Supplementary Methods

Patient samples
Tissue samples from patients with organ-confined prostate cancer treated with radical prostatectomy were collected with the approval of the Ethics Committee of the IRCCS Multimedica of the Regione Lombardia, Italy, and patient written informed consent. Tissue microarrays containing samples from normal, HGPIN and adenocarcinoma were prepared from paraffin-embedded tissues as previously described (1). Histopathological and clinical data were collected at Department of Pathology (IRCCS Multimedica, Italy) and are shown in Supplementary Figure 1. Tissue from the second cohort were taken from patients with organ-confined prostate cancer treated with radical prostatectomy at the Insespital (Bern, CH) with the approval of the Ethical Committee and patient’s written informed consent. Tissue microarrays containing samples from primary prostate tumours ($n=226$) were prepared from paraffin-embedded tissues as previously described (2). Histopathological and clinical data were collected and continuously updated at the Department of Urology (Insespital, Bern) and have been previously described (2) (Supplementary Figure 1).

Chromatin Immunoprecipitation
Computational search for ETS binding sites on selected gene promoters was performed using Motifviz (biowulf.bu.edu/MotifViz). Chromatin immunoprecipitation was performed with anti-ESE3 (Clone 5A5.5, Lab Vision, Fremont, CA USA) anti H3Ac (Up-state,Millipore); anti H3K9 2 met (Up-state,Millipore) and IgG control antibody. Samples were analyzed as previously described (3) by end-point PCR and primer sets reported in Supplementary Table 1.

In vitro sphere-forming and self renewal assay
Single cell suspensions were plated in ultra-low attachment dishes (Corning Life Sciences) in serum-free Mammary Epithelial Basal Medium (MEBM, Cambrex) (4). Prostato-spheres (PS) $\geq 50$ $\mu$m in diameter were counted after 7 to 14 days. PS were collected and filtered through 40-$\mu$m filters to isolate RNA or for re-plating. To test self-renewal capability, PS were serially passed for at least 5 generations. To assess clonality of the PS, cells were labeled separately with either PKH67 or PKH26 fluorescent dye (Sigma) and mixed in equal proportions prior to performing the assay. Single cell derived spheres were expected to be stained with only one fluorescent dye. For limiting dilution assay, single cells were plated in 96-well plates at a density of 1 cell/well. Wells containing single cells were monitored. The number of PS was determined after 7-10 days. Each experiment was carried out in triplicate and repeated at least three times.
Animals and tumour xenografts

Mice were purchased from the Harlan Laboratories. Mice were maintained under pathogen-free conditions with food and water provided *ad libitum* and their general health status was monitored daily. All protocols involving animals were conducted in conformity with the institutional guidelines for experiments on animals and in compliance with national and international policies. Study protocol was approved by the Swiss Veterinary Authority (No.5/2011). For subcutaneous tumour xenografts, PrEc and $ESE3^{kd}$-PrECs ($4 \times 10^6$ cells) were inoculated in the flank of SCID mice (CB.17/lcrHan Hsd-Prkds; $n=4$/group) and DU145 cells ($4 \times 10^6$ cells) in athymic male nude mice (Balb c nu/nu; $n=8$/group). Tumour size was monitored twice a week with a caliper. Cells in primary PS derived from $ESE3^{kd}$-PrECs and $ESE3^{kd}$-RWPE-1 cells were dissociated into single cell suspensions and implanted subcutaneously with Matrigel ($1.5 \times 10^5$ cells/injection site) in the flank of SCID mice (CB.17/lcrHan Hsd-Prkds; $n=3$/group). A similar number of cells derived from primary PS from control PrEC and RWPE-1 cells were injected in the other flank of each mouse. For metastasis assay, $ESE3^{kd}$-PrECs and control PrECs ($1 \times 10^6$) were injected into tail vein of SCID mice ($n=4$/group) twice with a 24-h interval between injections. Animals were sacrificed after 8 and 12 weeks. Lungs were collected, included in paraffin and stained with H & E or by immunohistochemistry. Multiple sections from different lobes for each mouse were scored.

Immunohistochemistry

Immunohistochemistry on clinical tissue samples was carried out using anti-ESE-3 (Clone 5A5.5) rat monoclonal antibody (1:75 dilution, LabVision Corporation.), anti-ERG mouse monoclonal (1:100 dilution, Epitomics) and anti- LTSV40T Rabbit Polyclonal Antibody (1:100 dilution, Ag Santa Cruz). We used, normal prostate and colon tissues as a positive and negative control. Formalin-fixed and paraffin-embedded pellets of LNCaP and PC3 cells were also used to test the optimal conditions for each antibody and, depending on the antibody, as positive and negative control samples. Samples, in which the primary antibody was omitted, were also included as negative controls. Staining was independently scored by three expert pathologists (FS, GD and MS).

Identification of gene signature and functional annotation analysis

To extract the $ESE3^{low}$ signature, filtered matrices were prepared as previously described (3). For the Wallace and Glinsky dataset, probesets with median intensity $>50$ and coefficient of variation $>0.3$ were selected and intensity values were $\log_2$ transformed. Differential gene expression analysis between $ESE3^{low}$ samples vs all the other tumours was performed using GEPAS (Gene Expression
Profile Analysis Suite, http://gepas.bioinfo.cipf.es/). Genes differentially expressed were obtained after filtering on q-value (q <0.05). Lists of up and downregulated genes obtained with GEPAS were then analyzed by means of Metacore software (portal.genego.com) searching for the mostly overrepresented biological maps (GeneGO Pathway Maps) within the distinct signatures (FDR <0.05 and p-value <0.01).

**Gene Set Enrichment analysis (GSEA)**

Log$_{10}$ (PCa vs Normal) and intensity matrices (for Kunderfranco, Wallace and Glinsky datasets, respectively) were loaded on GSEA tool (Gene Set Enrichment Analysis, broadinstitute.org/gsea). Gene lists used for the GSEA were as follow: 1) genes deregulated in $ESE3^{kd}$-PrECs vs control PrECs; 2) core EMT gene signature identified by Taube (5); 3) core breast cancer stem cell genes_Gupta (6).

**Hierarchical clustering**

Core cancer stem cell gene (6) were mapped to the Agilent and Affymetrix microarrays datasets analyzed. For the Kunderfranco dataset the log$_{10}$(PCa vs Normal) matrix was loaded on MEV (MultiExperiment Viewer, www-tm4.org). For the Wallace and Glinsky datasets, log$_2$ intensity matrices were loaded on MeV and adjusted by mean centering genes.

**Survival analysis**

Kaplan-Meier survival curves of patients groups were created using SPSS software. The log rank test was applied to examine the relationship between ESE3/EHF expression, biochemical progression-free and overall survival. Patients ($n=174$) included in the analysis had been treated with radical prostatectomy, had clinical follow-up data and evaluable ESE3 IHC staining. Biochemical recurrence was defined as a 0.2 ng/ml increase in PSA with a second confirmatory PSA measurement > 0.2 ng/ml or recurrence of disease after prostatectomy, such as development of metastatic cancer, if biochemical recurrence information was not available. For overall survival we included all the death occurred within 10 years after radical retropubic prostatectomy. Patients were censored at the time of the last clinical follow-up or at the date of death.

**References**


