SUPPLEMENTAL MATERIALS AND METHODS

Isolation and culture of mouse prostate epithelial cells. Mouse prostates were dissected and minced with scissors, then digested with Collagenase/Hyaluronidase (STEMCELL, 07912) for 30 minutes with every 10-minute shaking in cell incubator. And subsequently digested with TrypLE (GIBCO, 12605-028) for another 15 minutes with shaking. DMEM (GIBCO, C11995500BT) with 10% FBS was added to stop the digestion and then cells were centrifuged for collection. Cells were cultured and processed as described previously (12, S1).

Flow cytometry and fluorescence-activated cell sorting (FACS). Flow cytometry cell sorting and analysis of mouse prostate cells were performed on an Aria II (BD Biosciences). Single cell suspensions of T2Y mice anterior prostates were stained using CD49f-PE (eBioscience, eBioGoH3, 1:200) and DAPI (4',6-diamidino-2-phenylindole). *Rosa26*$^{ERG/ERG}$ mice prostate cells were stained using a CD49f-PE (eBioscience, eBioGoH3, 1:200), Cd24-FITC (biolegend, M1/69, 1:200) and DAPI. Sorted murine prostate basal cells and luminal cells were fixed in Foxp3 Fixation/Permeabilization buffer (eBioscience, 00-5521-00) and stained with CK5- Alexa Fluor® 488 antibody (Abcam, ab193894, 1:500) or CK8+18- Alexa Fluor® 488 antibody (Abcam, ab192467, 1:500).

UGSM tissue recombination assays. UGSM tissue recombination assays were performed as described previously (48). Briefly, urogenital sinus mesenchyme (UGSM) cells were dissociated from urogenital sinus of day 18 rat embryos. Added 2-4mL 1% trypsin to the 3.5 cm dish and digested around 20 minutes at 4 degrees until the UGSM was fluffy. Wash the UGSM with DMEM with 10% FBS for twice. Transfer the UGSM cells to 4-6mL 0.1% collagenase B with 1%DNase and incubated at 37 degrees. Shaking vigorously every 10 minutes and carefully collected supernatant, then filtered cell suspension through 70μm strainer. Wash the UGM cells with DMEM with 10% FBS for twice to remove collagenase B. Mix mouse prostate organoids with UGM cells at appropriate ratio. Resuspend the cell mix with collagen for culture overnight. The cell pellets were transplanted under renal capsule of 6-week old SCID mice. The grafts were collected and analyzed as indicated timeline after transplantation.

Lentiviral CRISPR/Cas9-mediated knock-out. To knock out ERG and AR in mice organoids, we designed three pairs of single guide RNA (sgRNA) sequences for human ERG and mouse AR using the design tool from the Feng Zhang Lab (MIT) and cloned the targeting sequences into the

LentiCRISPRv2 vector obtained from Addgene. Lentiviruses for ERG sgRNAs, AR sgRNAs or vector control were generated in 293T cells by standard methods using lentiviral packaging vector. Mice and human prostate cells were infected with lentivirus for 48 hours and selected with 2 μg/mL puromycin for 7 days. ERG protein level and organoids histology were analyzed 21 days after infection. To knock out the distal ERG-binding site with the length of 880 bp, four single guide RNAs were designed including two upstream sgRNAs and two downstream sgRNAs which were subsequently cloned into codon-optimized SpCas9 plasmids PX330-RFP and PM458-GFP (derivative of PX330 backbone) respectively. Transient transfection was conducted to transfect above plasmids into LCD-ERG organoids and *Pten*$^{-/-}$*; R26*^{ERG} organoids with X-tremeGENE™ 9 transfection reagent (Roche, 6365779001). After 2 days, GFP/RFP double positive cells were sorted into 96-well plate using FACS. Knock-out efficiency was identified by PCR with 3 EB-KO-identify primers and Sanger sequencing. The target guide sequences and EB-KO-identify primers are listed as followed:

sgERG-1-F: CACCGACACCGTTGGGATGAACTA;

sgERG-1-R: AAACTAGTTCATCCCAACGGTGTC;

sgERG-2-F: CACCGTTCCTTCCCATCGATGTTC;

sgERG-2-R: AAACGAACATCGATGGGAAGGAAC;

sgERG-3-F: CACCGTACAGACCATGTGCGGCAG;

sgERG-3-R: AAACCTGCCGCACATGGTCTGTAC;

sgAR-1-F: CACCGGTGGAAAGTAATAGTCGAT;

sgAR-1-R: AAACATCGACTATTACTTTCCACC;

sgAR-2-F: CACCGCACTACGGAGCTCTCACTTG;

sgAR-2-R: AAACCAAGTGAGAGCTCCGTAGTGC;

sgEB-1-F: CACCGATATAGCACCTCGGTTCCCA;

sgEB-1-R: AAACTGGGAACCGAGGTGCTATATC;

sgEB-2-F: CACCGGTGGAAGAGGCATCGAATAG;

sgEB-2-R: AAACCTATTCGATGCCTCTTCCACC;

sgEB-3-F: CACCGATGTGATGCCTTCAGGCACG;

sgEB-3-R: AAACCGTGCCTGAAGGCATCACATC;

sgEB-4-F: CACCGCTGGGAACCGAGGTGCTATA;

sgEB-4-R: AAACTATAGCACCTCGGTTCCCAGC;

EB-KO-identify-F: TTGACAATATTGGAATTAGACGATAT;

EB-KO-identify-R1: AGTCACTCATGAGCAGCGTC;

EB-KO-identify-R2: ACAACAACTTGACCGTGTGG;

sgControl-F: CACCGGGCGAGGAGCTGTTCACCG;

sgControl-R: AAACCGGTGAACAGCTCCTCGCCC;

Stable gene expression. cDNAs for human prostate cancer *TMPRSS2-ERG* fusion was cloned into retroviral-based vector MSCV-C-HA (Addgene). Retrovirus was produced in 293T cells by standard methods using Ampho packaging vector. Prostate organoids were infected and selected with 2 μg/mL puromycin for 7 days at 48 hours after infection for subsequent histology and graft studies.

Immunohistochemistry. Organoids were fixed in 4% paraformaldehyde (Electron Microscopy Sciences) for 15 minutes at room temperature. Mouse prostates and mouse organoid grafts derived from UGSM tissue recombination assays were fixed using 4% paraformaldehyde overnight at 4 degrees. Organoids and tissues were processed for paraffin embedding using Leica ASP6025 tissue processor (Leica Biosystems). Freshly cut 5 microns paraffin sections were stained on Leica Bond RX (Leica Biosystems) with appropriate negative and positive controls. The following antibodies were diluted in Antibody Diluent (Leica) as indicated: ERG (Abcam, ab92513, 1:100); p63 (Abcam , ab735, 1:500); CK5 (Covance, PRB-160P, 1:2,000); CK8 (Covance, cat# MMS-162P, 1:1,000); PTEN (Cell Signaling Technology, 9188, 1:100); GFP (Abcam, 13970, 1:200); phospho-AKT (Ser473) (Cell Signaling Technology, 4060, 1:50); HA (Cell Signaling Technology, cat# 3724, 1:200); Ki67 (Abcam, cat# ab16667, 1:200).

Immunofluorescence. Organoids were fixed in 4% paraformaldehyde (Electron Microscopy Sciences)

for 15 minutes at room temperature. Mouse prostates were fixed using 4% paraformaldehyde for 2 hours at 4 degrees. Organoids and mouse prostate tissue were embedded using Tissue-Tek O.C.T. Freshly cut 5 Micron paraffin sections were stained with CK5 antibody (Covance, PRB-160P, 1:1,000); CK8 antibody (Covance, cat# MMS-162P, 1:1,000); ERG (Abcam, ab92513, 1:100); p63 (Abcam , ab735, 1:500) on Leica Bond RX (Leica Biosystems) with appropriate negative and positive controls. After washing in PBS, slides were mounted with Mowiol® 4-88 (Millipore, 475904) and imaged with a Leica TCS SP5 II confocal microscope. Immunofluorescence was independently performed twice.

Target cell quantification. To calculate the number of Trp63 and ERG, Krt5 and ERG double positive cells in the T2PE mice prostate PIN lesions and intraductal carcinomas, slides were scanned with Pannoramic Confocal Scanner (3DHistech, Hungary) using 40x/1.2 water objective. Appropriate areas of the scanned tissue were exported to .tiff files and analyzed in ImageJ/FIJI (NIH). Appropriate thresholds were applied for each channel, and cells expressing ERG were segmented out from the T2PE mice prostate PIN lesions and intraductal carcinomas. Then analysis was performed to determine the percentage of those cells were also positive for basal markers Trp63 or Krt5.

Western blotting. Cell lysates were prepared in RIPA buffer supplemented with proteinase/phosphatase inhibitor. Proteins were resolved in NuPAGE Novex 4–12% Bis-Tris Protein Gels (Life Technologies) and transferred electrophoretically onto a PVDF 0.45 mm membrane (Millipore). Blocking was conducted for 1 hour at room temperature in 5% milk in TBST buffer and were incubated overnight at 4 degrees with the primary antibodies diluted in 5% milk in TBST buffer. The following primary antibodies were used: β-Actin (Sigma-Aldrich, A3854, 1:5,000), ERG (Abcam, ab92513, 1:1,000); CK5 (Covance, PRB-160P, 1:1,000); CK8 (Covance, MMS-162P, 1:1,000); PTEN (Cell Signaling Technology, 9188, 1:1,1000); phospho-AKT (Ser473) (Cell Signaling Technology, 4060, 1:1,1000). Immunoblots were independently performed at least twice.

Quantitative RT-PCR analysis. Total RNA was extracted with TRIzol reagent (ambion, 15596018) and reverse transcription was further performed with 500 ng total RNA as initiation material with PrimeScriptTM RT Master Mix (TaKaRa, RR036A). qRT-PCR was conducted with 2 x S6 Universal SYBR qPCR Mix (NovaBio, Q204) using the manufacture’s protocol. The primers sequences are listed as followed:

Krt5-qPCR-F: GAACAGAGGCTGAGTCCTGGTA;

Krt5-qPCR-R: TCTCAGCCTCTGGATCATTCGG;

Krt14-qPCR-F: GAAGAACCGCAAGGATGCTGAG;

Krt14-qPCR-R: TGCAGCTCGATCTCCAGGTTCT;

Trp63-qPCR-F: GTATCGGACAGCGCAAAGAACG;

Trp63-qPCR-R: CTGGTAGGTACAGCAGCTCATC;

Krt8-qPCR-F: TGGAAGGACTGACCGACGAGAT;

Krt8-qPCR-R: GGCACGAACTTCAGCGATGATG;

Krt18-qPCR-F: AATCAGGGACGCTGAGACCACA;

Krt18-qPCR-R: GCTCCATCTGTGCCTTGTATCG.

ChIP-seq library preparation. 10 million target cells were collected for centrifuging and then resuspended in 10mL freshly made 1% formaldehyde with incubation at room temperature for 10 minutes with rotation. 526ul 2.5M glycine was added to a final concentration of 125 mM to quench the formaldehyde at room temperature for 5 minutes with rotation. Cells were pelleted and washed in ice-cold PBS. 880μL of ice-cold cell lysis buffer (1% SDS; 10 mM EDTA; 50 mM Tris-HCl; 1X proteinase inhibitor) was added to lyse cross-linked cells with rotation at 4 degrees for 30 minutes. 880μL cell lysis was transferred into a Covaris milliTUBE 1mL AFA Fiber and sheared with Covaris S220 (Fill level: 10; Duty cycle: 5; PIP: 140; Cycles/Burst: 200; Time: 4 minutes). Samples was clarified for 15 minutes at 16100 rcf at 4 degrees. Another 1600μL ChIP Dilution Buffer (0.01% SDS; 1.1% Triton X-100 (fisher scientific, BR151-500); 1.2 mM EDTA; 16.7 mM Tris pH 7.5; 167 mM NaCl) was added to achieve SDS concentration of 0.33%. 60μL Protein A beads was pre-cleared in 500μL ChIP Dilution Buffer. Protein A beads (Invitrogen, 10001D) were resuspended in 60μL of ChIP Dilution Buffer and added to sample at 4 degrees for 1 hour. Samples with beads was put on magnet and the supernatant was transferred into new tubes. Target antibody (5 μg for H3K27ac (Abcam, ab4729), 8 μg for ERG (Abcam, ab92513)) was added at 4 degrees for incubation overnight with rotation. To bind target antibody, 60μL pre-cleared Protein A beads were added to samples with two-hour rotation at 4 degrees. Beads was washed three times each with Low Salt Wash Buffer, High Salt Wash Buffer and LiCl Wash Buffer. 100μL of freshly-

made DNA Elution Buffer (50 mM $NaHCO_3$; 1% SDS) was added to resuspend ChIP sample beads with incubation at RT for 10 minutes, followed by 3 minutes at 37 degrees. ChIP sample beads were placed on magnet and the supernatant was transferred to a new tube. Another 100μL of DNA Elution Buffer was added to ChIP samples and the same incubation protocol was conducted. Supernatant of ChIP samples was collected again to the new tube. 10μL of Proteinase K (Invitrogen, 25530049) was added to each sample with incubation at 67 degrees at least 4 hours with shaking. DNA was purified with QIAGEN purification kit (QIAGEN, 28106) and eluted in 20μL of nuclease-free water.

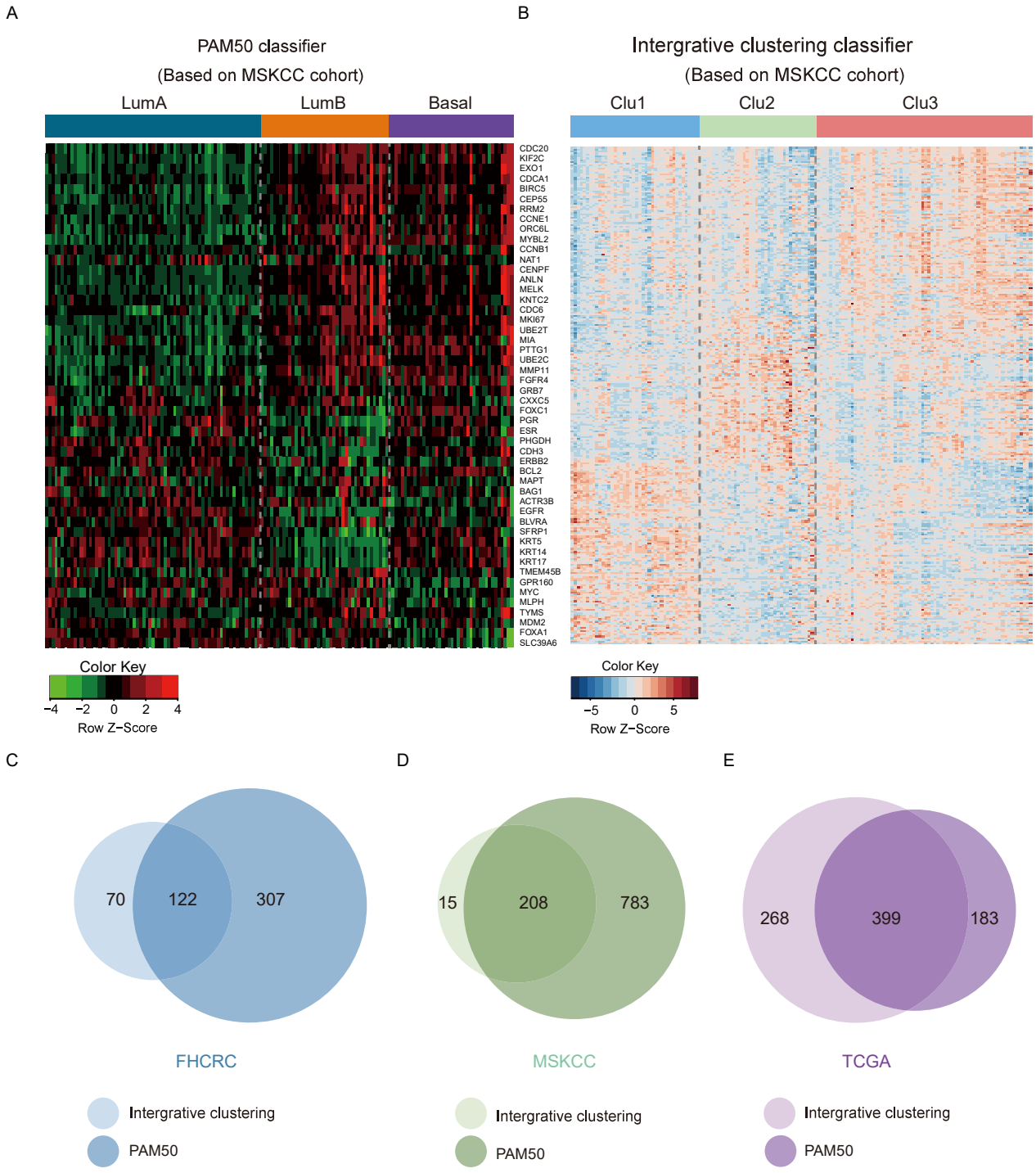
Bridge Linker-Hi-C (BL-Hi-C) assay. One million cells were cross-linked in 1mL 1% methanol free formaldehyde (Sigma-Aldrich, F8775) with shaking for 10 minutes at room temperature. 2.5M glycine was added to quench formaldehyde to a final concentration of 0.2M with shaking for 10 minutes at room temperature, and then the cells were put on ice for 5 minutes. Cell pellets were collected with centrifuging and then 1mL 0.1% SDS BL-Hi-C Lysis buffer (50 mM HEPES-KOH, pH 7.5; 150 mM NaCl; 1 mM EDTA; 1% Triton X-100; 0.1% Sodium Deoxycholate; 0.1% SDS) with proteinase inhibitor was added with incubation for 15 minutes at 4°C with shaking at 850 rpm. Cells were centrifuged and resuspended with 1mL of 0.55% SDS BL-Hi-C Lysis buffer (50 mM HEPES-KOH, pH 7.5; 150 mM NaCl; 1 mM EDTA; 1% Triton X-100; 0.1% Sodium Deoxycholate; 1% SDS; 1X proteinase inhibitor) for another 15-minute incubation at 4°C with shaking at 850 rpm. Cells were collected by centrifuging and resuspended with 100 μL of 0.3% SDS in 1×NEBuffer 2.1(NEB, B7002S) with following shaking at 37°C for 30 minutes at 900 rpm. 135μL nuclease-free water and 15μL 20% Triton X-100 were added and samples were kept at 37°C for 15min with shaking at 900rpm. Cell pellets were collected and resuspended by 76.5μL nuclease-free water, 10μL 10 X NEBuffer2.1, 2.5μL of 20% Triton X-100, 1μL BSA (100x, 10mg/mL) and 10μL HaeIII (10U/μL, NEB, R0108S) with shaking at 37°C overnight. 2μL 10mM dATP (NEB, N0440S) and 5μL Klenow Fragment (3′->5′ exo-, NEB, N0202S) for A-tailing were added at 37°C for 40 minutes. Cell pellets were collected and resuspended with 411μL nuclease-free water, 50μL T4 DNA ligase buffer (NEB, B0202S), 25μL 20% Triton X-100, 5μL BSA, 5μL T4 DNA ligase (NEB, M0202S) and 4μL Bridge Linker (Bridge Linker S2-F: /5P/CGCGATATC/iBIOdT/TATCTGACT; Bridge Linker S2-R: /5P/GTCAGATAAGATATCGCGT) with subsequent rotation at room temperature for 4 hours. Cell pellets were centrifuged at 1000g for 2 minutes at 4°C and resuspended with 88μL nuclease-free water, 10μL λ Exonuclease Buffer, 1μL λ Exonuclease (NEB, M0262L) and 1μL

Exonuclease I (NEB, M0293L) with shaking at 37°C for 60 minutes. DNA purification was conducted with Phenol: Chloroform: Isoamyl Alcohol 25:24:1(Sigma-Aldrich, cat. no. P3803) and eluted with 60μL of TE buffer. Add another 60μL 2XB&W buffer was added into DNA solution for sonication with setting Covaris parameters for the DNA size of 300 bp. Streptavidin C1 beads were washed (Invitrogen, 65001) with 1mL 1X TWB buffer for twice and resuspended with 20μL 1× B&W buffer. To bind target DNA, beads suspension was added into 130μL sonicated DNA solution with incubation at room temperature for 15 minutes with rotation at 900 rpm. Beads was then washed with 1X TWB buffer, BW buffer and nuclease-free water. Finally, DNAs-on-beads was resuspended in 50μL nuclease-free water.

Supplemental reference

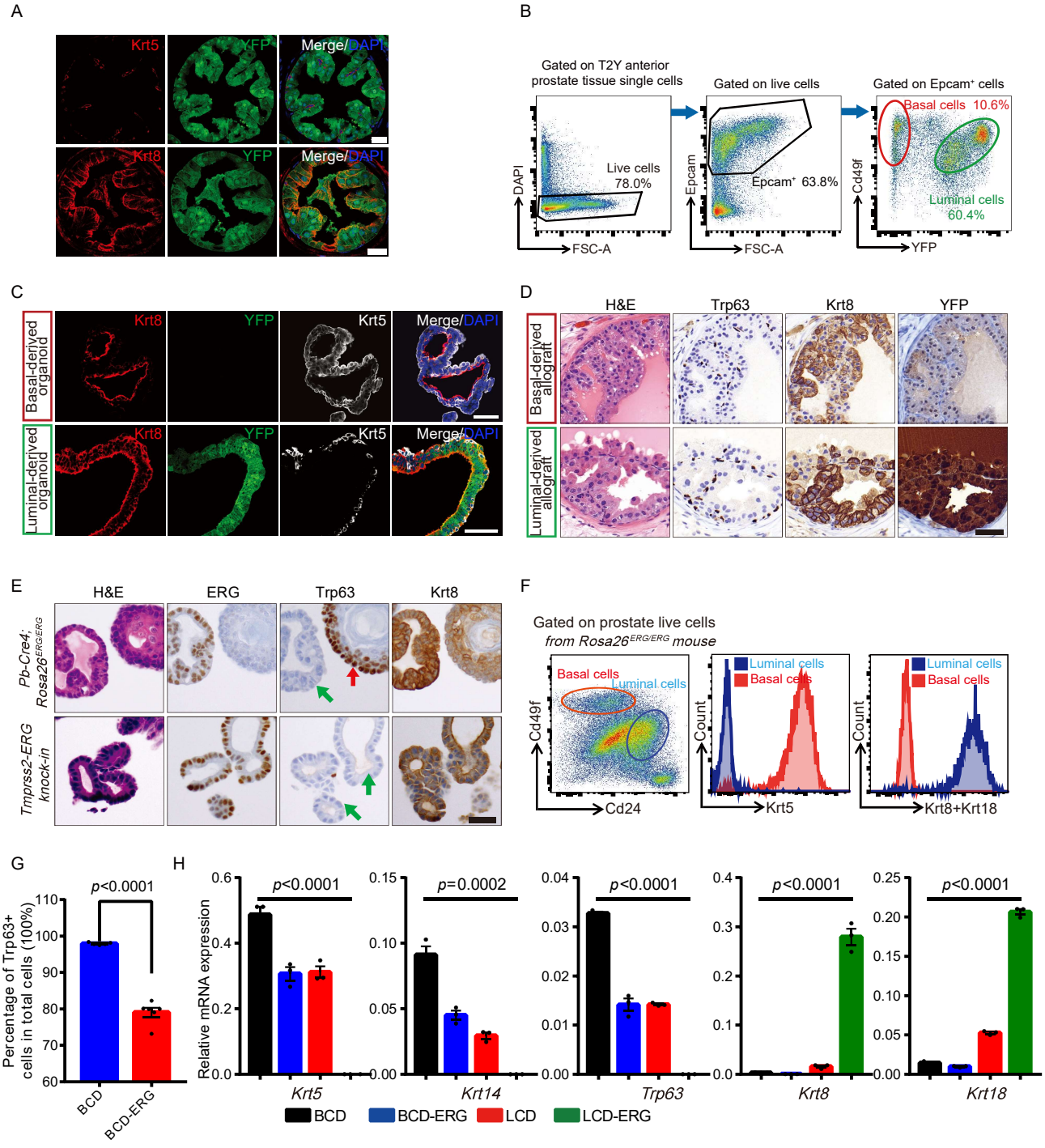
1. Drost J, Karthaus WR, Gao D, Driehuis E, Sawyers CL, Chen Y, et al. Organoid culture systems for prostate epithelial and cancer tissue. *Nat Protoc*. 2016;11(2):347-58.

Supplemental Figure. 1



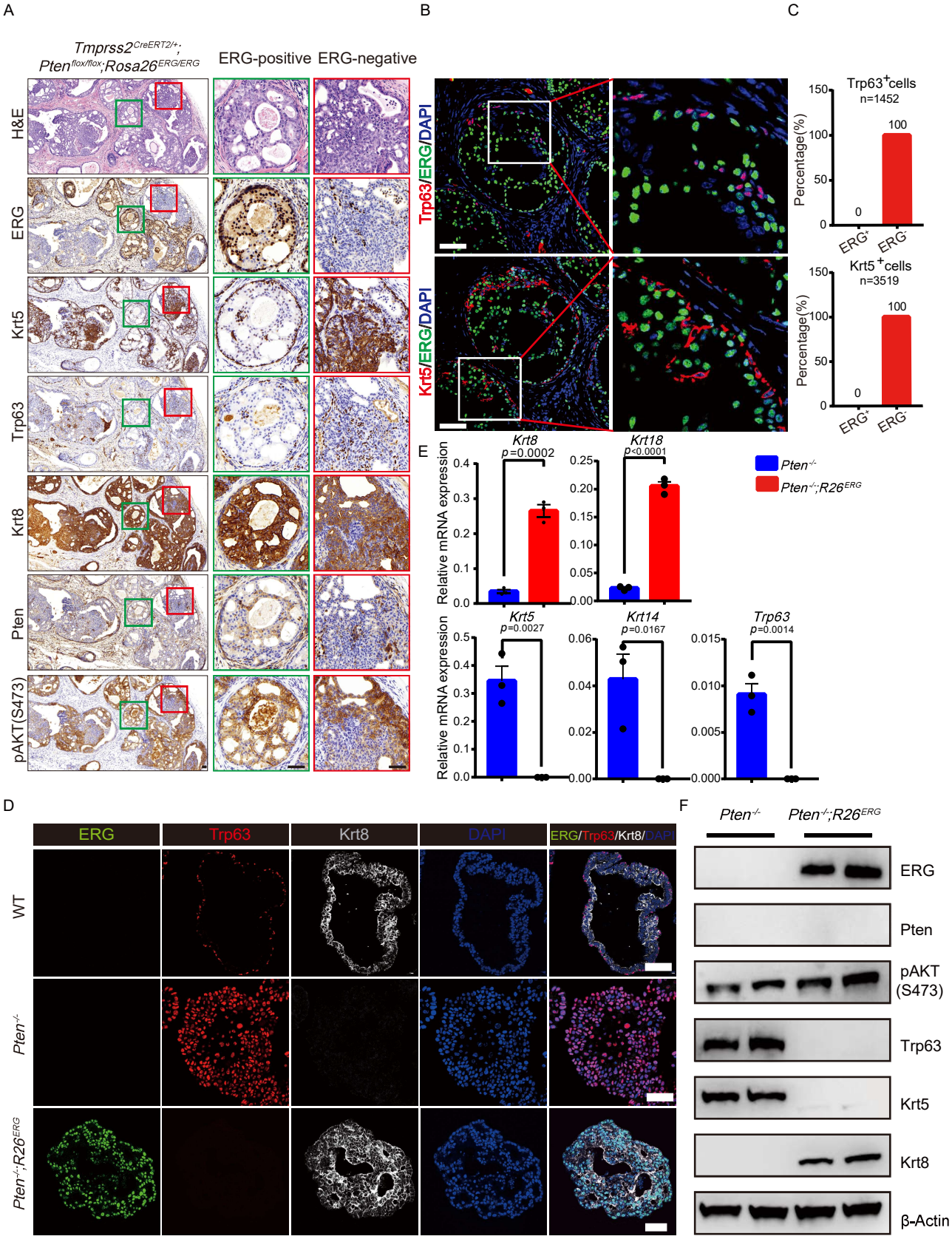
Supplemental Figure 1. Identification for the master transcription factors with the combination analysis of integrative classifier and PAM50 classifier. (A) Heatmap showing the expression of PAM50 classification marker genes in each sample of MSKCC cohort, with the results of three subtypes including Luminal A (dark blue), Luminal B (orange) and Basal (purple). The genes are represented with the same order as shown in (41). **(B)** Heatmap showing the expression of integrative classification marker genes in each sample of MSKCC cohort, with the results of three subtypes including Clu 1 (light blue), Clu 2 (light green) and Clu3 (red). The genes are represented with the same order as shown in (40). **(C-E)** Venn diagram showing the number of the overlapped TFs identified by integrative classifier and PAM50 classifier in three cohorts: 122 overlapped TFs of FHCRC cohort **(C)**, 208 overlapped TFs of MSKCC cohort **(D)**, 399 overlapped TFs of TCGA cohort **(E)**.

Supplemental Figure. 2



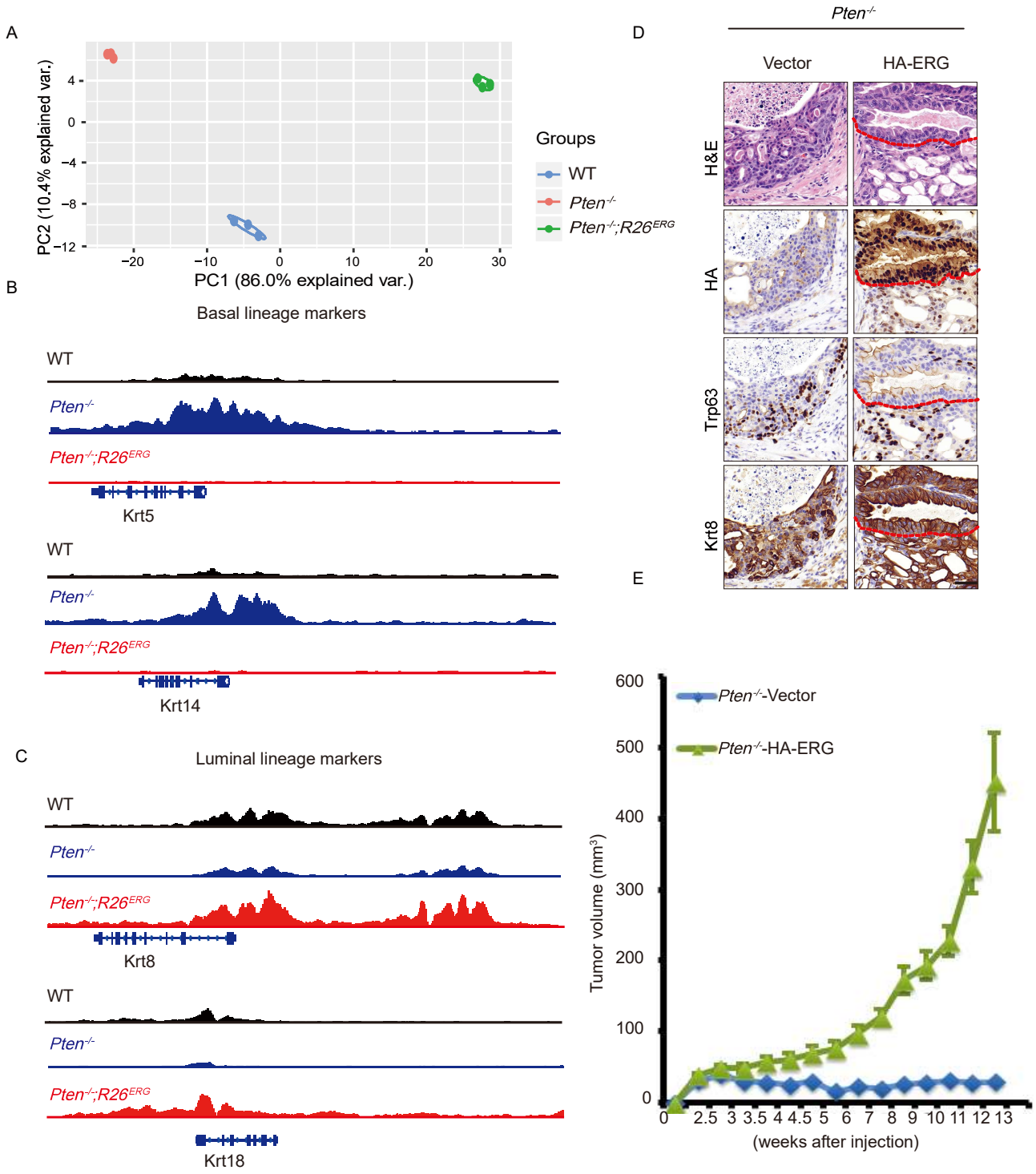
Supplemental Figure 2. Normal prostate epithelial cells have multiple cell fate and ERG overexpression promotes prostate luminal lineage differentiation. (A) Immunofluorescence staining of Krt5, Krt8 and YFP in T2Y mice prostates. **(B)** Isolation of $Cd49f^{low}/YFP^{+}$ luminal cells and $Cd49f^{high}/YFP^{-}$ basal cells by flow cytometry. **(C)** Immunofluorescence staining of Krt8, YFP and Krt5 of basal-cell-derived organoids (top) and luminal-cell-derived organoids (bottom). **(D)** H&E and Trp63, Krt8 and YFP IHC staining of allografts from UGSM tissue recombination assays after transplantation of basal- (top) or luminal-cell-derived organoid (bottom) generated from T2Y mice anterior prostate. **(E)** H&E and ERG, Trp63 and Krt8 IHC staining of prostate organoids derived from *Pb-Cre4; Rosa26*$^{ERG/ERG}$ and *Tmprss2-ERG* knock-in mice, respectively. Trp63-negative organoids with ERG expression were indicated by green arrows and Trp63-positive organoids without ERG expression were indicated by red arrows. **(F)** Sorting strategy for prostate basal ($Cd49f^{high}/Cd24^{low}$) and luminal ($Cd49f^{low}/Cd24^{high}$) cells from *Rosa26*$^{ERG/ERG}$ mice (left), intracellular flow cytometry for basal cell lineage marker Krt5 (middle) and luminal cell lineage markers Krt8/Krt18 (right) on *Rosa26*$^{ERG/ERG}$ mice prostate cells. **(G)** Quantification statistics for the percentage of $Trp63^{+}$ cells in total cells (analyses were performed based on 8613 BCD cells and 7300 BCD-ERG cells respectively, two-tailed t-test, mean ± sem, n=5-6). **(H)** QRT-PCR analysis for mRNA expression of basal cell lineage markers *Krt5*, *Krt14*, *Tr63* and luminal cell lineage markers *Krt8*, *Krt18* (one-way ANOVA, mean ± sem, n=3). Scale bars, 50 μm.

Supplemental Figure. 3



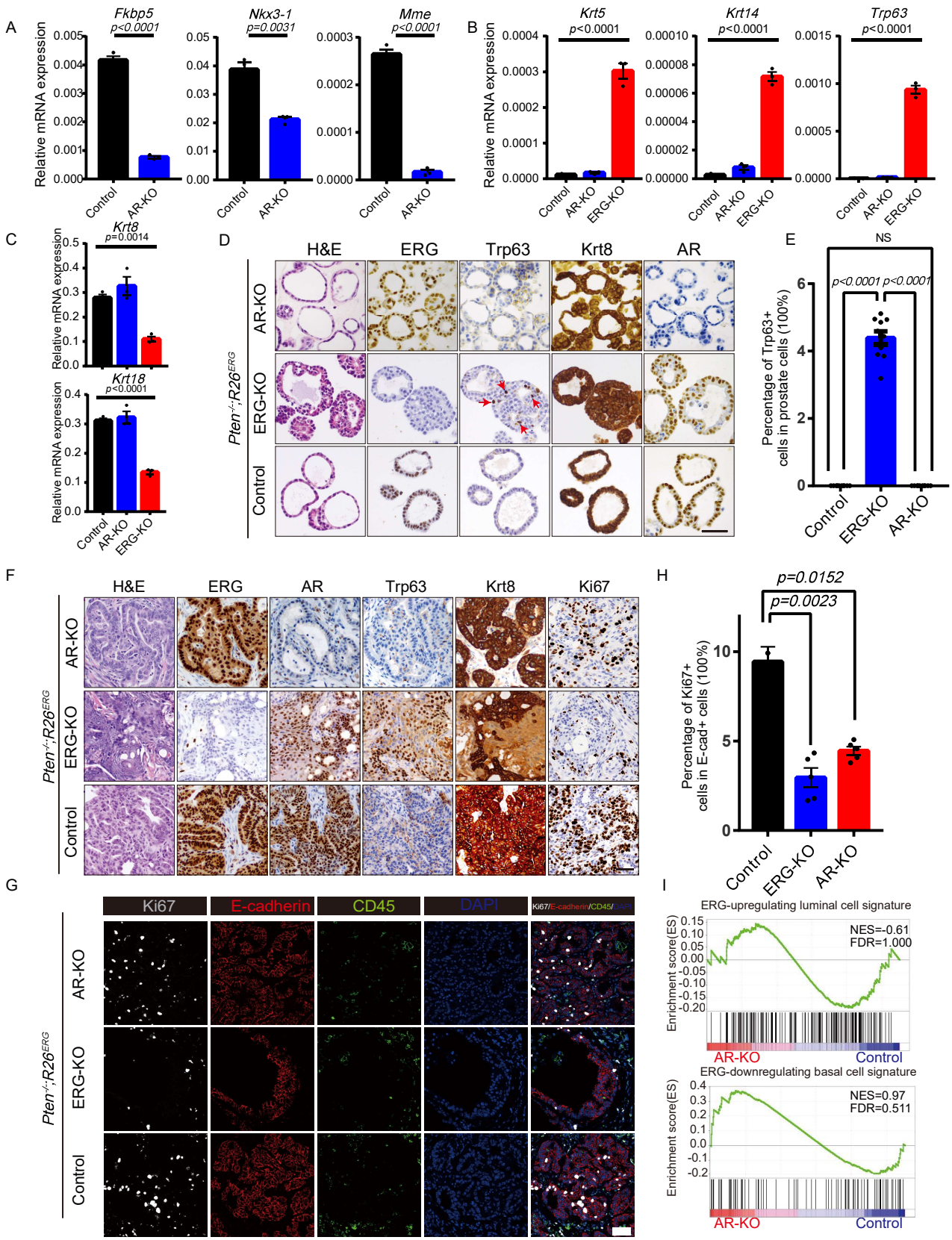
Supplemental Figure 3. ERG regulates prostate cancer cell lineage in the context of Pten loss. **(A)** H&E and ERG, Krt5, Trp63, Krt8, Pten and pAKT(S473) IHC staining of T2PE mice prostate tumors with 7 months post tamoxifen injection, green box indicates ERG-positive region and red box indicates ERG-negative region. **(B)** Immunofluorescence staining of ERG, Krt5 and Trp63 in T2PE mice prostate with 7 months post tamoxifen injection. **(C)** Histogram statistic for the percentage of ERG^+ cells and ERG^- cells in $Trp63^+$ cells (top) and $Krt5^+$ cells (bottom) respectively. **(D)** ERG, Trp63, Krt8 and DAPI IF staining for WT, *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids. **(E)** QRT-PCR analysis for mRNA expression of basal cell lineage markers *Krt5*, *Krt14*, *Trp63* and luminal cell lineage markers *Krt8*, *Krt18* (two-tailed t-test, mean ± sem, n=3). **(F)** Western blotting analysis of ERG, Pten, pAKT(S473), Trp63 and Krt5 in *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids.

Supplemental Figure. 4



Supplemental Figure 4. ERG enhanced prostate cancer luminal identity, validated by multi-omics evidences. (A) PCA plot of RNA-seq for WT, *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids using ERG-upregulating luminal cell signature genes and ERG-downregulating basal cell signature genes. **(B-C)** Snapshot representation of the H3K27ac peak profile (defined by H3K27ac ChIP-seq) associated with prostate basal cell lineage markers Krt5, Krt14 **(B)**and luminal cell lineage markers Krt8, Krt18 for WT, *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids **(C)**. **(D)** H&E and HA, Trp63 and Krt8 IHC 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 (right) or a control vector (left). **(E)** Subcutaneous tumor growth in SCID mice of Pten loss organoids overexpressing TMPRSS2-ERG fusion protein with HA tag or a control vector respectively (mean ± sem, n=10). Scale bars, 50 μm.

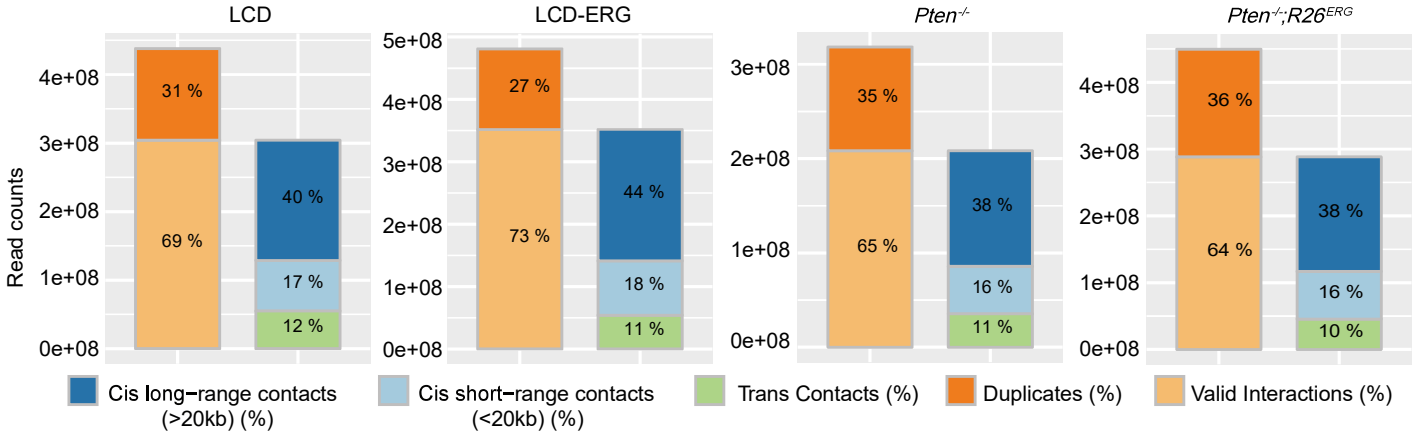
Supplemental Figure. 5



Supplemental Figure 5. ERG instead of AR is critical for prostate cancer cell lineage regulation in Pten loss context. (A) QRT-PCR analysis for mRNA expression of AR target genes *Fkbp5*, *Nkx3-1* and *Mme* (two-tailed t-test, mean ± sem, n=3). **(B-C)** QRT-PCR analysis for mRNA expression of basal cell markers *Krt5*, *Krt14* **(B)** and *Trp63* and luminal cell markers *Krt8*, *Krt18* **(C)** (one-way ANOVA, mean ± sem, n=3). **(D)** H&E and ERG, Trp63, Krt8 and AR IHC staining of $Pten^{-/-}; R26^{ERG}$ organoids infected with a lentiviral CRISPR/Cas9 carrying guide RNA targeting the AR (AR-KO, top) and ERG (ERG-KO, middle) and a control vector (Control, bottom), red arrows indicate Trp63 positive cells. **(E)** Quantification statistics for the percentage of $Trp63^+$ in total cells (analyses were performed based on 30127 AR-KO cells, 39558 ERG-KO cells and 29156 Control cells, two-tailed t-test, mean ± sem, n=10). **(F)** H&E and ERG, AR, Trp63, Krt8 and Ki67 IHC 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) respectively. **(G)** Ki67, E-cadherin, Cd45 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) respectively. **(H)** Quantification statistics for the percentage of $Ki67^+$ cells in E-$cadherin^+$ prostate cells (analyses were performed based on 4027 E-$cadherin^+$ cells of AR-KO allografts, 3934 E-$cadherin^+$ cells of ERG-KO allografts and 3764 E-$cadherin^+$ cells of Control allografts, two-tailed t-test, mean ± sem, n=5). **(I)** GSEA enrichment plot of AR-KO versus Control using ERG-upregulating luminal cell signature genes (top) and ERG-downregulating basal cell signature genes (bottom) respectively.

Supplemental Figure. 6

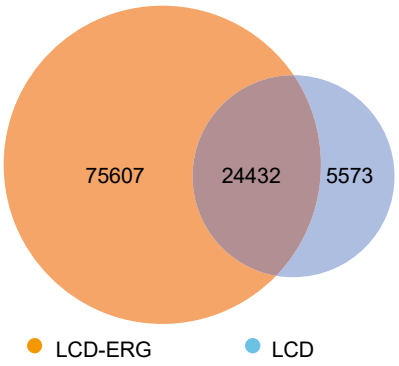
A



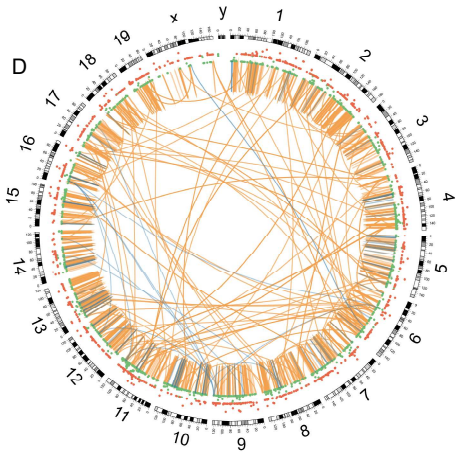
B

Sample	Unique Pairs	Valid Pairs	Cis Pairs	Percentage of Cis
LCD	487,636,911	304,119,352	249,087,773	81.90%
LCD-ERG	517,797,373	351,929,491	297,643,740	84.57%
Pten$^{-/-}$	380,991,969	208,194,264	172,818,759	83.01%
Pten$^{-/-}$;*R26*^{ERG}	520,754,171	288,278,089	242,728,002	84.20%

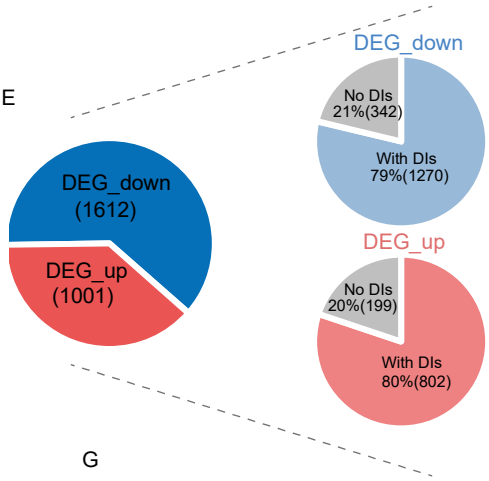
C



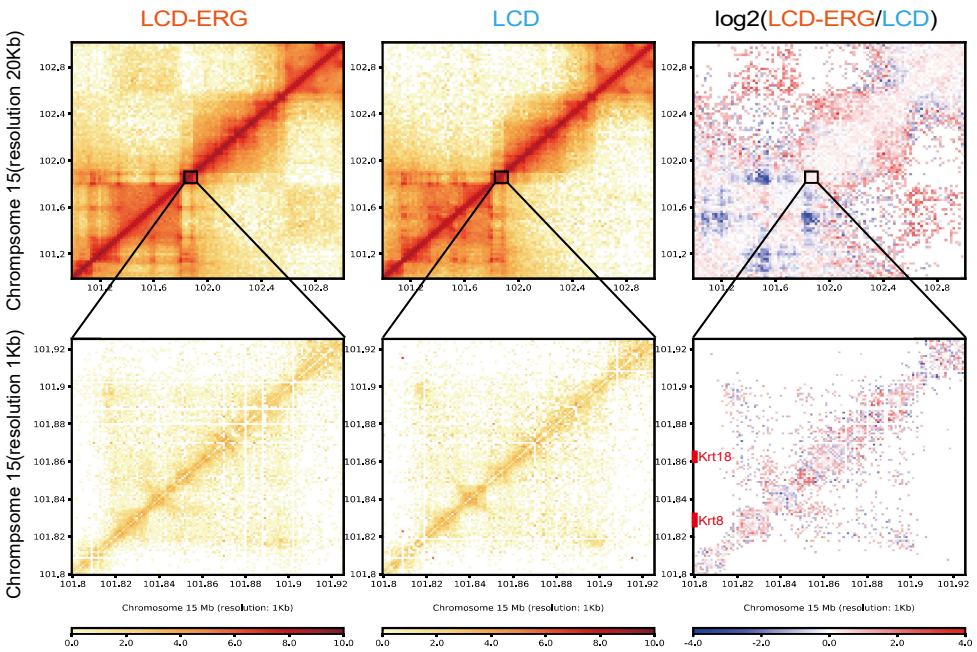
D



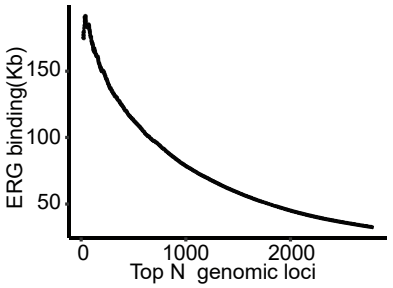
E



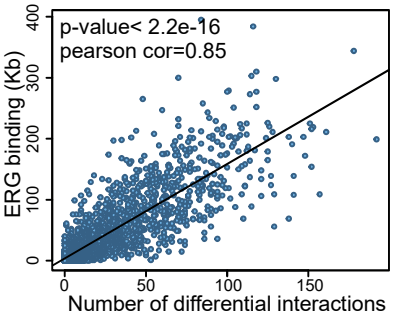
F



G

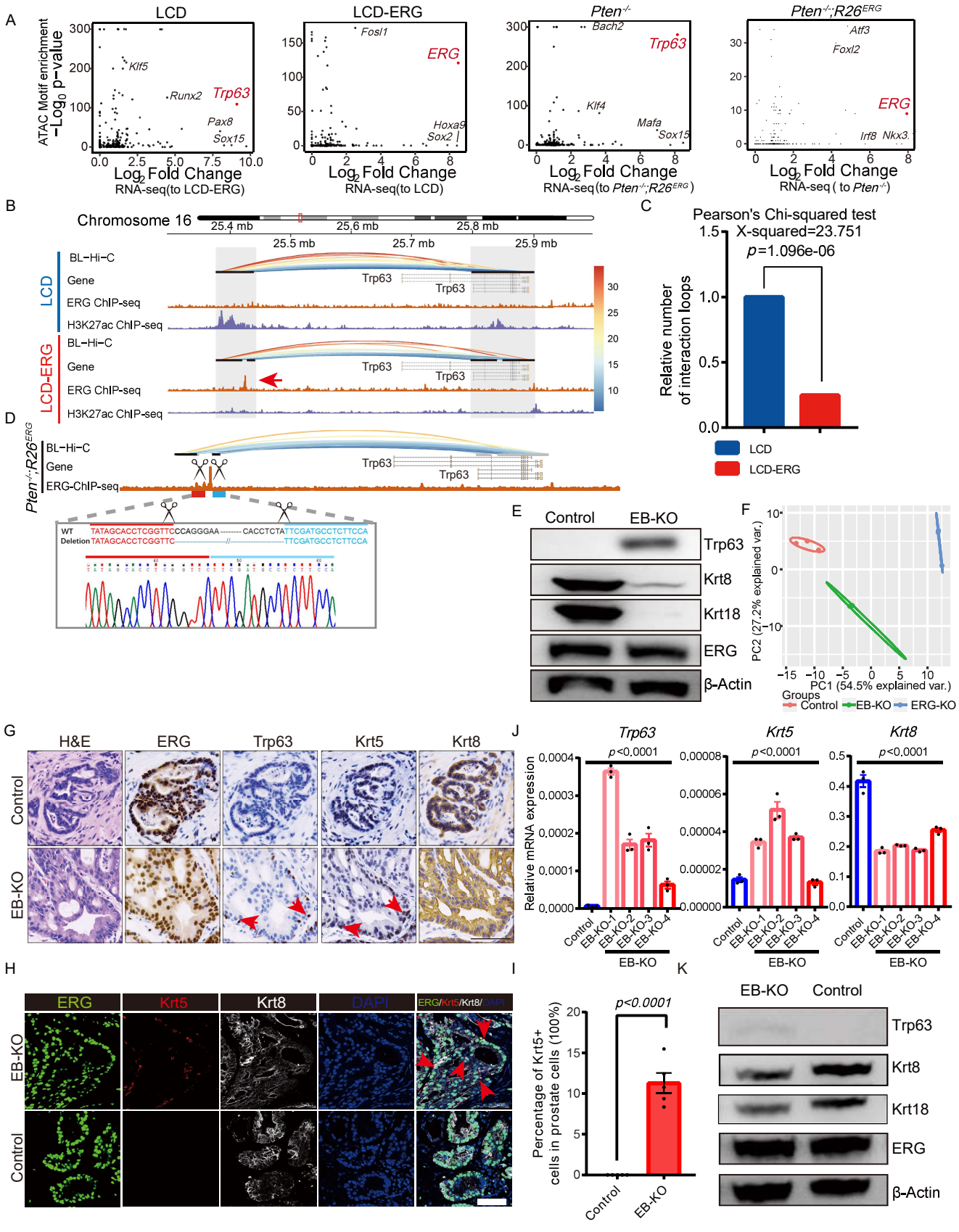


H



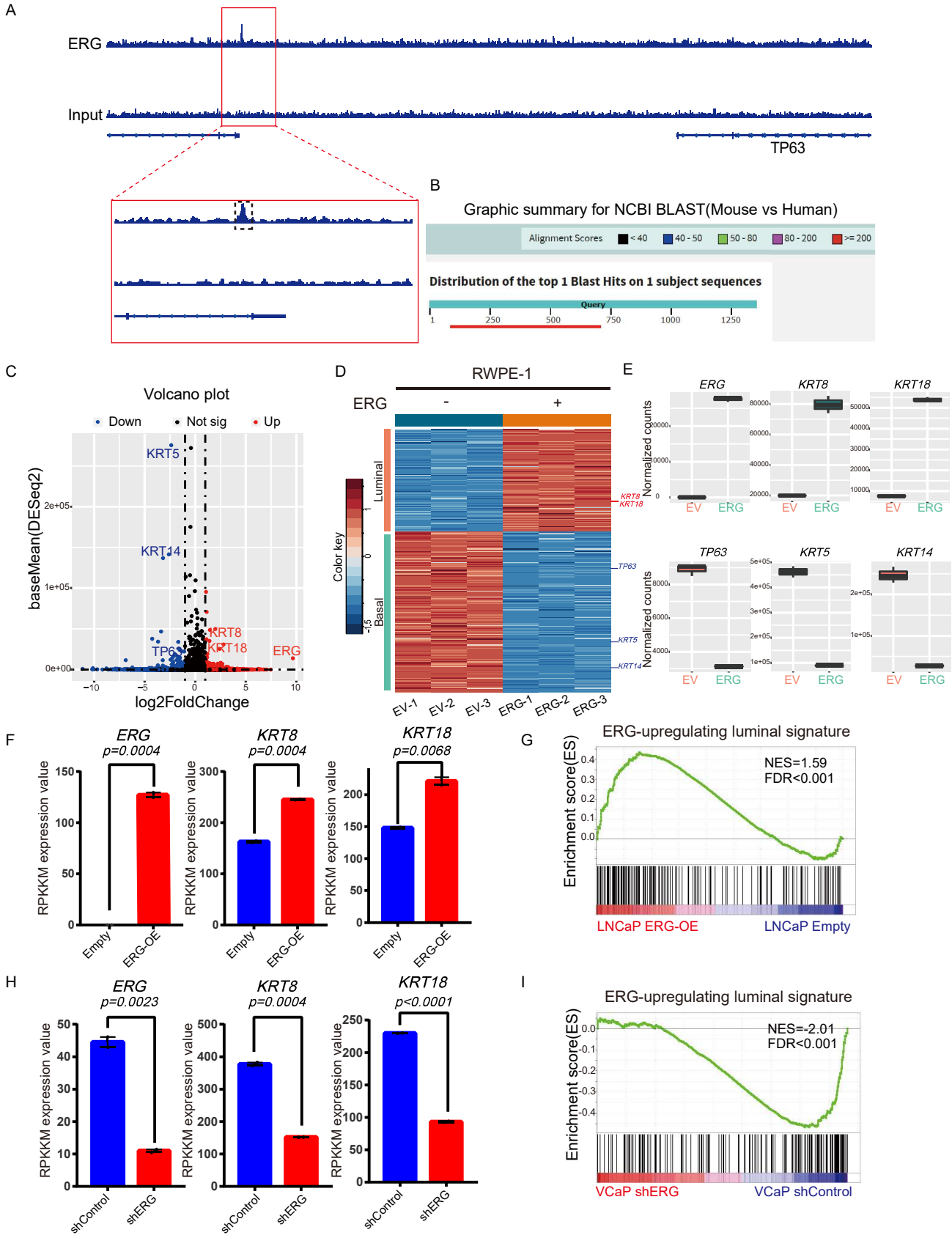
Supplemental Figure 6. ERG alters chromatin interactions associated with gene expression in normal prostate epithelial cells. (A) Histogram statistics for the composition of valid pairs in the libraries of LCD, LCD-ERG, *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids respectively. **(B)** Statistics on the number of unique pairs, valid read pairs, cis pairs and percentage of cis pairs in BL-Hi-C libraries of LCD, LCD-ERG, *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids respectively. **(C)** Venn plot showing differential chromatin interactions between LCD and LCD-ERG organoids. **(D)** Circos plot depicting chromosomes 1-19, X and Y on the basis of BL-Hi-C data and RNA-seq data, indicating DIs including LCD-ERG-specific DIs(orange) and LCD-specific DIs(light blue), DEGs including up-regulated DEGs of LCD-ERG (red) and down-regulated DEGs of LCD-ERG (green), respectively. **(E)** Venn plot showing the percentage of down-regulated DEGs(top) and up-regulated DEGs with DIs(bottom). **(F)** The normalized interaction heatmaps of LCD-ERG (left), LCD (middle), and the difference (right) at 20 kb resolution (top) and 1 kb resolution (bottom) of chromosome 15, including Krt8 and Krt18 genomic region. **(G)** Plot showing the density of ERG binding (kb) at each of the ranked (N) differential interacting chromatin loci of 1-Mb intervals. **(H)** Correlation plot showing the positive relationship between ERG binding density and the number of DIs in 1-Mb intervals.

Supplemental Figure. 7



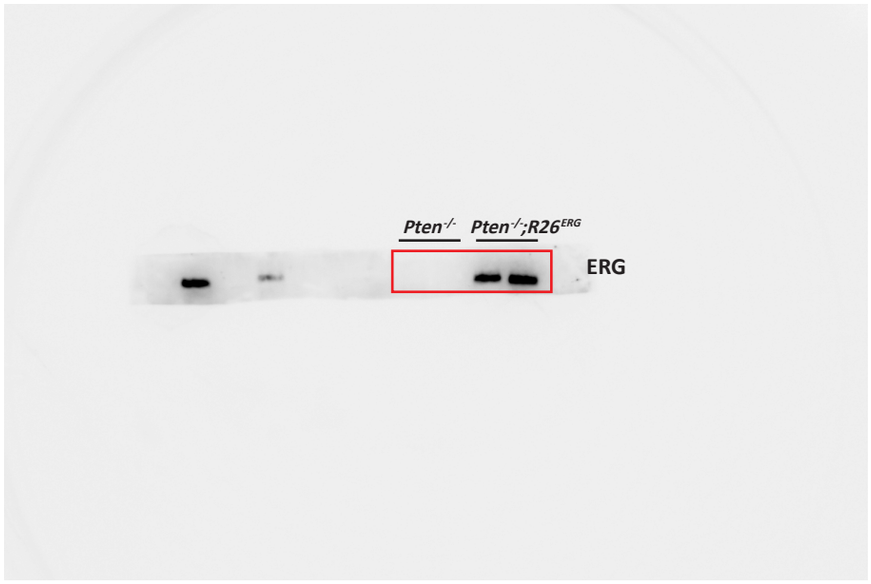
Supplemental Figure 7. Integrating analysis of ATAC-seq and RNA-seq reveals Trp63 and ERG as master regulators and ERG binding site knock-out impairs luminal lineage phenotype. (A) Dot plots for transcription factors showing the p-value of motif enrichment analysis and mRNA expression changes from ATAC-seq data and RNA-seq data respectively in LCD, LCD-ERG, *Pten*$^{-/-}$ and *Pten*$^{-/-}$*; R26*^{ERG} organoids respectively. **(B)** 3D signal of BL-Hi-C showing chromatin interactions between the distal ERG binding site and Trp63 gene body region in LCD (top) and LCD-ERG (bottom) organoids respectively. Red arrow indicates the distal ERG binding site. **(C)** Pearson's Chi-squared test to evaluate the differences of interaction loops density between LCD and LCD-ERG organoids. **(D)** Schematic diagram for the strategy to delete the distal ERG-binding site(top) and Sanger sequencing for identification of knock-out efficiency (bottom). **(E)** Western blotting analysis of Trp63, Krt8, Krt18 and ERG in EB-KO and Control of *Pten*$^{-/-}$*; R26*^{ERG} organoids. **(F)** PCA plot for EB-KO, ERG-KO and Control organoids using prostate cell lineage signature genes. **(G)** H&E and ERG, Trp63, Krt5, Krt8 IHC staining of allografts derived from EB-KO and Control of *Pten*$^{-/-}$*; R26*^{ERG} organoids using UGSM tissue recombination assays. **(H)** ERG, Krt5, Krt8 and DAPI IF staining for allografts of UGSM tissue recombination assays derived from EB-KO and Control organoids, red arrows indicate ERG^{+}/Krt5^{+} cells. **(I)** Quantification statistics for the percentage of ERG^{+}/Krt5^{+} in ERG^{+} prostate cells (analyses were performed based on 3486 ERG^{+} cells of EB-KO and 3591 ERG^{+} cells of Control, two-tailed t-test, mean ± sem, n=5). **(J)** QRT-PCR analysis of mRNA expression of *Trp63*, *Krt5* and *Krt8* in EB-KO and Control of LCD-ERG organoids (one-way ANOVA, mean ± sem, n=3). **(K)** Western blotting analysis of Trp63, Krt8, Krt18 and ERG in EB-KO and Control of LCD-ERG organoids. Scale bars, 50 μm.

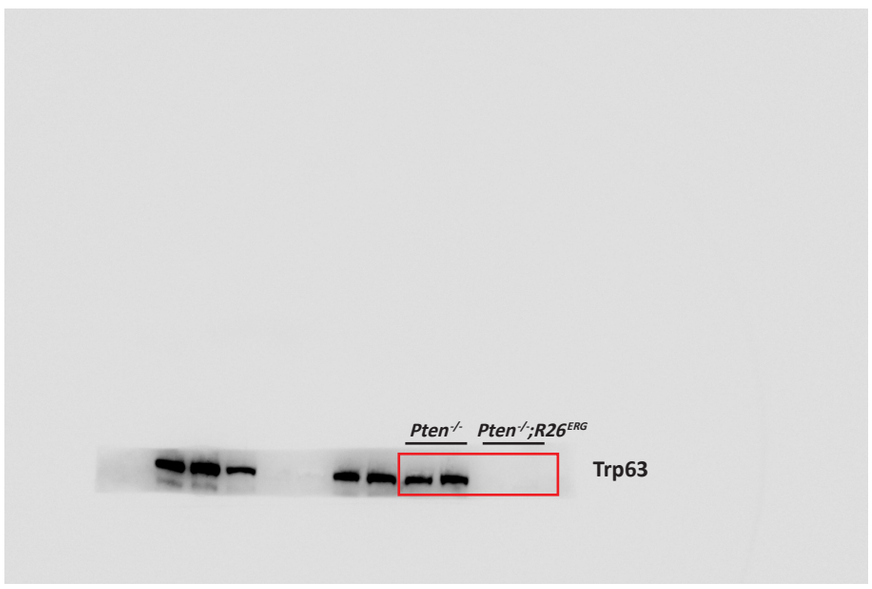
Supplemental Figure. 8

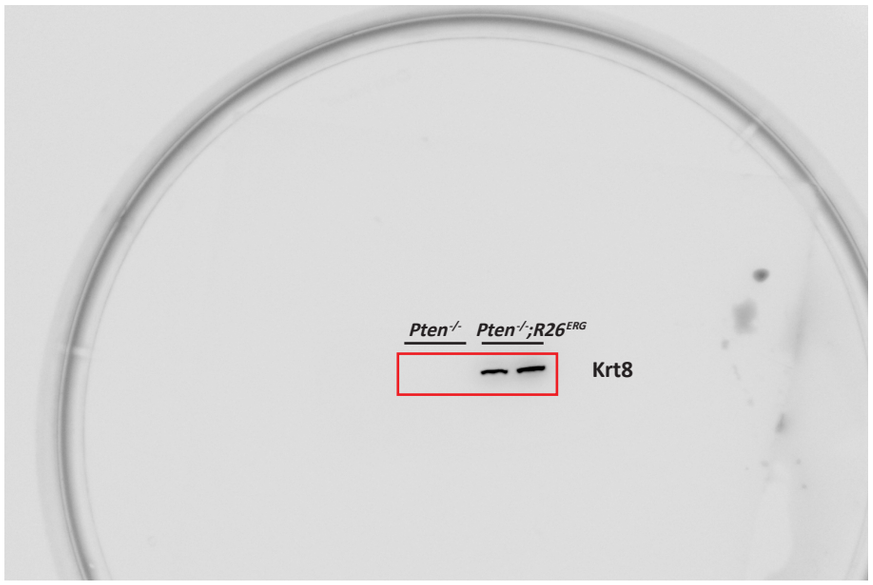


Supplemental Figure 8. Identification of the consistent function of ERG and the conserved distal ERG binding site in human prostate cells. (A) Snapshot representation of the ERG binding peak profile (defined by ERG ChIP-seq) in RWPE-1 cells with ERG expression. **(B)** Graphic summary for NCBI BLAST tools using the sequence of the distal ERG binding sites in mouse and human prostate cells. **(C)** Volcano plot for representative differentially expressed genes (DEGs) of RWPE1-ERG versus RWPE1-EV, up-regulated DEGs are marked by red, down-regulated DEGs are marked by blue. **(D)** Heatmap showing the expression of up-regulated luminal lineage genes and down-regulated basal lineage genes in RWPE1-ERG and RWPE1-EV cells respectively (RWPE1-ERG versus RWPE1-EV). **(E)** Box plots showing normalized counts (using DESeq2) of *ERG*, *KRT8*, *KRT18*, *TP63*, *KRT5* and *KRT14* respectively. **(F)** FPKM expression values for *ERG* (left), *KRT8* (middle) and *KRT18* (right) in empty LNCaP and ERG-OE LNCaP cells (two-tailed t-test, mean ± sem, n=2). **(G)** GSEA enrichment plot to compare ERG-OE LNCaP with empty LNCaP using the homologous ERG-upregulating luminal signature genes. **(H)** FPKM expression values for *ERG* (left), *KRT8* (middle) and *KRT18* (right) in shControl VCaP and shERG VCaP cells (two-tailed t-test, mean ± sem, n=2). **(I)** GSEA enrichment plot to compare shERG VCaP with shControl VCaP using the homologous ERG-upregulating luminal signature genes.

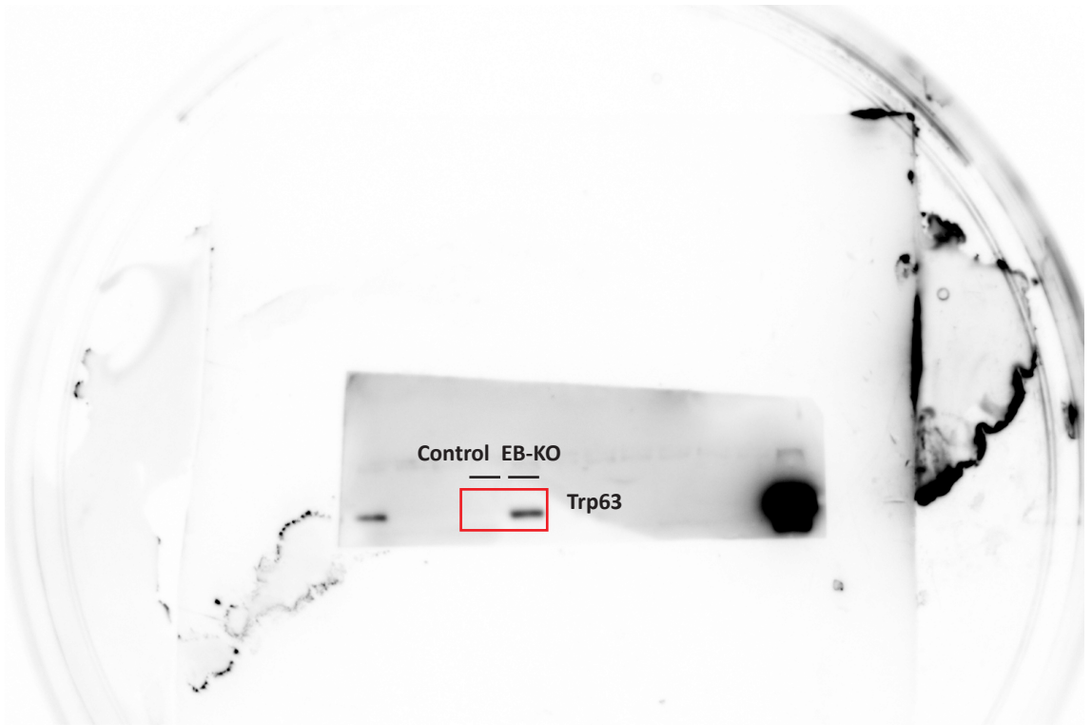
Full unedited gel for Supplemental Figure 3F

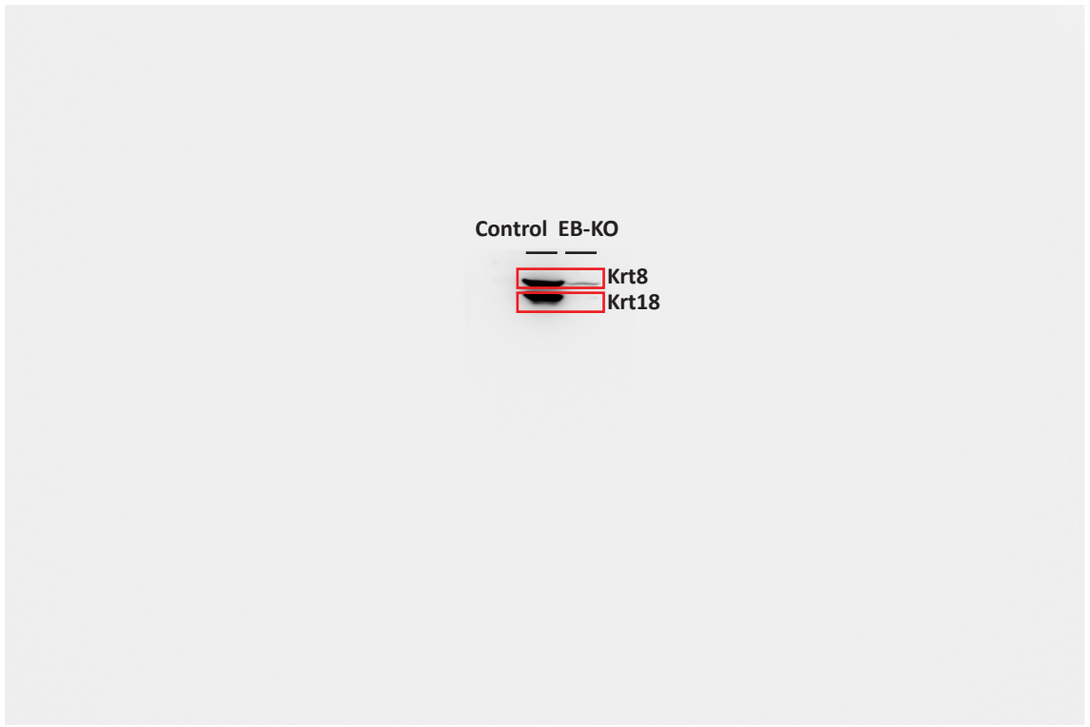


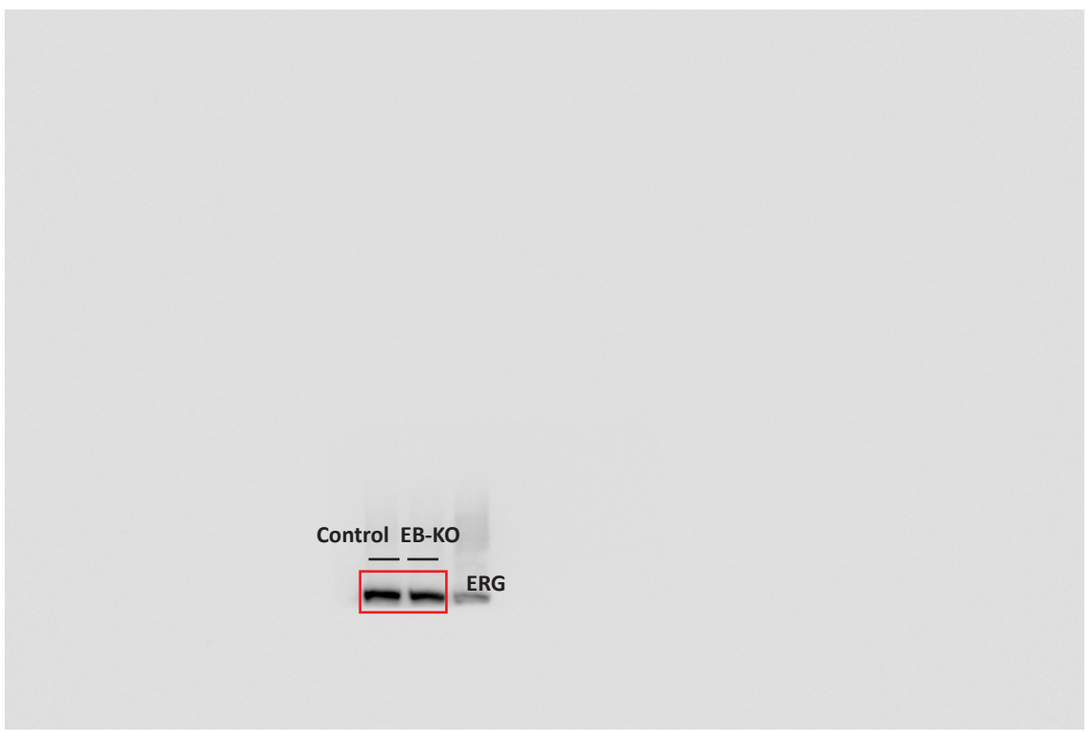




Full unedited gel for Supplemental Figure 7E







Full unedited gel for Supplemental Figure 7K