**Supplementary Methods**

**Epigenetic silencing of microRNA-338/-421 drives SPINK1-positive prostate cancer**

Vipul Bhatia1, 5, Anjali Yadav1, 5, Ritika Tiwari1, Shivansh Nigam1, Sakshi Goel1, Shannon Carskadon2, Nilesh Gupta3, Apul Goel4, Nallasivam Palanisamy2, Bushra Ateeq1\*

1Molecular Oncology Lab, Department of Biological Sciences and Bioengineering,

Indian Institute of Technology Kanpur, Kanpur, 208016, U.P., INDIA

2Vattikuti Urology Institute, Department of Urology, Henry Ford Health System,

Detroit, MI 48202, USA

3Department of Pathology, Henry Ford Health System,

Detroit, MI 48202, USA

4Department of Urology, King George's Medical University,

Lucknow, 226003, U.P., INDIA

5These authors contributed equally to this work.

**Running Title:** Silencing of miR-338-5p/-421 drives SPINK1-positive cancer.

**Financial support:** This work is supported by the Wellcome Trust/ DBT India Alliance grant (IA/I(S)/12/2/500635 to BA).

**Conflict of interest:** The authors declare no conflicts of interest or disclosures.

**\*Corresponding Author and Lead Contact**:

Bushra Ateeq,

Molecular Oncology Lab,

Department of Biological Sciences and Bioengineering,

Indian Institute of Technology Kanpur, Kanpur, 208016, U.P., INDIA

Phone: +91 512 2594083, Fax: +91 512 2594010

Email: [bushra@iitk.ac.in](mailto:bushra@iitk.ac.in) (B. Ateeq)

**Supplementary material and methods**

**Data Mining and Computational Analyses:**

**Integrative MicroRNA target Prediction**

MicroRNA prediction programs, namely miRanda, miRMap, PITA and RNA Hybrid were used to predict miRNAs targeting 3’UTR of *SPINK1* (Fig. 1A and Supplementary Table S1). The correlation between the expression of the predicted miRNAs and *SPINK1* was analyzed by employing RNA-Seq data for the TCGA-PRAD cohort. For Fig. 1A (lower panel) and Supplementary Table S3 miRanda was used to predict the putative binding sites of the miR-338-5p/-421 on the 3’UTR of target genes.

**Integrative analyses for** **The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA-PRAD) data**

For gene association studies between miRs-338-5p, -421, *SPINK1*, *ERG*, and *EZH2* Illumina HiSeq mRNA and miRNA-Seq data along with clinical information was downloaded from TCGA-PRAD dataset. Overexpression of SPINK1 in PCa exhibits outlier-expression in ~10-15% of the total PCa cases (1). Thus, to stratify patients with increased expression of *SPINK1*, we arranged TCGA patients’ samples on the basis of highest *SPINK1* expressing specimen as top candidates (descending order) and divided the dataset into four “Equal Quartiles” by employing Quartile-based normalization method (2). The top 25% of the patients (n=119) corresponding to the upper quartile (QU, log2 (RPM+1)>5.468 were assigned as *SPINK1* high samples, and the lower quartile (QL, log2 (RPM+1) <1.124 were considered as *SPINK1* low samples. Also, we found about 18 patients with outlier expression of *SPINK1* with log2 (RPM+1) of greater than 11.984, which were included in the heat map representation of *SPINK1* positive TCGA patients in Fig. 1B. No further cut-offs were applied for miR-338-5p, miR-421 and *ERG* expression, corresponding expression values (based on *SPINK1* cut-off) were considered for these genes for further analysis. Hierarchical Clustering of miRs-338-5p, -421, *SPINK1,* and *ERG* were employed using heatmap.2 of R’s gplot package.

For Kaplan-Meier survival analysis, the survival data included sample type (primary tumors), days to first biochemical recurrence and days to last follow-up for TCGA-PRAD patients was considered. The samples were divided into two groups, higher and lower miRNA expression groups according to the expression level of a miR-338-5p/-421 using Cox proportional hazards regression model in R. Next, we performed a 13-year survival analysis of these miRNAs using Kaplan-Meier survival analysis (3) by employing survival package (https://cran.r-project.org/web/packages/survival), and statistical significance was computed using the log-rank test. For clinical relevance of miR-338-5p/-421, TCGA-PRAD dataset was analyzed for the association of these miRNAs with clinical parameters such as primary Gleason score. Data analysis was performed by one-way analysis of variance with Tukey’s post hoc test for multiple comparisons, and student’s *t*-test was applied for comparison between two groups.

The MSKCC cohort (Cancer Cell, 2010) data was retrieved from cBioPortal (http://www.cbioportal.org/) for *SPINK1* and *EZH2* expression in the prostate cancer patients (n=85), and oncoprints were generated using default parameters (mRNA expression z-score threshold ±2 vs normal). Further, to ascertain possible association between *EZH2*, *SPINK1* and miR-338-5p/-421, TCGA patients’ samples were stratified on the basis of increasing *EZH2* expression, and divided the dataset into four equal quartiles, the top 25% of the patients (n=119) corresponding to the upper quartile (QU, log2 (normalized count+1) >7.313), were considered as *EZH2* high patient samples and the lower quartile log2 (normalized count+1) <6.36), were considered as *EZH2* low samples. The corresponding expression values for *SPINK1,* miR-338-5p and miR-421 in *EZH2* high and *EZH2* low groups (without further cut-offs) were considered to association with *EZH2* expression (related to Fig. 5B).

**Cancer Cell Lines culture condition and authentication**

Prostate cancer cell lines (22RV1, VCaP and PC3), Colorectal (WiDr), Pancreatic (CAPAN-1), Melanoma (SK-MEL-173), prostate epithelial cells (RWPE-1) and human embryonic kidney 293T cells (HEK293T) were procured from the American Type Culture Collection (ATCC) and were maintained using ATCC recommended medium supplemented with 10% fetal bovine serum and Gibco Penicillin-Streptomycin (Thermo-Fisher). Cell lines were cultured in CO2 incubator (Thermo-Fisher) supplied with 5% CO2 at 37°C temperature.

To ensure the identity, short tandem repeat (STR) profiling of all cell lines were performed at the Lifecode Technologies Private Limited, Bangalore and DNA Forensics Laboratory, New Delhi. The profiles were compared with reference STR genotypes available at ATCC, DSMZ-German Collection of Microorganisms and Cell Cultures, and Biosample databases to authenticate the identity and check for any cross contamination. All cell lines were routinely tested for *Mycoplasma* contamination using PlasmoTest mycoplasma detection kit (InvivoGen).

**Establishing stable miRNA-overexpressing cell lines**

Human pre-miR-338-5p and pre-miR-421 (miRBase accession IDs MIMAT0004701 and MIMAT0003339, respectively) was PCR amplified from human genomic DNA and ligated to *XhoI/NotI* digested lentiviral miRNA-expression, pLemiR vector with turbo RFP (Open Biosystems). Sequences of the positive pLemiR constructs containing miRNAs were confirmed by Sanger’s sequencing. To establish stable miRNA-338-5p and miR-421 overexpressing cells, 22RV1 (SPINK1+ prostate cancer cell line) and WiDr (SPINK1+ colorectal cancer cell line) cells were transfected with 2 µg of miRNA cloned pLemiR construct using FuGENE HD Transfection Reagent (Promega) as per manufacturer's instructions. 48-hours post-transfection cells were selected in 1µg/ml puromycin for a period of 3-4 weeks, finally pooled population of miRNA overexpressing cells were validated for the overexpression of miRNAs by TaqMan assay, described elsewhere. Further, to obtain single clones overexpressing these miRNAs, single cells from the pooled population were seeded into each well of a 96-well plate, and were grown under puromycin selection for another 2-3 weeks. Finally, single clones obtained were evaluated for the overexpression of miRNA-338-5p and miR-421 using TaqMan probes and SPINK1 expression by using quantitative PCR.

**Transfection of microRNA mimetics, antagomiRs and small interfering RNA**

Synthetic miRCURY LNA (Lock Nucleic Acid) microRNA mimics for human miR-338-5p (HSA-MIR-338-5p Catalog No. 470322-001), miR-421 (HSA-MIR-421 Catalog No. 470788-001) and non-targeting control mimic (479903-001) were used for miRNAs overexpression. miRCURY LNA inhibitors specific to miR-338-5p, miR-421 and non-targeting control antagomiRs (anti-miRs) were used for silencing the expression of the respective miRNAs (Exiqon Product Catalog Number: 4100185-000, 4102114-000 and 199006-001 respectively). Briefly, the cells were seeded at 40% confluency and next day cells were transfected with respective miRNA mimics or antagomiRs, followed by a second transfection after 24 hours. MicroRNA mimics or antagomiRs were transfected using Lipofectamine RNAiMAX (Invitrogen) with a final concentration of 30pmol. Subsequently, 48 hours post-transfection, cells were processed for quantitative analysis and functional assays. Same transfection protocol was followed for On-Targetplus small interfering RNA (siRNA) for knockdown of *SPINK1* (J-019724-07, GE Dharmacon).

**Real-Time Quantitative PCR**

Briefly, total RNA was extracted using miRNeasy Mini Kit(Qiagen) for miRNA related experiments or else TRIzol (Ambion), and 1μg of RNA with good integrity was reverse transcribed into cDNA using SuperScript III (Invitrogen) in the presence of random primers (Invitrogen). For Real Time Quantitative PCR (qPCR) all reactions were performed in triplicates using SYBR Green Master Mix (Applied Biosystems). The relative expression of the target gene was calculated for each sample by using the ΔΔCt method as described before (4). Sequences for all the primer sets used in this study are listed in theSupplementary Table S4.

Quantitative PCR for the miRNA stem-loop was performed using target-specific stem-loop reverse transcription primers using TaqMan microRNA reverse transcription kit (Thermo Fisher), followed by Taqman assays (Applied Biosystems) following manufacturer’s instructions. Relative expression of the target miR-338-5p, miR-421, and miR-876-5p (Applied Biosystems Assays IDs: 4427975, 4427975, 4427975 respectively) was normalized to RNUB6 (Assay ID: 4427975).

**Cell proliferation, invasion and migration assays**

For cell proliferation assay, cells were seeded in 12-wells culture plates (10,000 cells/well). At the indicated time points cells were trypsinized and counted on the Z-Series Coulter counter (Beckman Coulter). Cell invasion assays were performed using Transwell Boyden chambers of 8μm pore size (Corning) (4). Briefly, RPMI-1640 media supplemented with 20% FBS was added to the lower compartment, and 100,000 cells in serum-free media were added onto Transwell insert coated with Matrigel (BD Biosciences). After 24 hours incubation at 37°C with 5% CO2, the invaded cells were fixed in formaldehyde (4% in PBS) and stained with crystal violet (0.5% w/v). Images of the representative field were taken on the Axio Observer Z1 microscope (Zeiss). The invaded cells were quantified by de-staining with 10% (v/v) acetic acid in distilled H2O, and absorbance of de-staining solution was measured at 550nm. Same protocol was followed for cell migration assay, except no Matrigel was coated on the inserts.

**Foci formation assay**

For foci formation assay, cells (2×103) were plated in six-well culture dishes in cell line specific recommended culture media supplemented with 5% heat-inactivated fetal bovine serum (Invitrogen) and incubated at 37ᵒC, media was changed every third day. The assay was terminated after 3 weeks and cells were fixed in formaldehyde (4% in PBS) and stained with crystal violet solution (0.05% w/v).

**Soft agar colony assay**

For anchorage-independent growth assay, soft agar plates were prepared by pouring 2ml of 0.6% low melting-point agarose (Sigma) in RPMI-1640 medium in 6-well dishes, after polymerization, second layer containing 2ml of 0.3% agar in RPMI-1640 medium and stable 22RV1-CTL, 22RV1-miR-338-5p and 22RV1-miR-421 cells (~1.5×104) were poured on the top of the first layer. Soft agar assay plates were incubated at 37ᵒC for 20 days, and colonies greater than 40μm in size were counted.

**Chick Chorioallantoic Membrane (CAM) assay**

The chick embryo CAM assay was performed as explained previously (5). Briefly, fertilized eggs were incubated in a humidified incubator at 38°C for 10 days. The CAM was released by applying low pressure to the hole over the air sac and the shell was cut to make a square 1cm2 windows. Two million cells (22RV1-miR-338-5p, 22RV1-miR-421 or 22RV1-CTL) were implanted near the allantoic vein onto the CAM in 10 days post-fertilized eggs. The windows were subsequently sealed, and the eggs were incubated at 38oC. For intravasation experiments, genomic DNA from lower CAM was isolated using Phenol/chloroform method and presence of the tumor cells was quantified by performing quantitative human *Alu*-specific PCR. The upper CAMs were isolated, fixed and immunostained for human-specific cytokeratin-18 as previously described (6). To assess tumor growth and metastasis, the assay was terminated on 18th day post-implantation and extra-embryonic tumor mass were excised and weighed. For metastases, embryonic lungs and liver was harvested, genomic DNA was isolated and subjected to quantitative human *Alu*-specific PCR as mentioned previously (5).

**Mice Xenograft Studies**

Mice were anesthetized using a cocktail of ketamine/xylazine (50 and 5 mg/kg respectively, via intraperitoneal route) and were subcutaneously implanted with 22RV1-CTL, 22RV1-miR-338-5p or 22RV1-miR-421 cells (2×106) suspended in 100μl of saline with 20% Matrigel into the dorsal both flank sides of the mice. A blinded assessment of tumor growth was conducted twice a week using digital Vernier’s calipers, and tumor volumes were calculated using the formula (π/6) (L × W2), (L=length; W=width). Spontaneous metastasis to lungs and bone marrow of the xenografted mice was analyzed by performing qPCR using primers specific for human specific *Alu*-sequences as mentioned in the Supplementary Table S4.

**Ki-67 staining of tumor xenografts**

Tumor tissues excised from the xenografted mice were fixed in 10% buffered formalin overnight, followed by dehydration using increasing concentration of ethanol. Subsequently, tumor specimens were embedded in paraffin and serially sectioned at 3µm thickness using microtome (Leica) as described earlier (4). Briefly, tissue sections were deparaffinized and dehydrated/rehydrated using standard protocol, followed by antigen-retrieval in the citrate buffer (pH 6.0) for 10 minutes at 100°C. Endogenous peroxidase activity was quenched using 3% hydrogen peroxide for 5 minutes. Sections were then blocked with 10% goat serum and probed with anti-mouse Ki-67 (1:400, CST, 9449S) at 4°C overnight, followed by secondary horseradish peroxidase (HRP)-conjugated antibody (DAKO), and HRP activity was detected using DAB (3, 3 -diaminobenzidine) peroxidase (HRP) substrate kit (DAKO). Quantification of IHC staining was performed in a blindfolded manner. The numbers of cells positive for Ki-67 staining were manually counted from ten random histological sections for each mouse.

**Gene expression array analysis**

Differentially regulated genes were clustered using hierarchical clustering based on Pearson coefficient correlation algorithm to identify significant gene expression patterns. Further, multiple hypotheses testing adjustments were applied using Benjamini and Hochberg procedure to calculate the FDR-corrected *P*-values (with FDR< 0.05) for the differentially expressed genes. Data was filtered to include only features with significant differential expression (log2 fold change greater than 0.6 or less than -0.6, *P*< 0.05) i.e. ~1.6-fold average over- or under-expressed genes, were then used for the deregulated biological processes using DAVID bioinformatics platform. Further, molecular signatures that were enriched upon miRNA overexpression with respect to control were analyzed using Gene set enrichment analysis (GSEA). A network-based analysis of critical miR-338-5p/-421 overlapping biological pathways was generated using Enrichment Map, a plug-in for Cytoscape network visualization software (http://baderlab.org/Software/EnrichmentMap/).

**Western Blot analysis**

Cell lysates were prepared in radioimmunoprecipitation assay (RIPA) lysis buffer, supplemented with complete protease (Roche) and phosphatase inhibitors mixture (Calbiochem). Protein samples were prepared in 1X SDS sample loading buffer; size fractionated on the SDS-PAGE and transferred onto a Polyvinylidene Difluoride membrane (PVDF) membrane (GE Healthcare). The PVDF membrane was then incubated for 1 hour at room temperature in blocking buffer [Tris-buffered saline, 0.1% Tween (TBS-T), 5% non-fat dry milk], and were incubated overnight at 4oC with the following primary antibodies: anti-phosphor -MEK or -ERK rabbit (1:1000, 9121S or 4377S) or total-MEK or –ERK (1:1000, 9126S or 4695S), anti-E2F1 rabbit (1:1000, 3742S), anti-TET1 rabbit (1:2000, ab121587) and anti-β-Actin rabbit (1:3000, 4970S). Subsequently, blots were incubated with horseradish peroxidase-conjugated secondary anti-mouse or anti-rabbit antibody (1:5000, Jackson ImmunoResearch Laboratories) for 2 hours at room temperature, and were washed with 1X TBS-T buffer, and the signals were visualized by enhanced chemiluminescence system (GE Healthcare) as described by the manufacturer.

**Immunofluorescence analysis**

Cells were grown on the glass coverslips and fixed in 4% para-formaldehyde, washed with 1X PBS, permeabilized with 0.3% Triton X-100 in PBS for 10 min and blocked with 5% goat serum in 0.1% Triton X-100 in 1X PBS for 2 hours at room temperature. Subsequently, cells were incubated with the following primary antibodies: SPINK1 (1:100, Abnova, H00006690-M01), ERG (1:200, Abcam ab92513), E-cadherin (1:400, CST, 3195S), N-cadherin (1:400, Abcam ab98952), Slug (1:50, CST, 9585S), Snail (1:50, CST, 3895S), c-Kit or CD117 (1:400, CST, 3308S), SOX-9 (1:400, Merck millipore, AB5535, a kind gift from Dr. A. Bandyopadhyay, IITK), TET1 (1:500, Abcam, ab121587). Subsequently, cells were incubated with Alexa Fluor-488 conjugated secondary anti-mouse or anti-rabbit antibodies (1:600, CST, 4412 or 4408). The coverslips with the stained cells were mounted on the slides using Vectashield with DAPI (Vector laboratories).

**Prostatosphere Assay**

In order to access any decrease in self-renewal properties and stemness, we performed prostatosphere in miR-338-5p and miR-421 overexpressing 22RV1 cells as compared to control. Briefly, single cell suspension of miRNA-338-5p/-421 overexpressing 22RV1 cells was seeded in low-adherence plates (Corning® 3471 Costar® 6-Well) at a density of 10,000 viable cells/mL in serum-free DMEM-F12 media (Invitrogen) supplemented with B27 (1:50, Invitrogen), 20 ng/ml EGF (Invitrogen), 20 ng/ml FGF (Invitrogen) and Penicillin-Streptomycin (Thermo-Fisher Scientific) as previously described (7,8). Small population of cells which formed prostatospheres were collected after 6 days by gentle centrifugation (~800 rpm), mechanically dissociated into single cells suspension and then re-plated under the same condition for several generations. Further, the sphere initiation/forming efficiency or self-renewal capacity was assessed by counting spheres larger than 50µm in diameter and plotted as percent sphere-forming efficiency for respective clones. ImageJ plugin was used to calculate the mean area occupied by the prostatospheres. Total RNA was isolated at day 6 and 12 as described before and was subjected to qPCR for detecting the expression of cancer stem cell like markers.

**Flow cytometry**

For BrdU-7AAD cell cycle analysis, 22RV1 cells were transfected with mimics for miR-338-5p, miR-421 or control for 2 consecutive days, followed by BrdU pulse labelling for 2 hours. Subsequently, cells were stained with anti-BrdU antibody conjugated to fluorescein isothiocyanate (FITC) (BrdU-7AAD flow kit, BD Biosciences), and then with 7-Aminoactinomycin D (7-AAD), following manufacturer’s instructions. Samples were subjected to FACS (BD FACS Calibur) excited at 488nm. Forward Scatter (FSC) and Side Scatter (SSC) parameters were adjusted to gate the population of interest. The 7-AAD signals were recorded on the linear scale while BrdU-\FITC signals were recorded on the logarithmic scale.

For Hoechst side population assay, 22RV1-CTL, 22RV1-miR-338-5p or 22RV1-miR-421 cells were stained with Hoechst 33342 (5 μg/ml), in the presence or absence of an ABC efflux pump inhibitor, verapamil (used as a negative control) followed by incubating samples at 37ᵒC in water bath for 2 hours with periodic agitation. Later, cells were centrifuged at 2000 rpm for 5 min at 4ᵒC and resuspended in cold 1X PBS and washed twice. Samples were kept at 4◦C until analysis. Propidium iodide (PI) was added at a concentration of 5μg/ml to exclude dead cells. For detection of side population (SP), Hoechst blue and red signals were acquired using a 460/50 and 670/30 nm band-pass filters respectively. While 7-AAD was excited at 488 nm and its emission was measured in logarithmic scale through a band pass filter of 670/30. Since, Hoechst Red signals are comparatively lower than that of Blue, a relatively higher laser power was used, and an optimal resolution of the SP cells was found using 30-35mW of power with the UV laser. A dim tail of SP cells enriched faction was gated using a dot plot displaying Hoechst Blue and Hoechst Red scatter. A minimum of 100,000 live cell events were acquired to resolve the SP cell population in each sample.

For Annexin PI staining, 22RV1 cells were transfected with mimics for miR-338-5p, miR-421 or control miRNA. After 24 hours, cells were washed with cold 1X PBS and resuspended in 1X Binding Buffer at a concentration of 1x106 cells/ml. Next, 5µl of FITC Annexin V and PI was added and incubated at room temperature for 15min. Subsequently, 400µl of 1X Binding Buffer was added to the samples and were analysed on BD FACSCalibur. Data acquisition was performed on BD FACSCalibur platform, and analysed using FlowJo version 10.7 (TreeStar).

**Immunohistochemistry**

Tissue microarray (TMA) slides comprising prostate cancer specimens (n=238) were incubated at 60°C for at least 2 hours. Slides were then placed in EnVision FLEX Target Retrieval Solution, High pH (Agilent DAKO, K800421-2) in a PT Link instrument (Agilent DAKO, PT200) at 75⁰C, heated to 97⁰C for 20 minutes, and then cooled to 75⁰C. Slides were then washed in 1X EnVision FLEX Wash Buffer (Agilent DAKO, K800721-2) for 5 minutes. Slides were then treated with Peroxidazed 1 (Biocare Medical, PX968M) for 5 minutes and Background Punisher (Biocare Medical, BP974L) for 10 minutes with a wash of 1X EnVision FLEX Wash Buffer for 5 minutes after each step. Mouse monoclonal SPINK1 (Novus Biologicals, H00006690-M01) diluted 1:100 in EnVision FLEX Antibody Diluent (Agilent DAKO, K800621-2) was added to each slide, which were then cover slipped with parafilm, placed in a humidifying chamber, and incubated overnight at 4⁰C. The next day, slides were washed in 1X EnVision Wash Buffer for 5 minutes and then incubated in Mach2 Doublestain 1 (Biocare Medical, MRCT523L) for 30 minutes at room temperature in a humidifying chamber. Slides were then rinsed in 1X EnVision Wash Buffer 3 times for 5 minutes each. Slides were then treated with a Ferangi Blue solution (1 drop to 2.5ml buffer; Biocare Medical, FB813S) for 7 minutes. Slides were rinsed 2 times in distilled water, then treated with EnVision FLEX Hematoxylin (Agilent DAKO, K800821-2) for 5 minutes. Slides were rinsed several times in distilled water, immersed in a 0.01% ammonium hydroxide solution, and then rinsed twice in distilled water. Slides were then dried completely. Slides were dipped in xylene approximately 15 times. EcoMount (Biocare Medical, EM897L) was added to each slide, which was then cover slipped.

**RNA *in situ* hybridization**

TMA slides were incubated at 60°C for 1 hour. Tissues were then de-paraffinized by immersing in xylene twice for 5 minutes each with periodic agitation. The slides were then immersed in 100% ethanol twice for 3 minutes each with periodic agitation, then air-dried for 5 minutes. Tissues were further treated with H2O2 for 10 minutes. Slides were rinsed twice in distilled water, and then boiled in 1X Target Retrieval for 15 minutes. Slides were rinsed twice in distilled water, and then treated with Protease Plus for 15 minutes at 40°C in a HybEZ Oven (Advanced Cell Diagnostics, 310010). Slides were rinsed twice in distilled water, and then treated with EZH2 probe (Advanced Cell Diagnostics, probe ID: 405491) for 2 hours at 40°C in the HybEZ Oven. Slides were then washed in 1X Wash Buffer (Advanced Cell Diagnostics, 310091) twice for 2 minutes each. Slides were then treated with Amp 1 for 30 minutes, Amp 2 for 15 minutes, Amp 3 for 30 minutes, and Amp 4 for 15 minutes, all at 40⁰C in the HybEZ oven with 2 washes in 1X Wash Buffer for 2 minutes each after each step. Slides were then treated with Amp 5 for 30 minutes and Amp 6 for 15 minutes at room temperature in a humidity chamber with 2 washes in 1X Wash Buffer for 2 minutes each after each step. Red color was developed by adding a 1:60 solution of Fast Red B: Fast Red A to each slide and incubating for 10 minutes. Slides were washed twice in distilled water. Amps 1-6 and Fast Red are included in the RNAscope 2.5 HD Detection Reagents-RED (Advanced Cell Diagnostics, 322360). Slides were then treated with EnVision FLEX Hematoxylin (Agilent DAKO, K800821-2) for 5 minutes. Slides were rinsed several times in distilled water, immersed in a 0.01% ammonium hydroxide solution, and then rinsed twice in distilled water. Slides were then dried completely. Slides were dipped in xylene approximately 15 times. EcoMount (Biocare Medical, EM897L) was added to each slide, which was then cover slipped.

**EZH2 and SPINK1 staining Evaluation Criteria**

EZH2 expression intensity scoring by RNA-ISH for all the tumor foci was evaluated on the basis of the number of red dots/cell and were graded into five levels ranging from score of 0 to 4 as described previously (9). SPINK1 staining by IHC was used to evaluate SPINK1 positive and negative status of the PCa specimens. Further, an association between SPINK1 and EZH2 expression in patients’ samples was calculated by applying Chi-Squared contingency test on GraphPad Prism.

**Chromatin immunoprecipitation**

Briefly, cancer cells (~80-90% confluency) were crosslinked with 1% formaldehyde for 10 minutes, followed by quenching with Glycine (125mM) for 10 minutes at room temperature, followed by washing with 1X PBS twice. Next, cell lysis was performed using lysis buffer [1% SDS, 10mM EDTA, 50mM Tris-Cl and protease inhibitor (Roche)] followed by sonication using Bioruptor (Diagenode) to obtain an average length of ~500bp DNA fragments. Chromatin immunoprecipitation (ChIP) assays were carried out using antibodies against EZH2/KMT6 (Abcam, ab191250), H3K27me3 (CST, 9733) and control rabbit IgG (Invitrogen). Supernatant containing sheared chromatin were incubated at 4°C overnight with 4µg of EZH2 or H3K27me3 and IgG antibodies. Concurrently, the Protein G coated Dynabeads (Invitrogen) were blocked with 100µg/ml BSA (HiMedia) and 500µg/ml sheared salmon sperm DNA (Sigma) and incubated at 4°C overnight. Blocked beads were washed twice with 9:1 dilution buffer: lysis buffer [1% Triton X-100; 150mM NaCl; 2mM EDTA (pH 8.0); 20mM Tris-HCl (pH 8.0) with protease inhibitors] and were incubated with respective antibodies to form antibody-bead conjugates. The antibody-bead conjugates were then washed three times in a low salt wash buffer 1 [1% Triton X-100; 0.1% SDS; 150mM NaCl; 2mM EDTA (pH 8.0); 20mM Tris-HCl (pH 8.0) with protease inhibitors] and once in high salt wash buffer 2 (same as wash buffer 1, except 500mM NaCl). The antibody/protein/DNA complexes were eluted using elution buffer [100mM NaHCO3, 1% SDS, RNaseA and Proteinase K (500µg/ml each)]. DNA was isolated using phenol-chloroform-isoamyl alcohol extraction method, precipitated and washed with 70% ethanol, air-dried, and dissolved in nuclease free water (Ambion). QPCR was performed using appropriate primer sets as listed in Supplementary Table S4.

**Methylated DNA Immunoprecipitation (MeDIP)**

Genomic DNA was extracted from 22RV1 cells using QIAamp DNA Mini Kit and was sonicated to produce random fragments ranging from 300-1000 bp. About 4µg of fragmented DNA was used for MeDIP assay. The DNA was denatured for 10 min at 95ᵒC and immunoprecipitated with 4µg of monoclonal antibody against 5‐mC (Abcam ab10805) or 5‐hmC (Abcam ab106918) and IgG (Santa Cruz, sc-2027) in a final volume of 500µl IP buffer (10mM sodium phosphate (pH 7.0), 140 mM NaCl, 0.05% Triton X‐ 100) for 2 hours at 4ᵒC. Dynabeads (40µl) were washed twice with 800 µl PBS-BSA (0.1%) for 5 minutes at room temperature, and then were resuspended in 40µl of 1X IP buffer. The resuspended Dynabeads were added to the samples and incubated for 4-5 hours at 4ᵒC with end-over-end shaking using rotator stirrer. Beads were then collected and washed thrice with 700µl of 1X IP buffer. The beads were treated with proteinase K (500µg/ml) for 3 hours at 50ᵒC, subsequently immunoprecipitated DNA was recovered by phenol‐chloroform extraction followed by ethanol precipitation. Real‐time PCR reactions were carried out with 40 ng of input DNA and 2µl of the immunoprecipitated DNA following manufacturer’s instructions.

**Bisulfite sequencing**

Bisulfite conversion of the genomic DNA was carried out using EpiTect Bisulfite Kit (Qiagen) following manufacturer’s instructions. Briefly, bisulfite converted DNA was used as template for PCR amplification using primers (Macrogen Inc., South Korea) designed using the Methprimer software as listed in Supplementary Table S4. The amplified PCR product was purified using QIAquick PCR purification kit (Qiagen), cloned into pGEM-T Easy Vector (Promega) and transformed into One Shot TOP10 competent cells (Invitrogen). Plasmid DNA was isolated from eight independent colonies and was outsourced for conventional Sanger sequencing at Macrogen Inc., South Korea. The BiQ Analyzer online tool was used to calculate the methylation percentage and to generate the graphical plots.

**References:**

1. Tomlins SA, Rhodes DR, Yu J, Varambally S, Mehra R, Perner S*, et al.* The role of SPINK1 in ETS rearrangement-negative prostate cancers. Cancer cell **2008**;13(6):519-28.

2. Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N*, et al.* A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Briefings in bioinformatics **2013**;14(6):671-83 doi 10.1093/bib/bbs046.

3. Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. Journal of the American statistical Association **1988**;83(402):414-25.

4. Tiwari R, Pandey SK, Goel S, Bhatia V, Shukla S, Jing X*, et al.* SPINK1 promotes colorectal cancer progression by downregulating Metallothioneins expression. Oncogenesis **2015**;4:e162 doi oncsis201523 [pii]10.1038/oncsis.2015.23.

5. Mira E, Lacalle RA, Gomez-Mouton C, Leonardo E, Manes S. Quantitative determination of tumor cell intravasation in a real-time polymerase chain reaction-based assay. Clinical & experimental metastasis **2002**;19(4):313-8.

6. Brenner JC, Ateeq B, Li Y, Yocum AK, Cao Q, Asangani IA*, et al.* Mechanistic rationale for inhibition of poly (ADP-ribose) polymerase in ETS gene fusion-positive prostate cancer. Cancer cell **2011**;19(5):664-78.

7. Johnson S, Chen H, Lo P-K. In vitro tumorsphere formation assays. Bio-protocol **2013**;3(3).

8. Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ*, et al.* In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes & development **2003**;17(10):1253-70.

9. Warrick JI, Tomlins SA, Carskadon SL, Young AM, Siddiqui J, Wei JT*, et al.* Evaluation of tissue PCA3 expression in prostate cancer by RNA in situ hybridization—a correlative study with urine PCA3 and TMPRSS2-ERG. Modern Pathology **2014**;27(4):609.