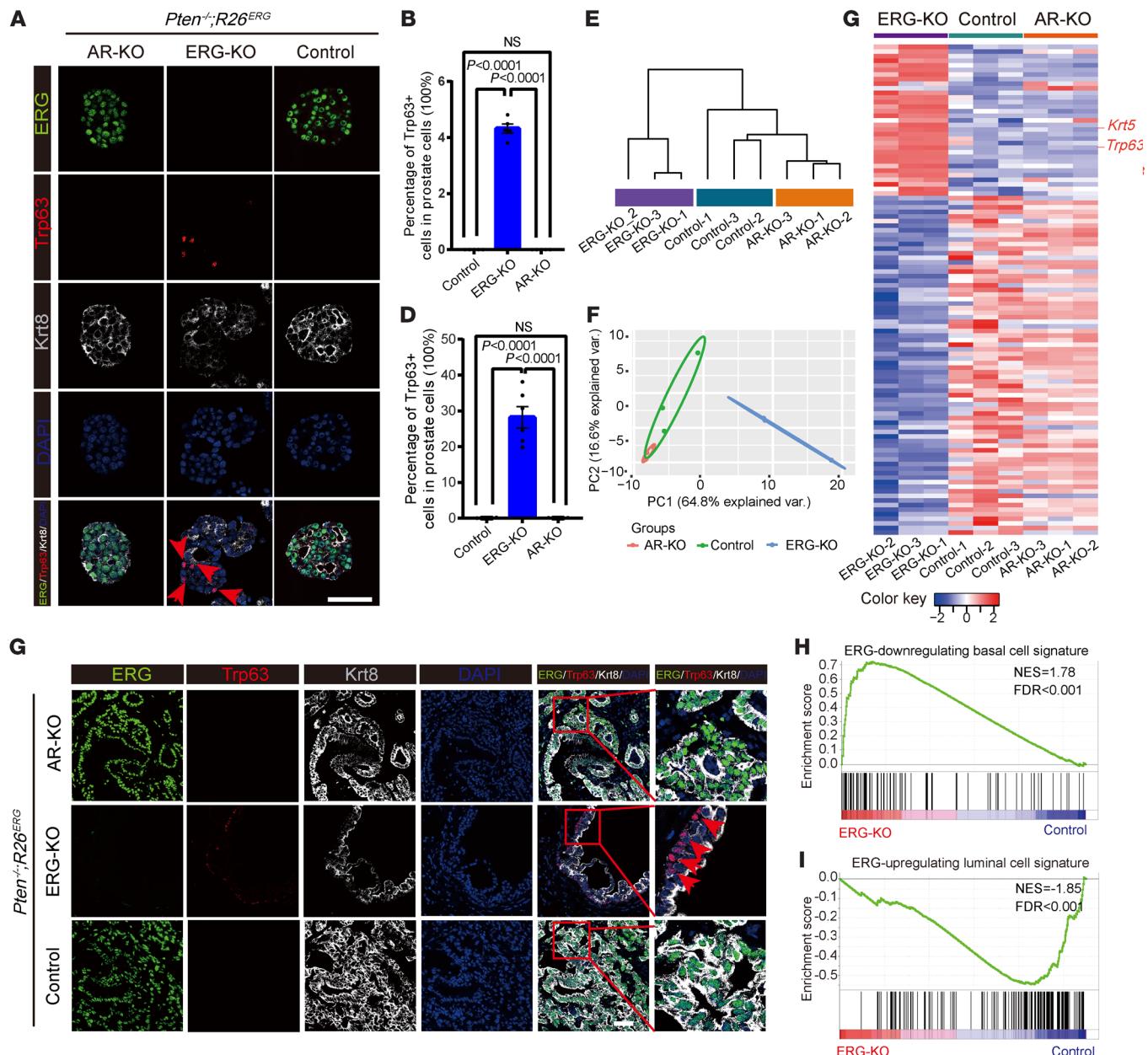


Figure 6. ERG, not AR, is required for sustaining the luminal phenotype of prostate cancer cells in the context of Pten loss. (A) ERG, Trp63, Krt8, and DAPI IF staining of *Pten*^{-/-} *R26*^{ERG} organoids infected with a lentiviral CRISPR/Cas9 carrying guide RNA targeting the AR (AR-KO, left) and ERG (ERG-KO, middle) and a control vector (control, right); red arrows indicate Trp63⁺ cells. (B) Quantification statistics for the percentage of Trp63⁺ cells in total cells in vitro (analyses were performed based on 9080 AR-KO cells, 8756 ERG-KO cells, and 6498 control cells, 1-way ANOVA and multiple comparisons, mean ± SEM, $n = 5$). (C) ERG, Trp63, Krt8, and DAPI IF staining of grafts derived from UGSM tissue recombination assays in SCID mice 8 weeks after transplantation of *Pten*^{-/-} *R26*^{ERG} organoids with AR-KO (top), ERG-KO (middle), and control (bottom); red arrows indicate Trp63⁺ cells. (D) Quantification statistics for the percentage of Trp63⁺ cells in total prostate cells in vivo (analyses were performed based on 3236 AR-KO cells, 2206 ERG-KO cells, and 3554 control cells, 1-way ANOVA and multiple comparisons, mean ± SEM, $n = 6$). (E and F) Clustering dendrogram (E) and PCA plot (F) for ERG-KO, AR-KO, and control organoids using prostate cell lineage signature genes. (G) Heatmap showing the expression of lineage-related differentially expressed genes in ERG-KO, AR-KO, and control organoids. (H and I) GSEA enrichment plot of ERG-KO versus control using ERG-downregulating basal cell signature genes (H) and ERG-upregulating luminal cell signature genes (I). Scale bars: 50 μ m.

Collectively, these results suggested ERG as a master regulator to manipulate the luminal lineage of prostate cancer cells, tightly associated with epigenetic regulation.

ERG but not AR is sufficient to maintain luminal lineage in Pten-loss prostate cancer. AR is a well-known TF highly expressed in luminal prostate cells, but is dispensable for Pten loss-mediated tumor-

igenesis in the mouse prostate (51, 52). To determine whether ERG or AR is required to maintain luminal differentiation in prostate cancer, we performed CRISPR/Cas9-mediated AR-KO and ERG-KO in *Pten*^{-/-} *R26*^{ERG} organoids. AR-targeted genes, such as *Fkbp5*, *Nkx3.1*, and *Mme*, were significantly decreased in *Pten*^{-/-} *R26*^{ERG} organoids with AR-KO (Supplemental Figure 5A). AR-KO in *Pten*^{-/-}

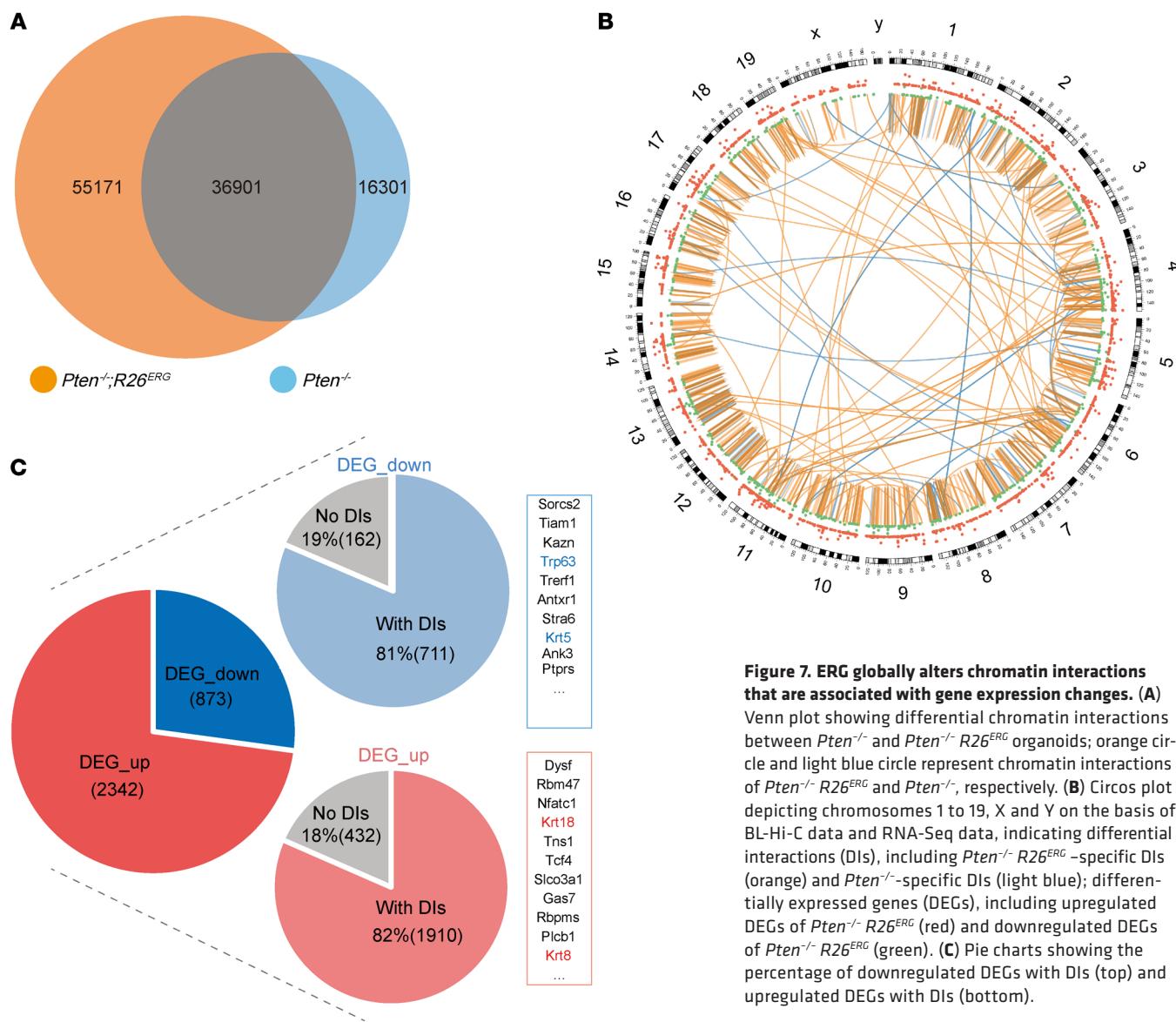


Figure 7. ERG globally alters chromatin interactions that are associated with gene expression changes. (A) Venn plot showing differential chromatin interactions between *Pten*^{-/-} and *Pten*^{-/-} *R26^{ERG}* organoids; orange circle and light blue circle represent chromatin interactions of *Pten*^{-/-} *R26^{ERG}* and *Pten*^{-/-}, respectively. **(B)** Circos plot depicting chromosomes 1 to 19, X and Y on the basis of BL-Hi-C data and RNA-Seq data, indicating differential interactions (DIs), including *Pten*^{-/-} *R26^{ERG}*-specific DIs (orange) and *Pten*^{-/-}-specific DIs (light blue); differentially expressed genes (DEGs), including upregulated DEGs of *Pten*^{-/-} *R26^{ERG}* (red) and downregulated DEGs of *Pten*^{-/-} *R26^{ERG}* (green). **(C)** Pie charts showing the percentage of downregulated DEGs with DIs (top) and upregulated DEGs with DIs (bottom).

R26^{ERG} organoids still maintained their pure prostate luminal histology (*Krt8*⁺/*Trp63*⁺) without obvious lineage changes, which was also evident in *in vivo* UGSM tissue recombination assays (Figure 6, A–D, and Supplemental Figure 5, D–F). On the contrary, ERG-KO in *Pten*^{-/-} *R26^{ERG}* organoids resulted in the loss of pure luminal differentiation and appearance of many cells that expressed basal lineage markers (*Trp63*, *Krt5*, and *Krt14*) in 3D organoids and renal grafts (Figure 6, A–D, and Supplemental Figure 5, D–F). In addition, the dramatic decrease in the percentage of *Ki67*⁺ cells was attributable to ERG-KO, reinforcing the oncogenic role of ERG in the context of *Pten* loss (Supplemental Figure 5, F–H).

Based on the expressions of prostate lineage genes, hierarchical clustering and PCA analyses were performed to evaluate the similarities among AR-KO, ERG-KO, and control *Pten*^{-/-} *R26^{ERG}* organoids. AR-KO showed only relatively small changes with control *Pten*^{-/-} *R26^{ERG}* organoids, whereas ERG-KO organoids were clearly separated from AR-KO and control *Pten*^{-/-} *R26^{ERG}* organoids (Figure 6, E and F). Moreover, remarkably increased basal cell lineage markers, such as *Krt5* and *Trp63*, were also verified in

ERG-KO organoids through RNA-Seq DEGs analyses of AR-KO, ERG-KO, and control *Pten*^{-/-} *R26^{ERG}* organoids (Figure 6G). In addition, GSEA was performed to evaluate the changes in prostate cell lineage, revealing that ERG-downregulating basal signature genes were significantly enriched in ERG-KO organoids, whereas ERG-upregulating luminal signature genes were enriched in control *Pten*^{-/-} *R26^{ERG}* organoids (Figure 6, H and I). Consistently, no significant differences were identified in the expression of prostate lineage genes between AR-KO and the control (Supplemental Figure 5I). These findings further confirmed the importance of ERG in the lineage regulation of prostate cancer cells. Moreover, in the context of both *Pten* loss and ERG expression, AR deletion had no significant effects on prostate cell lineage differentiation, suggesting that luminal lineage regulation in primary prostate cancer cells does not rely on AR.

ERG induces global changes in chromatin interactions. Chromatin dynamics are highly correlated with cell fate reprogramming (53–55). To examine whether ERG expression induces changes in chromatin interactions, we performed BL-Hi-C (56) in LCD, LCD-

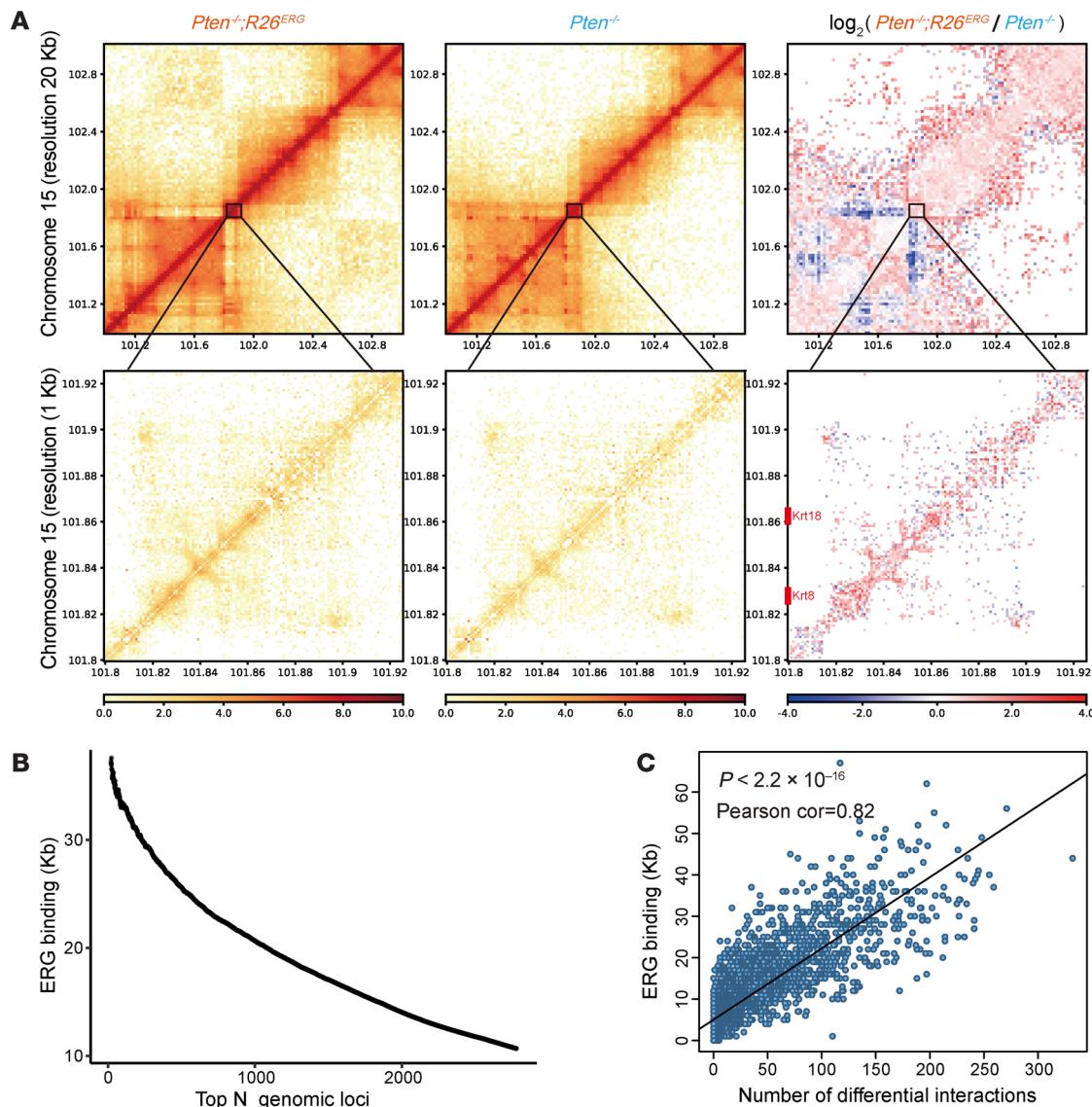


Figure 8. Differential chromatin interactions positively correlated with ERG binding activity. (A) The normalized interaction heatmaps of *Pten^{-/-};R26^{ERG}* (left), *Pten^{-/-}* (middle), and the difference (right) at 20 kb resolution (top) and 1 kb resolution (bottom) of chromosome 15, including *Krt8* and *Krt18* gene loci. (B) Plot showing the density of ERG binding (Kb) at each of the ranked (N) differential interacting chromatin loci of 1 Mb intervals. (C) Correlation plot showing the significant positive relationship between ERG binding density and the number of DIs in 1 Mb intervals.

ERG, *Pten^{-/-}*, and *Pten^{-/-};R26^{ERG}* organoids. On average, each library contained over 470 million unique pairwise contacts, which had high quality with over 80% of *cis*-pairs in total valid pairs (Supplemental Figure 6, A and B). After the systematic loop calling, we found that ERG expression resulted in an increased number of interaction loops in both Pten-intact and Pten-loss contexts (Figure 7A and Supplemental Figure 6C). A Circos plot globally visualized the differential interactions (DIs) across the 21 chromosomes (chromosome 1–19, X and Y) and demonstrated that ERG expression enhanced chromatin interactions (Figure 7B and Supplemental Figure 6D). To investigate the associations between chromatin interactions and gene expression, we next correlated DEGs with DIs of *Pten^{-/-};R26^{ERG}* organoids compared with *Pten^{-/-}* organoids (Figure 7C). Remarkably, the percentages of downregulated DEGs with DIs reached 81% (711 of 873, $P = 5.89 \times 10^{-83}$), including *Trp63*

and *Krt5*. Moreover, 82% (1910 of 2342, $P = 2.77 \times 10^{-176}$) of upregulated DEGs were found with DIs, including *Krt8* and *Krt18*. When similar analyses were performed on LCD-ERG and LCD organoids in the Pten-intact setting, we found that 79% (1270/1612, $P = 1.82 \times 10^{-118}$) of downregulated DEGs and 80% (802/1001, $P = 4.61 \times 10^{-89}$) of upregulated DEGs were mapped with DIs (Supplemental Figure 6E). To further explore the enrichment pattern of chromatin interactions in prostate lineage-related loci, the *Krt8* and *Krt18* genomic regions in chromosome 15 were chosen, and contact maps were shown at 20 kb and 1 kb resolution. Upon close inspection of these regions, we observed that enhanced chromatin interactions were detected in ERG-expressing organoids, including *Pten^{-/-};R26^{ERG}* organoids (Figure 8A) and LCD-ERG organoids (Supplemental Figure 6F). These observations indicate that the gene expression changes induced by ERG were highly associated

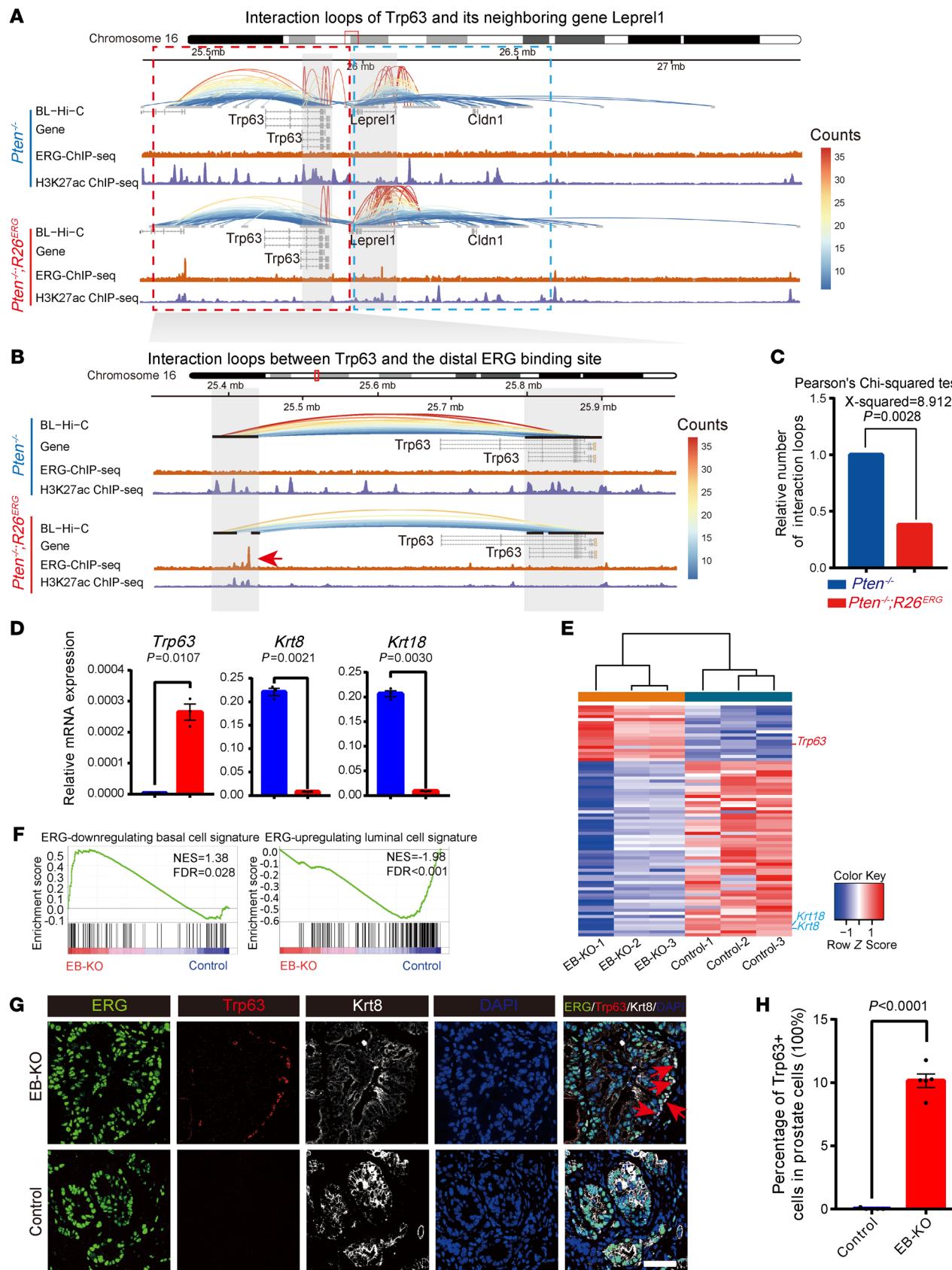


Figure 9. Deletion of a specific ERG binding site impaired the function of ERG in prostate lineage regulation. (A) 3D signal of BL-Hi-C showing chromatin interactions of *Trp63* loci and its neighboring gene *Leprel1* loci in *Pten*^{-/-} (top) and *Pten*^{-/-} *R26^{ERG}* (bottom) organoids; red box indicates the highly interacting region of *Trp63* loci, blue box indicates the highly interacting region of *Leprel1* loci. (B) 3D signal of BL-Hi-C showing chromatin interactions between the distal ERG binding site and *Trp63* gene body region in *Pten*^{-/-} (top) and *Pten*^{-/-} *R26^{ERG}* (bottom) organoids. Red arrow indicates the distal ERG binding site. (C) Pearson's χ^2 test to evaluate the differences of interaction loops density between *Pten*^{-/-} and *Pten*^{-/-} *R26^{ERG}* organoids. (D) qRT-PCR analysis of *Trp63*, *Krt8*, and *Krt18* mRNA expression in EB-KO and control of *Pten*^{-/-} *R26^{ERG}* organoids (2-tailed *t* test, mean \pm SEM, $n = 3$). (E) Heatmap of RNA-Seq for EB-KO and control of *Pten*^{-/-} *R26^{ERG}* organoids using differentially expressed prostate cell lineage signature genes. (F) GSEA enrichment plot of EB-KO organoids versus control organoids using ERG-downregulating basal cell signature genes (left) and ERG-upregulating luminal cell signature genes (right). (G) ERG, *Trp63*, *Krt8*, and DAPI IF staining for allografts of UGSM tissue recombination assays derived from EB-KO and control organoids; red arrows indicate ERG⁺*Trp63*⁺ cells. (H) Quantification statistics for the percentage of ERG⁺*Trp63*⁺ in ERG⁺ prostate cells (analyses were performed based on 3239 ERG⁺ cells of EB-KO and 3806 ERG⁺ cells of control, 2-tailed *t* test, mean \pm SEM, $n = 5$). Scale bars: 50 μ m.

with the alterations of chromatin interactions. To directly characterize the role of ERG in chromatin interactions, we binned the genome into 1 Mb intervals and analyzed the total DIs in these genomic bins. Importantly, we observed the preferential ERG binding occupancy of genomic bins with more DIs (Figure 8B and Supplemental Figure 6G); such positive correlation was also confirmed by Pearson correlation analysis (Figure 8C and Supplemental Figure 6H). Taken together, these results suggest that ERG binding occupancy significantly correlated with differential chromatin interactions, which were also highly associated with DEGs, indicating the potential role of ERG in transcriptional programs through reorganization of chromatin interactions to facilitate cell lineage regulation.

Deletion of a specific ERG binding site disrupts the function of ERG in prostate lineage regulation. Given the associations between transcriptional regulations induced by ERG and chromatin interactions, we next asked whether such associations were functionally related to prostate lineage regulation. Through integrating motif enrichment analysis with transcriptional expression changes generated from ATAC-Seq and RNA-Seq, we found that *Trp63* exhibited high potential as a master TF in LCD organoids and *Pten*^{-/-} organoids, both of which contained cells with basal cell differentiation (Supplemental Figure 7A). Concordantly, ERG played a pivotal role in prostate lineage regulation that was verified in LCD-ERG organoids and *Pten*^{-/-} *R26^{ERG}* organoids (Supplemental Figure 7A). Indeed, *Trp63* is a known master regulator of the prostate basal cell lineage, and *Trp63*-KO mice failed to develop basal cells (57, 58).

To determine whether *Trp63* expression could be regulated by ERG through altering chromatin interactions, we first examined the chromatin interactions of the *Trp63* (Δ Np63) loci in *Pten*^{-/-} and *Pten*^{-/-} *R26^{ERG}* organoids by BL-Hi-C. The attenuated chromatin interactions of the *Trp63* loci were identified, whereas the chromatin interactions of its neighboring gene loci, *Leprel1*, were remarkably increased in *Pten*^{-/-} *R26^{ERG}* organoids

compared with *Pten*^{-/-} organoids (Figure 9A). In addition, almost all the chromatin interactions of the *Trp63* loci were distributed between the loci and the region at 400 kb upstream of the *Trp63* promoter in *Pten*^{-/-} and *Pten*^{-/-} *R26^{ERG}* organoids. Intriguingly, this region was accompanied by a strong ERG binding site in *Pten*^{-/-} *R26^{ERG}* organoids (Figure 9A). This result indicated a potential role of this ERG binding site in mediating the associations between *Trp63* expression and chromatin interactions. We next specifically investigated the chromatin interactions and histone modifications between this distal ERG binding site and *Trp63* loci. Upon close inspection of this region, we observed an enhancer strongly enriched for the H3K27ac histone mark in ERG⁻ LCD organoids and *Pten*^{-/-} organoids, suggesting this was a bona fide enhancer for *Trp63* in prostate cells (Figure 9B and Supplemental Figure 7B). Upon ERG expression in LCD-ERG organoids and *Pten*^{-/-} *R26^{ERG}* organoids, we did not observe ERG binding to the *Trp63* gene body, but to its distal enhancer (Figure 9B and Supplemental Figure 7B). Remarkably, there were significantly decreased H3K27ac signals at the distal enhancer, as well as chromatin interaction loops with the *Trp63* promoter upon the ERG binding site in *Pten*^{-/-} *R26^{ERG}* (Figure 9, B and C) and LCD-ERG organoids (Supplemental Figure 7, B and C). These results indicate a functional link between ERG-directed rewiring of chromatin interactions and epigenetic modifications to regulate gene expression.

To further determine whether ERG could directly repress *Trp63* expression through the distal binding-induced attenuations on chromatin interactions, we used the CRISPR/Cas9 system to specifically delete this ERG binding site in *Pten*^{-/-} *R26^{ERG}* and LCD-ERG organoids. The results of Sanger sequencing confirmed the successful heterozygous KO of the ERG binding site in *Pten*^{-/-} *R26^{ERG}* organoid clones (Supplemental Figure 7D). Notably, both qRT-PCR and Western blotting assays revealed that deletion of the ERG binding site (EB-KO) resulted in the increased expression of *Trp63* and the decreased expression of *Krt8* and *Krt18* in *Pten*^{-/-} *R26^{ERG}* organoids (Figure 9D and Supplemental Figure 7E) and LCD-ERG organoids (Supplemental Figure 7, J and K; see complete unedited blots in the supplemental material). To further characterize the global changes in prostate lineage induced by EB-KO, we compared 3 independent EB-KO organoid clones with control *Pten*^{-/-} *R26^{ERG}* organoid clones using ERG-regulating prostate lineage genes. GSEA demonstrated that EB-KO was significantly associated with the reduced expression of ERG-upregulating luminal signature genes and the increased expression of ERG-downregulating basal signature genes (Figure 9, E and F). Furthermore, PCA analysis revealed the distinct relationships among EB-KO, ERG-KO, and the control in each of their *Pten*^{-/-} *R26^{ERG}* organoids. Intriguingly, the ERG control showed a closer relationship with EB-KO than with ERG-KO, suggesting that EB-KO could partially phenocopy the biological effects of ERG-KO (Supplemental Figure 7F). Given the effects of EB-KO on lineage changes in vitro, we next sought to investigate the effects of EB-KO in vivo using UGSM tissue recombination assays. Remarkably, the outer layer with *Trp63*⁺ERG⁺ and *Krt5*⁺ERG⁺ basal cells could be widely identified in EB-KO-derived allografts, indicating the EB-KO-induced differentiation of prostate basal lineage in vivo (Figure 9, G and H, and Supplemental Figure 7, G-I).

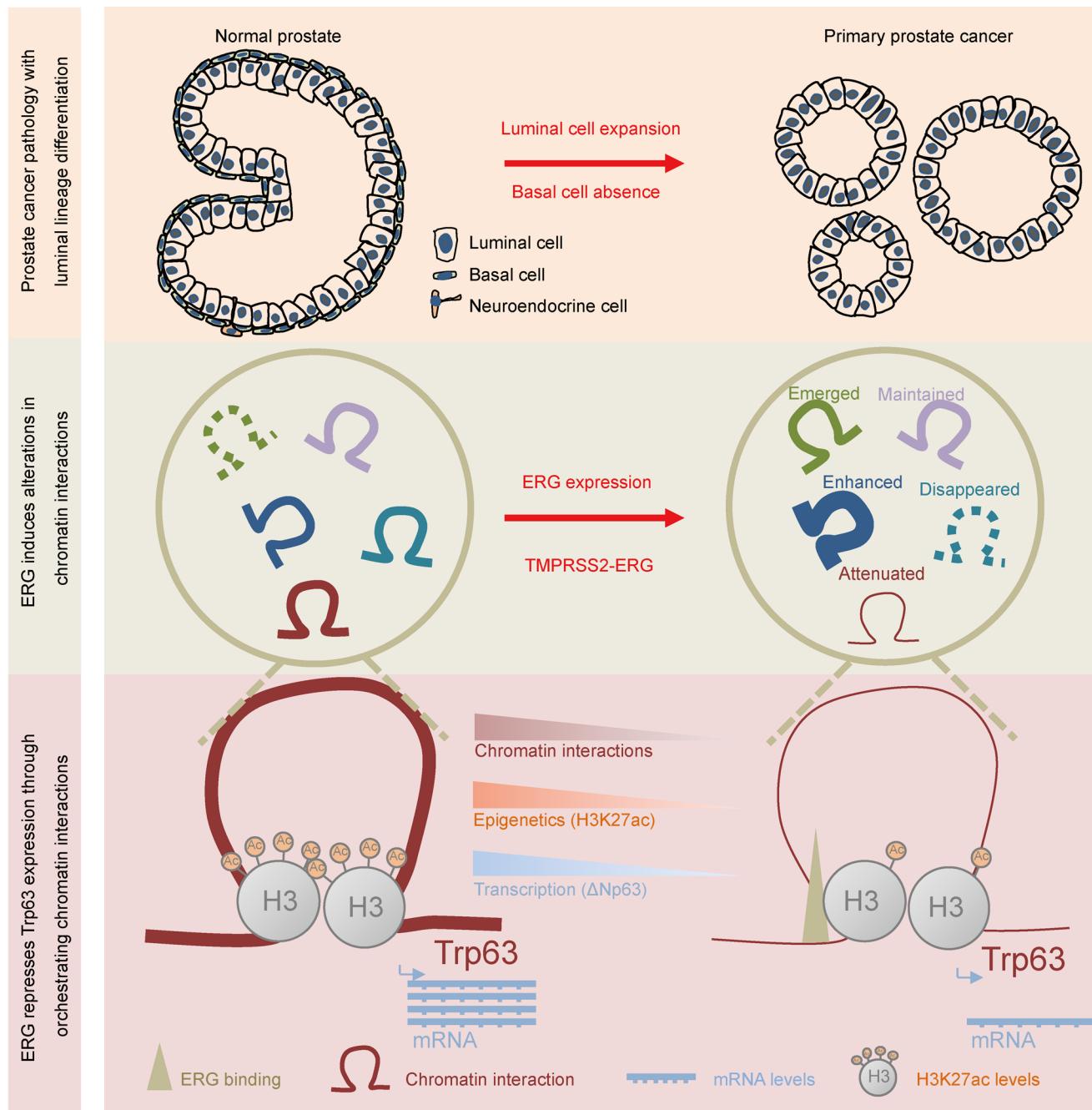


Figure 10. Schematic diagram of how ERG drives prostate cell fate reprogramming through orchestration of chromatin interactions. Most prostate cancers are characterized by luminal cell expansion and basal cell absence, compared with normal prostate architecture composed of both luminal cells and basal cells (top). ERG overexpression driven by TMPRSS2-ERG fusion is one of the most common genetic alteration events in prostate cancer, which can alter chromatin interactions (middle). Since chromatin architecture is closely associated with epigenetic modifications and mRNA transcription, ERG-induced alterations in chromatin interactions may cause dysregulation of genes including Trp63. ERG overexpression reduces chromatin interactions and H3K27ac levels across the region from a distal ERG binding site to the Trp63 gene body, which further causes decreased mRNA levels of Trp63 to facilitate the function of ERG in promoting luminal lineage differentiation (bottom).

To validate the existence of the distal ERG binding site in human prostate cells, we analyzed a data set that was previously generated from RWPE-1 cells with ERG overexpression (59). Remarkably, we found the actual existence of the distal ERG binding site in ERG-expressing RWPE-1 cells (Supplemental Figure 8A). Moreover, the homologies for these binding sites between

human and mouse prostate cells were also confirmed by additional analyses using NCBI BLAST tools (Supplemental Figure 8B). We next sought to characterize the lineage changes induced by ERG expression in human prostate cells. Consistent with the results found in mouse prostate cells, ERG expression resulted in the enhanced luminal phenotype with the increased expression of

KRT8 and *KRT18*, and attenuated the basal phenotype indicated by the reduced expression of *TP63*, *KRT5*, and *KRT14* (Supplemental Figure 8, C–E).

To provide more evidence for ERG in stimulating prostate cell luminal differentiation, we analyzed another publicly available data set, including LNCaP with ERG overexpression and VCaP with ERG knockdown (60). In LNCaP cells, ERG overexpression significantly enhanced the expression of canonical luminal makers *KRT8* and *KRT18* (Supplemental Figure 8F). Moreover, ERG overexpression significantly promoted luminal differentiation on the basis of ERG-upregulating luminal signature genes (Supplemental Figure 8G). In accordance with results from LNCaP cells, ERG knockdown in VCaP cells reduced *KRT8* and *KRT18* expression as well as luminal activity (Supplemental Figure 8, H and I).

In summary, our above results demonstrated the function of the distal ERG binding site in ERG-mediated maintenance and regulation of prostate luminal cell features, reflecting that ERG orchestrates the plasticity of prostate luminal lineage through chromatin interactions. In addition, the existence of the distal ERG binding site and the function of ERG in promoting luminal differentiation in human prostate cells revealed a conserved role of ERG in prostate luminal lineage regulation.

Discussion

Definitive evidence collected in previous studies supports the close associations between the activity of TFs and cell lineage determination in various biological processes, including development, immune response, and cancer progression (61–64). Particularly, primary prostate cancer is characterized by luminal cell expansion and loss of basal cells. Therapeutic treatments for prostate cancers can select for lineage alterations with the transitions from luminal cell lineage toward neuroendocrine and basal differentiation. Numerous studies have focused on lineage transitions in castration-resistant prostate cancer. However, the lineage-determining mechanisms of primary prostate luminal cancers are still largely unknown. Here, we have successfully identified ERG as a master regulator in regulating prostate cancer cell luminal lineage through chromatin interaction changes.

TMPRSS2-ERG fusion is a common genetic alteration event (~50%), which drives ERG expression occurring in the early stage of prostate cancer (30). We identified ERG as a master regulator in prostate cancer lineage regulation through the integrating analysis of 3 high-quality human prostate cancer cohorts (Figure 1C). It is widely accepted that prostate basal and luminal cells have bipotential plasticity, which was found in 3D organoids and the UGSM tissue recombination assay (12, 16). In this study, we found that ERG expression strongly facilitated the differentiations toward the luminal phenotype in luminal organoids and basal organoids, consistent with previous findings that ERG expressions induced a significant decrease in the proportion of prostate basal cells (65, 66). Moreover, our current study indicated that luminal cells tended to be more liable for lineage regulation conducted by ERG compared with basal cells. Together with their clinical relevance, our findings suggest the important role of ERG in initiation of primary prostate cancer with luminal cell features.

Previous studies have provided some insights into the functional role of AR in cell lineage regulation in normal prostate development and in prostate cancer. *In vivo* tissue recombination modeling suggests that stromal AR, but not epithelial AR, is essential for prostate developmental growth and morphogenesis (67, 68). Consistent with these findings, recent mouse lineage-tracing studies have demonstrated that in the adult prostate, specific AR deletion in luminal cells has little effect on luminal cell differentiation (51). As for prostate cancer, highly analogous to the previous findings that prostate tumors with AR knockout were characterized by luminal features in mouse models of *Pb-Cre*⁺ *Pten*^{f/f} *AR*^{f/Y} (52) and *Nkx3.1*^{CreERT2/+} *AR*^{f/Y} *Pten*^{f/f} *R26*^{YFP/+} (51, 55), our results (Figure 6, A and C) indicate that the luminal lineage differentiations for prostate cancer cells are not directly dependent on AR expression, providing the insight that ERG can directly determine the prostate cancer cell luminal lineage through changes in global chromatin interactions. Consistently, deletions of ERG or the ERG-specific binding site disrupted the prostate luminal lineage, leading to the differentiation of prostate basal cell lineage (Figure 6, A and C, and Figure 9). However, there are also limitations to our study. Compared with human prostate cancer, ERG expression in our murine prostate cancer models was driven by the promoter of the *Rosa26* locus; therefore, models with other forms of ERG expression should be further developed.

Essentially all localized, untreated prostate cancers exhibit a luminal phenotype, suggesting that prostate cancer tumor-initiating genetic lesions may inhibit the bipotential plasticity programs in normal prostate epithelial cells. In addition to TMPRSS2-ERG fusion in 50% of primary prostate cancers, other prostate cancer-specific mutations, such as ERF, FOXA1, and SPOP, have been reported to regulate prostate cell luminal differentiation in prostate cancer. ERF is a suppressive ETS factor and exhibits deleterious point mutations and focal deletions in prostate cancer (69, 70). ERF loss shares some functional consequences with ERG overexpression, including luminal morphology regulation, expanded activity of AR, and oncogenic ability in the setting of *Pten* loss (69). Notably, FOXA1 mutations were recently found in 41% of Asian patients with prostate cancer, an ethnic group with low frequency of ETS rearrangements (71, 72). In addition to coding mutations, the FOXA1 locus is involved in duplications and translocations that likely result in overexpression of the WT protein (71, 73). WT FOXA1 and 12 of the FOXA1 mutations have been demonstrated to have the function of promoting prostate cell luminal differentiation (46). FOXA1 was also included in the identified 154 master TFs in our present study, which was consistent with its pioneer function in lineage regulation for prostate cancer. SPOP is mutated in approximately 10% of prostate cancers, and SPOP mutants may play an important role in suppressing basal cell differentiation in ERG prostate tumors (74, 75). Further investigations to dissect cancer stage-specific roles of luminal cell AR in primary prostate cancer and advanced prostate cancer will be necessary.

PTEN deletion is another common genetic alteration event in primary prostate cancer (31, 76). It was demonstrated that *Pten* deletion led to basal differentiation, validated by a significant increase of *Krt5*⁺/*Trp63*⁺ cells with disease progression (77). Consistently, in comparison with WT luminal organoids, our *Pten* null organoids also exhibited basal differentiation (Figure 4A). This

could explain the clinical relevance that PTEN loss significantly occurs with ERG fusion, which may facilitate cancer cells to maintain both proliferation capacity and luminal characteristics. Gradually increased incidents of PTEN gene deletion and PI3K signaling pathway activation were identified during prostate cancer progression (49, 50, 78). Therefore, our study demonstrates a possible molecular mechanism underlying the basal lineage plasticity in advanced prostate cancer.

TMPRSS2-ERG translocation represents a distinct subset of the *cis*-regulatory landscape in primary prostate tumors (39). ERG overexpression is known to induce global changes in chromatin conformation (79). Here, we have further proved that ERG overexpression globally induces chromatin interaction changes (Figure 7, A and B, and Supplemental Figure 6, C and D). Moreover, these chromatin interaction changes are associated with coordinated DEG expressions. Through binding site at a distal regulatory element of Trp63, ERG can regulate Trp63 expression by chromatin interactions. Importantly, deletion of this binding site remarkably reverses the lineage plasticity toward basal differentiation. Compelling data in support of this hypothesis has also been obtained from the reanalysis of the publicly available human data sets with ERG ChIP-Seq, which can validate the conserved existence of the ERG binding site in human prostate cells (59). Therefore, we have identified the conserved ERG binding site that contributes to prostate lineage plasticity. In addition, we have also provided a potentially novel research paradigm for the investigation of how TFs regulate their responsive genes through chromatin interactions instead of direct binding at the gene body regions.

Here, ERG was identified as a master TF to manipulate plasticity in prostate cell lineage differentiation toward proluminal programming through chromatin interactions. Our findings propose a potentially novel working model, which elucidated the detailed mechanisms of ERG's role as a master TF in order to pursue the fundamental, longstanding goal of determining how prostate cancer cells actively maintain luminal lineage identities. In addition, our study provides further supporting research on the role of lineage plasticity in prostate cancer initiation.

Methods

Analysis pipeline for identification of the master TFs. Prostate cancer cohorts with more than 100 samples and RNA expression profiles were selected for downstream analysis. After prostate cancer cohort filtering, further cancer subtyping was performed based on the 3 selected cohorts: FHCRC (158 samples), MSKCC (150 samples), and TCGA (498 samples). For each cohort, the integrative classifier was performed to identify the TFs that correlated with epigenetic modifications, and the PAM50 classifier was performed to define prostate cancer lineage-related TFs. In detail, to identify the potential of a TF as master regulator, samples of every cohort were first divided into 3 groups according to TF expression levels, termed as TF-high, TF-medium, and TF-low. Meanwhile, samples were also categorized into another 3 groups using the PAM50 classifier or integrative classifier using their own marker genes. The classifier was performed originally based on the risk model that incorporates the gene expression-based "intrinsic" subtypes (43). We downloaded source code from the University of North Carolina Microarray Database (GSE10886; <https://genome.unc.edu/pub>

sup/breastGEO) (80). We excluded the normal-like subtype and HER2 subtype, similar to a previous study (42). The integrative classifier was performed based on 285 genes using unsupervised hierarchical clustering (41). Next, Pearson's χ^2 test was performed to evaluate the correlation between TF expression levels and subtypes, and overlapped TFs in the integrative classifier and PAM50 classifier were defined as overlapped TFs. Taking confidence into consideration, overlapped TFs that occurred in at least 2 cohorts were defined as master TFs for further study. To visualize the significance of each master TF in all of the 3 cohorts, we created a bubble plot based on the transformed *P* value of each cohort [*P* value = $(p_{PAM} \times p_{INT})^{1/2}$, p_{PAM} was calculated by Pearson's χ^2 test based on the PAM50 classifier, p_{INT} was calculated by Pearson's χ^2 test based on the integrative classifier].

Gene targeting and mouse breeding. Mice were bred and maintained according to Shanghai Laboratory Animal Center Institutional Animal Regulations. *Tmprss2*^{CreERT2/+} (47), *Pten*^{fl/fl} (81), *Rosa26*^{EVYFP/EVYFP} (82), *Rosa26*^{ERG/ERG} (38), and *Pb-Cre4* (83) mice were previously described and all in the C57/B6 background. The *Tmprss2*^{CreERT2/+} *Rosa26*^{EVYFP/EVYFP} (T2Y), *Tmprss2*^{CreERT2/+} *Rosa26*^{EVYFP/ERG} (T2YE), *Tmprss2*^{CreERT2/+} *Pten*^{fl/fl} *Rosa26*^{ERG/ERG} (T2PE), *Rosa26*^{ERG/ERG}, *Tmprss2-ERG* knockin and *Pb-Cre4 Rosa26*^{ERG/ERG} mice were generated through standard mouse breeding within the Shanghai Institute of Biochemistry and Cell Biology (SIBCB) animal facility.

Mouse procedures. Mice were raised in accordance with the recommendations of MSKCC veterinary services and the MSKCC IACUC (11-12-029). For tamoxifen (Toronto Research Chemicals) treatment of T2Y, T2YE, and T2PE mice, tamoxifen was dissolved in 20 mg/mL corn oil and i.p. injected into 8-week-old mice at a dose of 3 mg every other day for 3 doses. Mice were euthanized according to the indicated timeline after the first tamoxifen dose.

RNA-Seq data processing and analysis. RNA sequencing libraries were prepared with the VAHTS mRNA-Seq V3 Library Prep Kit for Illumina (Vazyme, NR611). Sequencing was performed by Berry Genomics. Low-quality sequences and adapters were filtered by cutadapt-1.15. Clean reads were mapped to the mm9 genome using hisat2-2.1.0 (84). Gene expression was quantified at the gene level using featureCounts (85). DEGs were analyzed by DESeq2 (86) using raw counts. An adjusted *P* value of less than 0.05 was set as the threshold to define DEGs. GSEA (87) was conducted to determine statistically significant defined signatures based on the normalized expression value calculated from DESeq2. Enriched pathway analysis was performed using Metascape (88).

ATAC-Seq library preparation. First, 50,000 cells were collected by centrifuging and washed with ice-cold PBS. Cells were lysed in 50 μ L ice-cold lysis buffer (10 mM pH7.4 Tris-HCl; 10 mM NaCl; 3 mM MgCl₂; 0.5% NP-40) for 10 minutes on ice. Immediately after lysis, nuclei were spun at 500 \times g for 5 minutes using a refrigerated centrifuge at 4 degrees. The steps to generate sequencing libraries were performed with TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme, TD501).

ATAC-Seq data processing and analysis. We used Partek Genomics Suite to map sequencing reads and removed duplicate reads to mouse reference genome mm9. Peaks were identified using HOTSPOT with default parameters (<http://www.uwencode.org/proj/hotspot/>). HOTSPOT analysis generates 2 types of peaks: the narrow peak and hotspot region (broad peak). In this study, we used the narrow peak for downstream analysis.

Quantification of chromatin accessibility of ATAC-Seq. We referred to the narrow peak as regulatory element (RE). ATAC-Seq will measure the accessibility in a given RE. We quantified the openness of the RE by a simple fold change score, which computes the enrichment of read counts by comparing the RE with a large background region. Briefly, let N be the number of reads in RE of length L and G be the number of reads in the background window of length W (1 Mb in our case) around this RE. The openness of RE O is defined as follows: $O = (N/L)/(G/W)$.

Differential ATAC-Seq peak analysis. Differential ATAC-Seq peak analysis was performed by comparing LCD organoids with LCD-ERG organoids, as well as by comparing *Pten*^{-/-} organoids with *Pten*^{-/-} *R26^{ERG}* organoids. We defined the sample specific peaks with a greater than 1.5 fold-change and an openness value greater than 1.

ChIP-Seq library preparation. The libraries of ChIP-Seq were prepared using protocols as previously described (38). Descriptions of additional histology procedures are in the Supplemental Methods. Sequencing libraries were prepared from TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme, TD503) using the manufacturer's recommended protocol.

ChIP-Seq data processing and analysis. The ChIP-Seq pipeline was based on the ENCODE (phase-3) TF and histone ChIP-Seq pipeline specifications (by Anshul Kundaje) (<https://github.com/ENCODE-DCC/chip-seq-pipeline2>). Sequencing reads were mapped to mm9 using bwa (89) with parameters *aln -q 5 -l 32 -k 2*. Unmapped, multi-mapping, low-quality reads, and duplicates were filtered by SAMtools (90). Peaks were identified using MACS2 (91) with an FDR of 0.05.

BL-Hi-C assay. The libraries of BL-Hi-C were generated using the 2-step ligation protocols as previously described (56). Descriptions of additional histology procedures are in the Supplemental Methods. Library construction for sequencing was conducted with VAHTS Universal DNA Library Prep Kit for Illumina (Vazyme, ND607). The BL-Hi-C library was sequenced with the Illumina Sequencer NovaSeq (PE 2 × 150 bp reads).

BL-Hi-C data processing. We first trimmed the linkers of BL-Hi-C sequence using the trimLinker function of CHIA-PET2 (92). HiC-Pro (93) was then performed to process Hi-C data through several main steps, including mapping raw reads to the mm9 reference genome, detecting valid ligation products, performing quality control, and generating raw contact maps. To evaluate the library quality, we first removed duplicated reads, and then divided the remaining valid reads into several groups, including cis long-range (> 200k), cis short-range (< 200k), and trans contacts. Each group was indicated by a different color. Our BL-Hi-C libraries required approximately 40% of the long-range valid pairs to ensure high quality (93). For raw contact matrix generation, we set a variety of resolutions, including 10 kb, 20 kb, 40 kb, 150 kb, 500 kb, and 1 Mb, which denoted that the genome was divided into bins with the above equal sizes. Homer (94) was performed to further identify significant interactions (loops) (FDR < 0.0001) based on contact maps using 10 k resolution.

Identification of differential chromatin interactions. We performed differential chromatin interactions by comparing LCD organoids with LCD-ERG organoids, as well as by comparing *Pten*^{-/-} organoids with *Pten*^{-/-} *R26^{ERG}* organoids. We referred to the set of significant interactions identified by Homer in each sample as its loop set, and differential loops were defined by the difference of 2 sets.

Identification of the relationship between chromatin interactions and ERG binding. To characterize the relationship between chromatin interaction and ERG binding, we divided the genome into 1 Mb bins and sorted the genomic bins by the number of *Pten*^{-/-} *R26^{ERG}*/LCD-ERG-specific loops. We next performed Pearson correlation analysis based on the number of *Pten*^{-/-} *R26^{ERG}*/LCD-ERG-specific loops and ERG binding density, generated from BL-Hi-C and ERG ChIP-Seq data, respectively. We also calculated the average number of ERG binding sites for the first N bins with *Pten*^{-/-} *R26^{ERG}*/LCD-ERG-specific loops (N ranges 20 to 2,779). The results showed a strong association between the hotspots of differential chromatin interactions and enrichment of ERG binding.

Accession numbers. The raw data for RNA-Seq, ChIP-Seq, BL-Hi-C, and processed data for ATAC-Seq have been deposited in NODE (<http://www.biosino.org/node>). All data can be viewed by pasting the accession number (OEP000693) into the text search box or through the URL: <http://www.biosino.org/node/project/detail/OEP000693>, including ATAC-Seq data (OEX002111), ChIP-Seq data (OEX002110), RNA-Seq data (OEX002109), and BL-Hi-C data (OEX002216).

Statistics. For significance tests, 2-tailed Student's *t* test and Pearson's χ^2 test were used for comparing differences between 2 groups, and 1-way ANOVA tests were used for multiple groups. *P* values of less than 0.05 were considered significant.

Study approval. All animal experiments were approved by the IACUC of the Center for Excellence in Molecular Cell Science (Shanghai, China).

Author contributions

DG and YC conceived and designed the experimental approach. YW, BC, and PC provided advice about experimental design. FL, WD, and XYX performed most of the experiments. QYY, FL, LL, and YW contributed to the computational analysis and statistical analysis. CFL and JH generated the expression vectors and lentiviral CRISPR/Cas9 vectors. ZL, NHM, YGL, WXG, and SQW prepared mouse organoid RNA and sequencing. XYZ, ZL, YQZ, and RA helped with allograft experiments and mouse experiments. DG, FL, and YC wrote the manuscript. All authors discussed the results and edited the manuscript.

Acknowledgments

We would like to acknowledge Baojin Wu, Guoyuan Chen, and Wei Tang for the animal husbandry and Wei Bian for technical help at the SIBCB Core Facility. We would like to thank the Genome Tagging Project (GTP) Center, SIBCB, CAS for technical support (Shanghai, China). We thank the MSKCC Molecular Cytology (Ning Fan, Mesruh Turkekul, Sho Fujisawa), MSKCC Integrated Genomics Operation (Daoqi You, Agnes Viales), and MSKCC Epigenomics Core Facility (Yang Li) (New York, USA). We thank Wilbert Zwart and Suzan Steloo at the Netherlands Cancer Institute for sharing analysis methods of the integrative classifier (Amsterdam, Netherlands). This study was supported by grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB19000000 and XDA16020905), the Basic Frontier Science Research Program of the Chinese Academy of Sciences (ZDBS-LY-SM015), the National Key Research and Development Program of China (No. 2017YFA0505500, 2017YFC0908400, and 2018ZX10302207-00*-001), the National Natural Science Foundation of China (81830054, 81772723, 11871463, and 61621003), the Prostate Can-

cer Foundation, the Starr Cancer Consortium, Geoffrey Beene Cancer Research Center, the Gerstner Family Foundation, and the US National Cancer Institute (K08CA140946, R01CA208100, R01CA193837, P50CA092629 and P30CA008748).

Address correspondence to: Dong Gao, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, 320 YueY-

ang Road, Shanghai 200031, China. Phone: 86.21.54921117; Email: dong.gao@sibcb.ac.cn. Or to: Yu Chen, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, New York 10065, USA. Phone: 646.888.3356; Email: cheny1@mskcc.org. Or to: Yong Wang, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 55 Zhongguancun East Road, Beijing 100190, China. Phone: 86.10.82541372; Email: ywang@amss.ac.cn.

- Varga J, Greten FR. Cell plasticity in epithelial homeostasis and tumorigenesis. *Nat Cell Biol.* 2017;19(10):1133–1141.
- Restivo G, et al. Low neurotrophin receptor CD271 regulates phenotype switching in melanoma. *Nat Commun.* 2017;8(1):1988.
- Davies AH, Beltran H, Zoubeidi A. Cellular plasticity and the neuroendocrine phenotype in prostate cancer. *Nat Rev Urol.* 2018;15(5):271–286.
- Arozarena I, Wellbrock C. Phenotype plasticity as enabler of melanoma progression and therapy resistance. *Nat Rev Cancer.* 2019;19(7):377–391.
- Suva ML, Rigg N, Bernstein BE. Epigenetic reprogramming in cancer. *Science.* 2013;339(6127):1567–1570.
- Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 2006;126(4):663–676.
- Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell.* 2013;12(1):15–30.
- Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. *Cell.* 2016;166(1):21–45.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7–34.
- Zhang D, et al. Stem cell and neurogenic gene-expression profiles link prostate basal cells to aggressive prostate cancer. *Nat Commun.* 2016;7:10798.
- Gao D, et al. Organoid cultures derived from patients with advanced prostate cancer. *Cell.* 2014;159(1):176–187.
- Karthaus WR, et al. Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell.* 2014;159(1):163–175.
- Xin L, Lukacs RU, Lawson DA, Cheng D, Witte ON. Self-renewal and multilineage differentiation in vitro from murine prostate stem cells. *Stem Cells.* 2007;25(11):2760–2769.
- Garraway IP, et al. Human prostate sphere-forming cells represent a subset of basal epithelial cells capable of glandular regeneration in vivo. *Prostate.* 2010;70(5):491–501.
- Parsons JK, Gage WR, Nelson WG, De Marzo AM. p63 protein expression is rare in prostate adenocarcinoma: implications for cancer diagnosis and carcinogenesis. *Urology.* 2001;58(4):619–624.
- Chua CW, et al. Single luminal epithelial progenitors can generate prostate organoids in culture. *Nat Cell Biol.* 2014;16(10):951–961.
- Labrecque MP, et al. Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. *J Clin Invest.* 2019;130(10):4492–4505.
- Tan HL, et al. Rb loss is characteristic of prostatic small cell neuroendocrine carcinoma. *Clin Cancer Res.* 2014;20(4):890–903.
- Mosquera JM, et al. Concurrent AURKA and MYCN gene amplifications are harbingers of lethal treatment-related neuroendocrine prostate cancer. *Neoplasia.* 2013;15(1):1–10.
- Dardenne E, et al. N-Myc induces an EZH2-mediated transcriptional program driving neuroendocrine prostate cancer. *Cancer Cell.* 2016;30(4):563–577.
- Lee JK, et al. N-Myc drives neuroendocrine prostate cancer initiated from human prostate epithelial cells. *Cancer Cell.* 2016;29(4):536–547.
- Ku SY, et al. Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science.* 2017;355(6320):78–83.
- Mu P, et al. SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and Rb1-deficient prostate cancer. *Science.* 2017;355(6320):84–88.
- Zou M, et al. Transdifferentiation as a mechanism of treatment resistance in a mouse model of castration-resistant prostate cancer. *Cancer Discov.* 2017;7(7):736–749.
- Baltzinger M, Mager-Heckel AM, Remy P. Xl erg: expression pattern and overexpression during development plead for a role in endothelial cell differentiation. *Dev Dyn.* 1999;216(4–5):420–433.
- Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol.* 2001;2(11):827–837.
- Chi P, et al. ETV1 is a lineage survival factor that cooperates with KIT in gastrointestinal stromal tumours. *Nature.* 2010;467(7317):849–853.
- Barozzi I, et al. Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol Cell.* 2014;54(5):844–857.
- Morita R, et al. ETS transcription factor ETV2 directly converts human fibroblasts into functional endothelial cells. *Proc Natl Acad Sci U S A.* 2015;112(1):160–165.
- Tomlins SA, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science.* 2005;310(5748):644–648.
- Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell.* 2015;163(4):1011–1025.
- Leshem O, et al. TMPRSS2/ERG promotes epithelial to mesenchymal transition through the ZEB1/ZEB2 axis in a prostate cancer model. *PLoS One.* 2011;6(7):e21650.
- Adamo P, Ladomery MR. The oncogene ERG: a key factor in prostate cancer. *Oncogene.* 2016;35(4):403–414.
- Tomlins SA, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia.* 2008;10(2):177–188.
- Hollenhorst PC, Ferris MW, Hull MA, Chae H, Kim S, Graves BJ. Oncogenic ETS proteins mimic activated RAS/MAPK signaling in prostate cells. *Genes Dev.* 2011;25(20):2147–2157.
- Carver BS, et al. Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat Genet.* 2009;41(5):619–624.
- Yu J, et al. An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell.* 2010;17(5):443–454.
- Chen Y, et al. ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. *Nat Med.* 2013;19(8):1023–1029.
- Kron KJ, et al. TMPRSS2-ERG fusion co-opts master transcription factors and activates NOTCH signaling in primary prostate cancer. *Nat Genet.* 2017;49(9):1336–1345.
- Wasmuth EV, Hoover EA, Antar A, Klinge S, Chen Y, Sawyers CL. Modulation of androgen receptor DNA binding activity through direct interaction with the ETS transcription factor ERG. *Proc Natl Acad Sci U S A.* 2020;117(15):8584–8592.
- Stelloo S, et al. Integrative epigenetic taxonomy of primary prostate cancer. *Nat Commun.* 2018;9(1):4900.
- Zhao SG, et al. Associations of luminal and basal subtyping of prostate cancer with prognosis and response to androgen deprivation therapy. *JAMA Oncol.* 2017;3(12):1663–1672.
- Parker JS, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27(8):1160–1167.
- Kumar A, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat Med.* 2016;22(4):369–378.
- Taylor BS, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell.* 2010;18(1):11–22.
- Adams EJ, et al. FOXA1 mutations alter pioneering activity, differentiation and prostate cancer phenotypes. *Nature.* 2019;571(7765):408–412.
- Gao D, et al. A Tmprrs2-CreERT2 knock-in mouse model for cancer genetic studies on prostate and colon. *PLoS One.* 2016;11(8):e0161084.
- Xin L, Ide H, Kim Y, Dubey P, Witte ON. In vivo regeneration of murine prostate from dissociated cell populations of postnatal epithelia and urogenital sinus mesenchyme. *Proc Natl Acad Sci U S A.* 2003;100 Suppl 1:11896–11903.
- Han B, et al. Fluorescence in situ hybridization study shows association of PTEN deletion with ERG rearrangement during prostate cancer progression. *Mod Pathol.* 2009;22(8):1083–1093.
- Jamaspishvili T, et al. Clinical implications of

PTEN loss in prostate cancer. *Nat Rev Urol.* 2018;15(4):222–234.

51. Xie Q, Liu Y, Cai T, Horton C, Stefanson J, Wang ZA. Dissecting cell-type-specific roles of androgen receptor in prostate homeostasis and regeneration through lineage tracing. *Nat Commun.* 2017;8:14284.

52. Mulholland DJ, et al. Cell autonomous role of PTEN in regulating castration-resistant prostate cancer growth. *Cancer Cell.* 2011;19(6):792–804.

53. Stadhouders R, Filion GJ, Graf T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature.* 2019;569(7756):345–354.

54. Cuvier O, Fierz B. Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells. *Nat Rev Genet.* 2017;18(8):457–472.

55. Chua CW, et al. Differential requirements of androgen receptor in luminal progenitors during prostate regeneration and tumor initiation. *Elife.* 2018;7:e28768.

56. Liang Z, et al. BL-Hi-C is an efficient and sensitive approach for capturing structural and regulatory chromatin interactions. *Nat Commun.* 2017;8(1):1622.

57. Signoretti S, et al. p63 is a prostate basal cell marker and is required for prostate development. *Am J Pathol.* 2000;157(6):1769–1775.

58. Kurita T, Medina RT, Mills AA, Cunha GR. Role of p63 and basal cells in the prostate. *Development.* 2004;131(20):4955–4964.

59. Kedage V, et al. An interaction with Ewing's sarcoma breakpoint protein EWS defines a specific oncogenic mechanism of ETS factors rearranged in prostate cancer. *Cell Rep.* 2016;17(5):1289–1301.

60. Sandoval GJ, et al. Binding of TMPRSS2-ERG to BAF chromatin remodeling complexes mediates prostate oncogenesis. *Mol Cell.* 2018;71(4):554–566.

61. Rothenberg EV, Moore JE, Yui MA. Launching the T-cell-lineage developmental programme. *Nat Rev Immunol.* 2008;8(1):9–21.

62. Yui MA, Rothenberg EV. Developmental gene networks: a triathlon on the course to T cell identity. *Nat Rev Immunol.* 2014;14(8):529–545.

63. Bushweller JH. Targeting transcription factors in cancer – from undruggable to reality. *Nat Rev Cancer.* 2019;19(11):611–624.

64. Hemberger M, Hanna CW, Dean W. Mechanisms of early placental development in mouse and humans. *Nat Rev Genet.* 2020;21(1):27–43.

65. Klezovitch O, et al. A causal role for ERG in neo- plastic transformation of prostate epithelium. *Proc Natl Acad Sci USA.* 2008;105(6):2105–2110.

66. Blee AM, et al. TMPRSS2-ERG controls luminal epithelial lineage and antiandrogen sensitivity in PTEN and TP53-mutated prostate cancer. *Clin Cancer Res.* 2018;24(18):4551–4565.

67. Wang Y, et al. A human prostatic epithelial model of hormonal carcinogenesis. *Cancer Res.* 2001;61(16):6064–6072.

68. Marker PC, Donjacour AA, Dahiya R, Cunha GR. Hormonal, cellular, and molecular control of prostatic development. *Dev Biol.* 2003;253(2):165–174.

69. Bose R, et al. ERF mutations reveal a balance of ETS factors controlling prostate oncogenesis. *Nature.* 2017;546(7660):671–675.

70. Huang FW, et al. Exome sequencing of African-American prostate cancer reveals loss-of-function ERF mutations. *Cancer Discov.* 2017;7(9):973–983.

71. Parolia A, et al. Distinct structural classes of activating FOXA1 alterations in advanced prostate cancer. *Nature.* 2019;571(7765):413–418.

72. Li J, et al. A genomic and epigenomic atlas of prostate cancer in Asian populations. *Nature.* 2020;580(7801):93–99.

73. Zhou S, et al. Noncoding mutations target cis-regulatory elements of the FOXA1 plexus in prostate cancer. *Nat Commun.* 2020;11(1):441.

74. Gan W, et al. SPOP promotes ubiquitination and degradation of the ERG oncoprotein to suppress prostate cancer progression. *Mol Cell.* 2015;59(6):917–930.

75. Barbieri CE, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet.* 2012;44(6):685–689.

76. Berger MF, et al. The genomic complexity of primary human prostate cancer. *Nature.* 2011;470(7333):214–220.

77. Wang S, Garcia AJ, Wu M, Lawson DA, Witte ON, Wu H. Pten deletion leads to the expansion of a prostatic stem/progenitor cell subpopulation and tumor initiation. *Proc Natl Acad Sci USA.* 2006;103(5):1480–1485.

78. Grasso CS, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature.* 2012;487(7406):239–243.

79. Rickman DS, et al. Oncogene-mediated alterations in chromatin conformation. *Proc Natl Acad Sci USA.* 2012;109(23):9083–9088.

80. Parker JS, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27(8):1160–1167.

81. Carver BS, et al. Reciprocal feedback regulation of PI3K and androgen receptor signaling in PTEN-deficient prostate cancer. *Cancer Cell.* 2011;19(5):575–586.

82. Srinivas S, et al. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol.* 2001;1:4.

83. Wu X, et al. Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. *Mech Dev.* 2001;101(1-2):61–69.

84. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–360.

85. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–930.

86. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

87. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545–15550.

88. Zhou Y, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun.* 2019;10(1):1523.

89. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–1760.

90. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–2079.

91. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.

92. Li G, Chen Y, Snyder MP, Zhang MQ. ChIA-PET2: a versatile and flexible pipeline for ChIA-PET data analysis. *Nucleic Acids Res.* 2017;45(1):e4.

93. Servant N, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 2015;16:259.

94. Heinz S, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell.* 2010;38(4):576–589.