HORM CANC (2017) 8:28–48
DOI 10.1007/s12672-016-0279-9

ORIGINAL PAPER

Micro-RNA-204 Participates in TMPRSS2/ERG Regulation and Androgen Receptor Reprogramming in Prostate Cancer

Krassimira Todorova[1] · Metodi V. Metodiev[2] · Gergana Metodieva[2] · Milcho Mincheff[3] · Nelson Fernández[2] · Soren Hayrabedyan[1]

Received: 26 July 2016 / Accepted: 20 December 2016 / Published online: 3 January 2017

Abstract Cancer progression is driven by genome instability incurred rearrangements such as transmembrane protease, serine 2 (TMPRSS2)/v-ets erythroblastosis virus E26 oncogene (ERG) that could possibly turn some of the tumor suppressor micro-RNAs into pro-oncogenic ones. Previously, we found dualistic miR-204 effects, acting either as a tumor suppressor or as an oncomiR in ERG fusion-dependent manner. Here, we provided further evidence for an important role of miR-204 for TMPRSS2/ERG and androgen receptor (AR) signaling modulation and fine tuning that prevents TMPRSS2/ERG overexpression in prostate cancer. Based on proximity-based ligation assay, we designed a novel method for detection of TMPRSS2/ERG protein products. We found that miR-204 is TMPRSS2/ERG oncofusion negative regulator, and this was mediated by DNA methylation of TMPRSS2 promoter. Transcriptional factors runt-related transcription factor 2 (RUNX2) and ETS proto-oncogene 1 (ETS1) were positive regulators of TMPRSS2/ERG expression and promoter hypomethylation. Clustering of patients' sera for fusion protein, transcript expression, and wild-type ERG transcript isoforms, demonstrated not all patients harboring fusion transcripts had fusion protein products, and only few fusion positive ones exhibited increased wild-type ERG transcripts. miR-204 upregulated AR through direct promoter hypo-methylation, potentiated by the presence of ERG fusion and RUNX2 and ETS1. Proteomics studies provided evidence that miR-204 has dualistic role in AR cancer-related reprogramming, promoting prostate cancer-related androgen-responsive genes and AR target genes, as well as AR co-regulatory molecules. miR-204 methylation regulation was supported by changes in molecules responsible for chromatin remodeling, DNA methylation, and its regulation. In summary, miR-204 is a mild regulator of the AR function during the phase of preserved AR sensitivity as the latter one is required for ERG-fusion translocation.

Electronic supplementary material The online version of this article (doi:10.1007/s12672-016-0279-9) contains supplementary material, which is available to authorized users.

✉ Soren Hayrabedyan
shayrabedyan@ibir.bas.bg

[1] Institute of Biology and Immunology of Reproduction "Acad. Kiril Bratanov", Laboratory of Reproductive OMICs Technologies, Bulgarian Academy of Sciences, 73 Tsarigradsko shosse blvd, 1113 Sofia, Bulgaria

[2] School of Biological Sciences, University of Essex, Colchester, UK

[3] Cellular and Gene Therapy Ward, National Specialized Hematology Hospital, Sofia, Bulgaria

Introduction

Prostate cancer (PCa) is a leading cancer in men, with mortality rate only second to that from lung cancer. Prostate organ genesis is tightly controlled by androgens through androgen receptor (AR) signaling. AR is paramount for the lineage-specific differentiation of the prostate, inducing the expression of prostate-specific genes, such as prostate-specific antigen (PSA) and transmembrane protease, serine 2 (TMPRSS2), and maintaining the differentiated prostate epithelial phenotype [1]. AR is a nuclear transcription factor (TF) that could be modified by co-activators and co-repressors or binding by other TFs. It controls multiple target genes by binding consensus sequences within them termed as androgen response elements (ARE). Cancer genesis and progression originate from complex epigenome and genome instability causing intra- and inter-chromosome gene fusions, DNA methylation changes, and non-coding regulatory RNA perturbations. Genome rearrangement and epigenome deregulation result in complex AR

signaling reprogramming resulting in changes of AREs that AR binds to. Androgen deprivation therapy usually suppress oncogenic signaling, until new driver mutations are acquired and AR signaling becomes reinstated in condition referred to as castration-resistant PCa. At this point, AR is amplified and overexpressed [2, 3], mutated to lack ligand binding site, and its downstream signaling reprogrammed. Recent studies demonstrate that nearly half of the prostate cancers overexpress v-ets erythroblastosis virus E26 oncogene (ERG) by harboring various chromosomal rearrangements involving up to five exons of TMPRSS2 gene and majority of the exons of ERG gene, of which greater than 90% over expressed ERG [4, 5]. The resulting often *in-frame* fusion [6, 7] is under AR control using androgen-regulated TMPRSS2 promoter [8] and could also overexpress wild-type ERG binding to its locus [9, 10]. ERG gene is a part of the carcinogenesis regulatory E26 transformation-specific (ETS) transcription factor family, being the most commonly overexpressed proto-oncogene, found in 72% of the cases [11]. In androgen-dependent PCa harboring TMPRSS2/ERG, the overexpressed ERG disrupts AR signaling by inhibiting AR expression and directly binding to AR and inhibiting its target genes activation. ERG directly activates the histone H3K27 methyltransferase enhancer of zeste homolog 2 (*Drosophila*) (EZH2), a polycomb group (PcG) protein, inducing repressive epigenetic programs and facilitating a stem cell-like de-differentiation program. On the contrary, overexpression of AR or androgen induction therapy result in TMPRSS2/ERG downregulation [9].

The methylation may play an important role in regulating the TMPRSS2-ERG fusion. Fusion-negative tumors are heavily methylated when compared to fusion-positive samples, as found by methylated DNA immunoprecipitation sequencing [12]. Interestingly, PcG proteins can modulate the hyper-methylation of ERG promoters in prostate cancer cells, which also indicates the ERG gene is a hotspot of DNA methylation, especially for tumors with DNA methylation of ERG [13]. TMPRSS2/ERG activated PcG protein EZH2 directly controls DNA methylation of its targets by binding and recruitment of DNA methyltransferases (DNMTs).

In our previous study, we found that a tumor suppressor micro-RNA (miRNA, miR) termed miR-204 obtains pro-oncogenic properties in prostate cancer cell lines harboring the ERG fusion and that is also DNA methylation-dependent [14]. Micro-RNAs are a class of small non-coding RNAs that downregulate gene expression by hybridizing to complementary target mRNAs, resulting in either downregulation of translation or in mRNA degradation [15], thus serving as tumor suppressors. Upon cancer progression, genome instability and rearrangement could possibly dysregulate and overexpress miRNAs targeting tumor suppressors thus serving as cancer potentiating or onco-miRs [16, 17]. In prostate cancer cell lines like PC3 and DU145, miR-204 is overexpressed [18], compared to its slight expression in intact human prostate tissue [15]. Its basal expression is significantly increased in TMPRSS2/ERG fusion harboring VCaP and NCI-H660 cell lines compared to fusion-free LNCaP and PC3 cell lines [14]. We previously showed that dysregulated miR-204 loses its proliferation restriction, upregulating tumorigenesis important transcription factors ETS proto-oncogene 1 (ETS1) [19] and myeloblastosis proto-oncogene (c-MYB) [20], and bone metastasis important runt-related transcription factor 2 (RUNX2) [21] in TMPRSS2/ERG-positive PCa cell lines [14, 22]. ETS1 is the archetypical ETS factor of the TF family where ERG also belongs. ETS1 participates in a negative feedback loop with miR-204, as miR-204 resides within introns of ETS1 transactivated target gene [15], but it is at the same time ETS1 repressor [14]. RUNX2 and ETS1 also interact directly with the AR and regulatory loops between them and AR exist, promoting castration-resistant phenotype [19, 23]. Interestingly, both c-MYB and RUNX2 are positive miR-204 regulators, while miR-204 is their suppressor only in ERG fusion-negative cell lines, while in TMPRSS2/ERG fusion PCa cells it actually upregulated them.

Flinging that the inhibition of DNA (cytosine-5)-methyltransferase 1 (DNMT1) methyltransferase in PCa cells directly affects the regulatory properties of miR-204 in TMPRSS2/ERG-dependent or AR-dependent fashion, we sought to investigate miR-204 epigenetic regulatory properties on AR-regulated TMPRSS2/ERG, AR itself, AR-interacting TFs ETS1 and RUNX2 as well as AR replacing in castration-resistant phenotypes c-MYB.

In this report, we provide further evidence for an important role of miR-204 for TMPRSS2/ERG and AR signaling modulation, fine tuning and preventing TMPRSS2/ERG overexpression in prostate cancer. A novel method for detection of TMPRSS2/ERG protein products is provided. We demonstrate that miR-204 upregulates AR and downregulates TMPRSS2/ERG through direct regulation of their promoter methylation and set of TFs. Proteomics studies provide evidence that miR-204 has a dualistic role in AR cancer-related reprogramming.

Materials and Methods

Cell Lines

Vertebral bone metastasis-derived, AR+ (VCaP) and lymph-node metastasis-derived (LNCaP) prostate cancer cell lines were purchased from the ATCC (VA). VCaP cells were grown in DMEM media (ATCC Catalog No. 30–2002) (Sigma, St. Louis, MO) LNCaP cells were grown in RPMI media (ATCC,

Catalog No. 30–2001). All cell lines were cultivated in a humidified atmosphere at 5% CO2 and 37 °C.

Reagents and Transfection

MiScrpit miR-204 mimic and inhibitor (QIAGEN, Hilden, Germany) are synthetic RNAs resembling upon transfection of target cells the effects of mature miR-204. miR-204 inhibitor (anti-miR-204) and mimic (syn-miR-204), and their respective negative controls (MiScript Inhibitor Negative control and AllStars Negative siRNA) (QIAGEN, Hilden, Germany) were transfected into VCaP and LNCaP cells by HiPerFect (QIAGEN, Hilden, Germany) for 24 h, according to the manufacturer's instruction. The cells were seeded a day before transfection in 24-well plate for RT-qPCR, in 6-well plate for flow cytometry (FCS) using plastic ware from (TPP, Trasadingen, Switzerland). All experiments were done in triplicates.

RUNX2, *ETS1*, and *cMYB* Knockdown with Small-Interfering RNA

For small-interfering RNA (siRNA) knockdown of *RUNX2*, *ETS1*, and *cMYB*, we used human MISSION esiRNAs (Sigma, St. Louis, MO) against the target molecules with the following design:

- *RUNX2*: GGTACCAGATGGGACTGTGGTTACTGTCATGGCGGGTAACGATGAAAATTATTCTGCT
- *ETS1*: TGAGACCTTCCAAGGACAGCCGTGTTGGTTGGACTCTGAATTTTGAATTGTTATTCTAT
- *cMYB*: GGGCAGTAGAGCTTGGACAGAAAGAAAA- GAAACTTGGTGGTAGGTAATTGACTATGCA

AllStars Negative Control siRNA (5 nmol) and AllStars HsCell Death Control siRNA (5 nmol) were used as a control of transfection (QIAGEN, Hilden, Germany). VCaP and LNCaP cells were seeded a day before transfection with 80% confluence in a 24-well plate (TPP, Trasadingen, Switzerland) for RT-qPCR. Transfections were done according to the manufacturer's instruction with HiPerFect (QIAGEN, Hilden, Germany). Then, 200 nM of siRNA were used for transfection for 72 h. Experiments were done in triplicates. siRNAs knockdown of *RUNX2*, *cMYB*, and *ETS1* were tested by RT-qPCR.

Patients' Sera

The blood sera samples from diagnosed for advanced prostate cancer patients, staged III and IV, were supplied by the National Centre of Hematology and Transfusion Medicine (Sofia, Bulgaria) according to the rules of the Ethical Committee. Patients ranged from 49 to 80 years old, and averaged 65 years of age. None had received any chemotherapy during the course of the study.

Blood sera samples were collected from all 35 patients, and sera from healthy men and women were used as controls. All patients were routinely screened for the level of prostate-specific antigen (PSA) at the National Centre of Hematology and Transfusion Medicine (Sofia, Bulgaria). All sera were collected with the signed consent of the patients after they were informed about the aims of the study.

Real-Time Reverse Transcription Quantitative PCR Analysis

The expression of *TMPRSS2/ERG*, *ERG*, and *AR* genes was detected by RT-qPCR. After 24-h mimic or inhibitor and 72-h esiRNAs transfection, messenger RNA (mRNA) was isolated from VCaP and LNCaP cells using Midiprep Kit (QIAGEN, Hilden, Germany). From each sample, a 500 ng total RNA was used to synthesize complementary DNA (cDNA) by Sensiscript Reverse Transcription Kit (QIAGEN, Hilden, Germany). First-strand cDNA synthesis was performed according to the manufacturer's instruction (QIAGEN, Hilden, Germany) and 500 ng cDNA was used for PCR reactions (SYBR Green QuantiTect RT-PCR Master Mix, QIAGEN, Hilden, Germany). Total reaction volume was 50 μL. Primer concentrations were 0.5 μM (Fw and Re). RT-PCR Cycler (Agilent Technologies MX3005P, Stratagene, Santa Clara, CA) were used in this study. PCR reactions were started at 95 °C for 15 s, then 50 °C for 30 s, 72 °C for 30 s, followed by 45 cycles. The mRNA transcript expression levels of all studied genes were normalized towards transcript levels of endogenous reference gene glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). The following primer sequences, designed by us and produced by Biomers (Ulm, Germany) were used:

- *TMPRSS2/ERG*, variant 1 Fw: 5′-ctg gag cgc ggc agg aa-3′ Re: 5′-gta ggc aca ctc aaa caa cga ctg-3′
- *TMPRSS2/ERG*, variant 2 Fw: 5′-gat ggc ttt gaa ctc aga agc-3′ Re: 5′-tcc gta ggc aca ctc aaa caa-3′
- *TMPRSS2/ERG*, variant 3 Fw: 5′-tgg agc gcg gca ggt tat t-3′ Re: 5′-ttg tct tgc ttt tgg tca aca cg-3′
- *TMPRSS2/ERG*, variant 4 Fw: 5′-gga gcg cgg cag gaa-3′ Re: 5′-gtt cat ccc aac ggt gtc tgg-3′
- *ERG* isoform 1 Fw: 5′- cgc aga gtt atc gtg cca gca gat-3′ Re: 5′-cca tat tct ttc acc gcc cac tcc-3′
- *ERG* isoform 2 Fw: 5′- agc tac aac gcc gac atc c-3′ Re: 5′-gaa gtc aaa tgt gga aga gga gtc-3′
- *GAPDH* Fw: 5′-aag gtc gga gtc aac gga ttt-3′ Re: 5′-acc aga gtt aaa agc agc cct g-3′
- *AR* Fw: 5′-cgc tga agg gga aca gaa gta- 3′, Re: 5′-tct cct tcc tcc tgt agt ttc- 3′

Methylation-Specific RT-qPCR

VCaP and LNCaP cells were seeded in 24-well plate (TPP) on the day before transfection with miR204 mimic (QIAGEN) using HiPerFect (QIAGEN) with established protocol by manufacturer. control and transfected cells were collected by trypsinization and centrifugation. Then, the cells were treated with reagents from Methylamp Whole Cell Bisulfite Modification Kit (Epigentek, USA). Modified DNA was eluted and RT-qPCR (EpiQuik Quantitative PCR Fast Kit, Epigentek, USA) was performed to measure the expression levels of AR in methylated cells. The program started at 95 °C for 7-min activation, 95 °C for 10 s, 55 °C for 10 s, 72 °C for 8 s followed 45 cycles and final extension on 72 °C for 1 min. Experiments were done in duplicates.

The primer sequences amplifying methylated and non-methylated bisulfite converted sequence were designed using MethylPrimer (ABI, USA) and produced by Biomers (Germany). The targeted CpG islands belonged to *TMPRSS2* (chr21:42836478:42882086:-1) promoter region and region spanning between 5′ UTR and exon-1 of androgen receptor (chrX:66763874:66950461). The quantification cycle (Cq) was determined for each reaction with methylation-specific primers (MSP) and bisulfite-specific primers (BSP) and the ratio of unmethylated to total amplifiable bisulfite-treated DNA was calculated by methylation-specific RT-qPCR (MS-qPCR) adopted ΔΔCq method, using fully methylated and non-methylated reference DNA templates [24]. The relative expression was assessed for non-targeting control and miR-204 mimic transfected LNCaP and VCaP cells.

MS-qPCR Specific Primers

- MF_TMPRSS2:ERG 5′-gta cgt ttc gag ggt ttt gac-3′
- MR_ TMPRSS2:ERG 5′-aat aaa ccc gaa acc ccg-3′
- UF_TMPRSS2:ERG 5′-tta gta tgt ttt gag ggt ttt gat-3′
- UR_TMPRSS2:ERG 5′-aaa aat aaa ccc aaa acc cca-3′
- MF_AR 5′-ggg ttt tag gta ttt aga ggt cgc-3′
- MR_AR 5′-acc aaa taa cct ata aaa cct cta cg −3′
- UF_AR 5′-ggg ttt tag gta ttt aga ggt tgt-3′
- UR_AR 5′-acc aaa taa acc tat aaa acc tct aca-3′

Flow Cytometry

Control and treated (miR-204 mimic or miR-204 inhibitor) prostate cancer VCaP and LNCaP cells were detached with Accutase (eBioscience, Frankfurt, Germany) and washed with cold 1% BSA-PBS. Specific antibodies (Abs) (Santa Cruz Biotechnology, Dallas, TX) were used for detection of AR. After IC fixation/permeabilization solution (eBioscience, Frankfurt, Germany) wash, and flow cytometry staining buffer (eBioscience, Frankfurt, Germany), the specific mouse monoclonal anti-AR (Santa Cruz Biotechnology, Dallas, TX) or the appropriate isotype control Abs were used at concentration of 0.5 mg/10^6 cells for 60 min on ice, followed by BSA-PBS wash and secondary antibody (rabbit anti-mouse FITC conjugated IgG, Santa Cruz Biotechnology, Dallas, TX) incubation at 0.25 mg/10^6 cells for 30 min on ice (in the dark). Cells were gated using forward vs. side scatter to exclude dead cells and debris. After washing, cells were analyzed with a BD FACSCalibur flow cytometer (Becton Dickenson, Franklin Lakes, NJ). Fluorescence of 10^4 cells per sample was acquired in logarithmic mode for visual inspection of the distributions and for quantifying the expression of the relevant molecules by calculating the median fluorescence intensity (referred to as MFI) in a histogram overlay graphics.

Total Demethylation

VCaP and LNCaP cells were seeded in 24-well plate (TPP, Trasadingen, Switzerland) for RT-qPCR and in 6-well plates for FCS. The cells were treated with 6 μM 5-Azacitidine (Sigma, St. Louis, MO) (5-AzaC) for 4 days. The media was changed every day with fresh 5-AzaC solution. Briefly, total RNA was isolated from the cells and cDNA was synthesized following the same protocol described above. RT-qPCR was performed to measure the expression of *TMPRRS2/ERG* and *AR*. Using standard protocol described above, FCS assay was performed to evaluate the effect of demethylation in AR expression in VCaP and LNCaP cells.

Proximity Ligation Assay-Based Detection of TMPRSS2/ERG Fusion Protein Products

Proximity ligation assay (PLA)-based method for detection TMPRSS2/ERG protein product used polyclonal antibody that target epitope near the N-terminus of human TMPRSS2 and second polyclonal antibody that target epitope at the C-terminus of human Erg-1: TMPRSS2 Antibody (N-13): sc-19686, goat polyclonal IgG (stock concentration 200 μg/ml); Erg-1/2/3 Antibody (C-20): sc-353, rabbit polyclonal IgG (stock concentration 100 μg/ml).

The antibodies were conjugated to specific oligoes allowing for qPCR amplification-based detection. PLA is based on the detection by two antibodies of close molecular distance epitopes and subsequent amplification of the signal by looping of the conjugated to the antibodies oligoes and further second-level amplification of the loop sequence using standard qPCR. The expression level is calculated using regression of Ct values of PLA qPCR amplification of serial dilution of protein lysate of TMPRSS2/ERG fusion gene harboring VCaP cells.

The probes for conjugating TMPRSS2 antibody and ERG antibody were created using Proseek oligo probes with Assay

Solution according to the Olink Bioscience instruction (Proseek, Assay Development Kit, Olink Bioscience). Polyclonal IgGs were pre-concentrated to 1 mg/ml using Millipore concentration columns. Serial dilutions of VCaP cell lysate and oligo-conjugated antibodies were mixed and the extension master mix (contains polymerase 5 U/μl) were added to the wells. This process extended the looping oligoes of the two antibody-based probes that were bound to a target protein through a DNA polymerization event, creating the RT-PCR amplification. The following extension program was used: polymerization (100 μL) = 20 min, 37 °C; inactivation of enzyme (100 μl) = 10 min, 85 °C; cooling = 4 °C. RT-PCR master mix (polymerase = 1 U/μm) was prepared. Master mix (6 μl) and 4 μl of each extension product were mixed in a new optical 96-well plate. DNA amplification was done, using standard real-time PCR instrument: 95 °C, 5 min; 95 °C, 15 s and 60 °C, 1 min and 45 cycles.

Control and transfected (miR-204 mimic and inhibitor; esiRNAs) VCaP cells were detached with Accutase (eBioscience, Frankfurt, Germany) and prepared for serial dilutions in 96-well plate. LNCaP cells were used as a negative control.

Sera from prostate cancer patients (n = 35) and sera from healthy men and women were diluted 1:10 and 1:100. All sera samples were used as test samples following the same protocol.

Quantitative Proteomics

Protein samples were prepared and analyzed in triplicate as described in Metodieva et al. [25]. Briefly, the analysis of protein digests by electrospray ionization MS was performed on a hybrid linear trap quadrupole (LTQ)/Orbitrap Velos instrument (Thermo Fisher) interfaced to a split-less nano-scale HPLC (Ultimate 3000, Dionex). The samples were automatically injected by the autosampler and desalted and concentrated online at a flow of 1 ml min^{-1} on a 2-cm-long, 0.1 mm i.d. trap column packed with 5 mm C18 particles (Dionex). The peptides were then eluted from the trap column and separated in a 90-min gradient of 2–30% (v/v) acetonitrile in 0.1% (v/v) formic acid at a flow rate of 0.3 ml min^{-1}. A 15-cm-long, 0.1 mm i.d. pulled tip packed with 5 mm C18 particles was used as separation column (Nikkyo Technos Co., Tokyo, Japan). The eluting peptides were electrosprayed into the LTQ/Orbitrap Velos mass spectrometer by applying a voltage of 1.75 kV via a liquid junction interface. The LTQ/Orbitrap Velos was operated in the Top20 data-dependent mode with two high-resolution scans (resolution of 30,000 at 400 m/z) followed by 20 MS/MS scans for the 20 most abundant peptide ions having a charge state >1. During the high-resolution scans, the Orbitrap analyzer was set to accumulate 10^6 ions for the maximum of 0.5 s. During MS/MS scans, the LTQ was set to accumulate 5000 precursor ions for the maximum of 0.1 s. The normalized collision energy was 30; minimum signal intensity required was 500; activation time was 10 ms; and activation Q was 0.250. A dynamic exclusion was implemented for 30 s. Internal mass calibration was implemented by means of lock mass using the ambient ion of 445.12 m/z.

Bioinformatics

Bioinformatics Analysis of Proteomics Primary Data

All LC-MS/MS data were processed by MaxQuant [26] using the latest International Protein Index (IPI) fasta file for protein identification. Relative abundance was estimated by the spectral counting method [27]. Spectral counts of proteins with probability of detection over 75% were used for relative fold change (FC) expression assessment. Generally, a log_2FC of the averaged total per protein spectral counts was estimated, considering the uniqueness and coverage of detection. Data were log-transformed, missing values imputed, and t tests for all detected proteins, and calculates adjusted p values using the false discovery rate algorithm [28].

Bioinformatics Analysis of the differentially Expressed Proteomics Data

Gene inter-dependency differential expression and pathway analysis: We selected differentially expressed genes, based on pathway enrichment that considers inter-dependency between proteins, the size, role, and position of each gene on the pathway when modeling high-throughput expression data, and based on the fold change of individual genes, adjusts the model of its downstream targets using iPathwayGuide (www.advaitabio.com). This advanced approach allows for quick targets and pathways prioritization, avoiding false positive and negative results. iPathwayGuide identifies significantly impacted pathways based on two forms of evidences: over representation and accumulated perturbation analyses. We used iPathwayGuide analysis of differential expression to plot -log10 (adjPVal) against logFC (x-axis) based on the log_2 (miR-204 mimic/non-targeting mimic analog spectral counts) and p value (Student t test corrected by ad hoc Bonferroni) (Supporting Information Fig. S8).

Differential expression of proteins based on equal probability of expression: using spectral counts preprocessed data, were assessed using log-fold change (logFC) with significant probability. The expression data was plotted using a derivative of the MA Plot FCDEGUST (www.vicbioinformatics.com/degust/) that calculates the differential expression analysis using t test (R framework, voom/limma package), plotting MA plot (Supporting Information Fig. S9).

EGAN: Exploratory Gene Association Networks (*http://akt.ucsf.edu/EGAN/*) is used to build an interactive hypergraph of genes, relationships (protein-protein

interactions, literature co-occurrence, etc.) and meta-data (annotation, signaling pathways, etc.). miR-204 overexpression modulated proteins in LNCaP were analyzed for meta-data coincidence (over-representation, enrichment) for multiple annotation (NCBI Entrez Gene, PubMed, KEGG, Gene Ontology, iHOP, Google, etc.). We have supplied *t* test compared differential expression data with custom gene sets annotations downloaded from Molecular Signatures Database (MSigDB) at Broad Institute related to androgen receptor signaling and DNA methylation regulation. The pathways were enriched using the EGAN implementation of gene set enrichment analysis (GSEA) algorithm, based on the logFC converted data, ordered by probability (PermuSEED q-val, or FDR <0.05).

TMPRSS2 Multiple Protein Sequence Alignment

Nucleotide and amino acid sequences of all human TMPRSS2 and ERG isoforms as well as nucleotide and available amino acid sequences of different fusion variants of TMPRSS2/ERG were downloaded from NCBI and UNIPROT using UGENE (ugene.net). Sequencing data from VCaP cell line TMPRSS2/ERG were used to generate synthetic fusion of TMPRSS2 and ERG genomic versions as well. The sequences "aaaactccaggaggttaggactgca" from *TMPRSS2* and "tacattgacctatttggagtc" from *ERG* are found fused together using single nucleotide resolution Illumina next-generation sequencing [7]. Constructed sequence was used along with others for multiple sequence alignment–MSA (MUSCLE algorithm) as well as structural modeling of TMPRSS2/ERG protein product folding. ORF analysis was also done for elucidating the consensus sequences belonging to an encoding region in the fusions.

TMPRSS2 Epitope Mapping

Full protein sequences of human TMPRSS2 and ERG were subject to extensive B cell epitope prediction analysis using IEDB (http://tools.iedb.org/bcell). The following algorithms were used: Bepipred Linear Epitope Prediction (predicts the location of linear B cell epitopes using a combination of a hidden Markov model and a propensity scale method), Chou and Fasman beta turn prediction (predict beta turns), Emini surface accessibility scale (surface accessibility), Karplus and Schulz Flexibility Prediction (mobility of protein segments on the basis of the known temperature B factors of the a-carbons), Kolaskar and Tongaonkar Antigenicity (semi-empirical method which makes use of physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes), and Parker Hydrophilicity Prediction. Predicted peptide sequences were aligned to MSA of fusion variants vs. native sequences.

TMPRSS2 de Novo Structural Modeling

Using published, sequencing detected fusion variant of TMPRSS2/ERG and synthetic construct (genome sequences of TMPRSS2 and ERG extracted and "fused" based on fusion data of VCaP), in silico modeling of TMPRSS2/ERG fusion was conducted. Two approaches were used—homology modeling combining also ab initio modeling implemented by Phyre-2 (http://www.sbg.bio.ic.ac.uk/phyre2) and homology combined with de novo modeling implemented by I-Tasser (http://zhanglab.ccmb.med.umich.edu/I-TASSER/). Phyre-2 models represented TMPRSS2/ERG NCBI sequence lacking ETS-DNA binding site, while I-Tasser models represented TMPRSS2/ERG containing ETS-DNA binding site.

Statistical Analysis

The data were generated from three independent experiments, each performed in triplicates. ANOVA test with respective multiple comparison post-tests (Greenhouse-Geisser correction for non-data sphericity, Tukey correction post-test, adjusted p value and family-wise significance, confidence level of 0.05) was used to analyze the data (GraphPad Prism 6, La Jolla, CA). $p < 0.05$ was considered significant.

Results

TMPRSS2/ERG Fusion Product and Native TMPRSS2 and ERG Are Predicted to Have Putative Shared Epitopes, Based on Sequence Alignment and de Novo Structural Models, Allowing for Proximity Ligation Assay Detection

Although TMPRSS2/ERG fusion is considered to produce an increased high level of ERG protein in many prostate cancer patient cases [29], there are other patients having different rearrangement variants of the fusion that are either non-sense or frame-shift, or containing more than just the first TMPRSS2 exon. For that very reason, so far, there is no practical method to measure for the actual full TMPRSS2/ERG protein product. Therefore, we used multiple sequence alignment (MSA) of native TMPRSS2 and ERG isoform protein sequences, together with several reported TMPRSS2/ERG mutant variants detected in cancer cases, to seek for shared amino acid sequences between the native and fusion forms. We additionally added synthetic TMPRSS2/ERG protein sequences resembling fusions of TMPRSS2 and ERG as described [6, 30–35], using native genes genomic sequences and published fusion points sequences [7]. In TMPRSS2, we found three N-terminal sub-regions that partially shared common sequences, while in ERG C-terminal region, we found

shared full amino acid sequences matching between cancer and native variants (Fig. 1a). Then, we used four algorithms for B cell epitope prediction to assess the probability whether the shared between native and mutant TMPRSS2 and ERG regions could harbor amino acid sequences that could be detected by antibodies specifically raised to these genes native protein products N- and C-terminus, respectively (Fig. 1b). We found high probability of putative epitope availability. We used the most probable amino acid sequences predicted as putative epitopes and using MSA mapped them to the TMPRSS2 and ERG sequences, respectively (Fig. 1c, Supporting Information Fig. S1). Using a combination of template-based and de novo modeling, we created in silico a putative TMPRSS2/ERG protein molecular models and mapped the predicted consensus epitopes probing the molecular distance between them (Fig. 1d). The models suggested that the epitopes are located in high enough proximity for eventual antibody detection based on proximity ligation assay method.

Development of Proximity Ligation-Based Assay for Assessment of Protein Products of TMPRSS2/ERG Fusion in VCaP Cells and Patients' Sera

In order to detect TMPRSS2/ERG protein product, we developed a proximity ligation assay (PLA)-based method, using polyclonal antibody that target epitope near the N-terminus of human TMPRSS2 and second polyclonal antibody that target epitope at the C-terminus of human ERG-1. The antibodies were conjugated to specific oligoes allowing for qPCR amplification-based detection (Fig. 1e). PLA is based on the detection by two antibodies of close molecular distance epitopes and subsequent amplification of the signal by looping of the conjugated to the antibodies oligoes and further second-level amplification of the loop sequence using standard qPCR. The expression level is calculated using regression of Ct values of PLA qPCR amplification of serial dilution of protein lysate of TMPRSS2/ERG fusion gene harboring VCaP cells (Fig. 1f). Protein lysate of TMPRSS2/ERG fusion gene-free LNCaP cells was used as a negative control, rendering no amplification (Fig. 1e).

The micro-RNA-204 Is a Negative Regulator of TMPRSS2/ERG Oncofusion

We followed the effect of miR-204 in prostate cancer cells harboring the *TMPRSS2/ERG* fusion. We deliberately used previously published primer sets detecting *TMPRSS2/ERG* fusion transcripts in prostate cancer patients [36], that were proven specific until Ct 30 (Fig. 2a, b). Our results showed that miR-204 artificial overexpression using synthetic mimic resulted in significant downregulation of *TMPRRS2/ERG* transcript in VCaP cells and vice versa, miR-204 suppression using synthetic complementary inhibition produced significant increase in *TMPRRS2/ERG* mRNA levels (Fig. 2c). Similarly, miR-204 overexpression downregulated wild-type isoforms of *ERG* mRNA (Supporting Information Figs. S2, S3).

Using the proximity ligation assay (PLA) for protein detection, we followed up the oncofusion expression in VCaP cells after miR-204 synthetic analog transfection. The results were compared to VCaP cells transfected with non-targeting control. We found similar to mRNA level miR-204 incurred significant suppression of TMPRSS2/ERG protein product by 80%. No TMPRSS2/ERG protein product expression was detected in LNCaP cell line control. The data were represented as relative fold change of TMPRSS2/ERG expression compared to mean expression in transfected with non-targeting RNA VCaP cells, where the expression was 100% (Fig. 2d).

Inhibition of DNA Methyltransferase Abrogates miR-204 Mediated TMPRSS2/ERG Downregulation

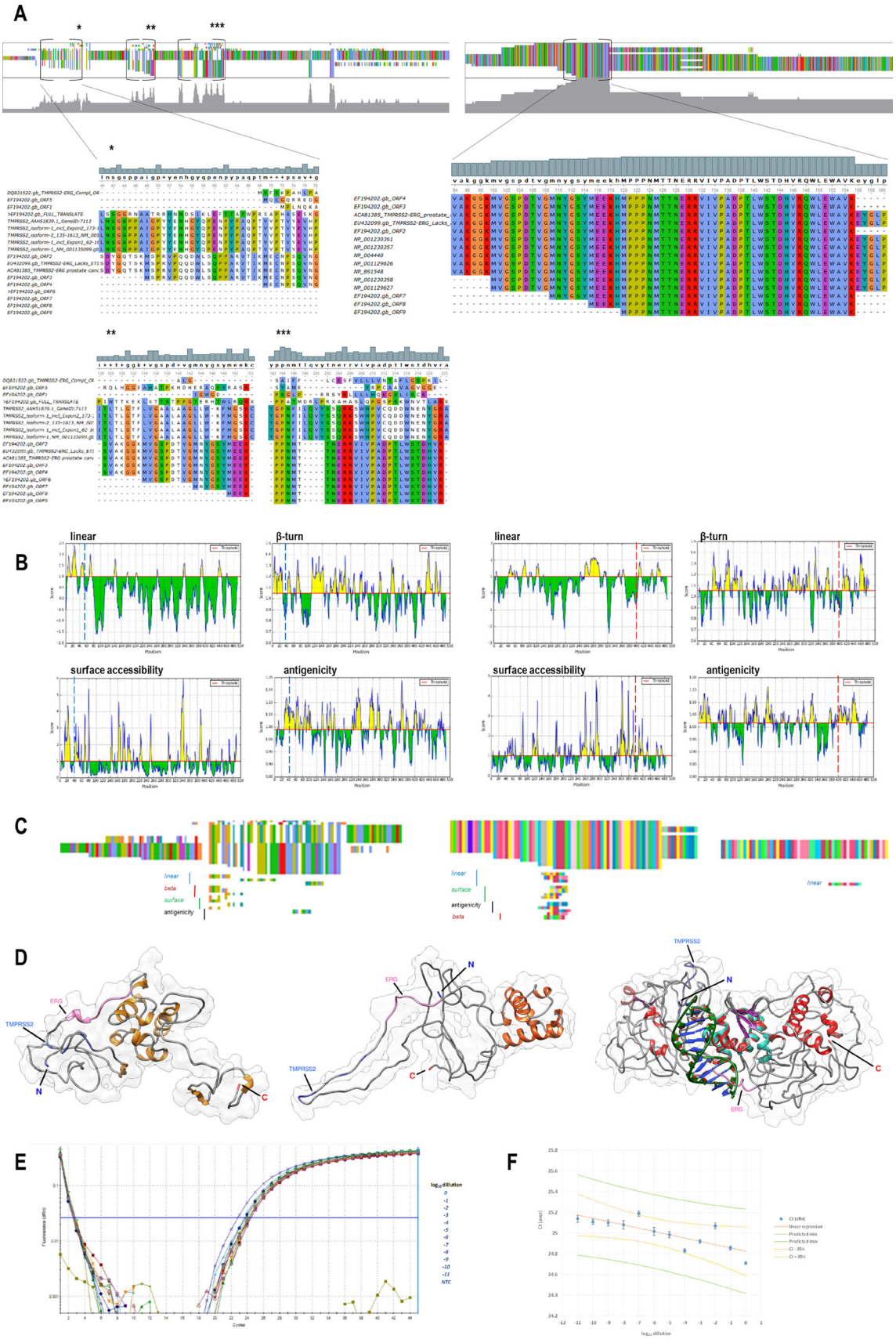
Total genome demethylation by inhibiting DNA (cytosine-C5) methyltransferase (DNMT) using 5-azacytidine resulted in abrogation of the effect of miR-204 overexpression in VCaP cells on the *TMPRSS2/ERG* mRNA downregulation (Fig. 2e), suggesting gene methylation as possible mechanism of miR-204 action. Similarly, 5-azacytidine alone was able to upregulate the fusion gene mRNA compared to its basal levels. Conversely, the overexpression of gene fusion mRNA caused by a miR-204 inhibitor sequence was not affected by total demethylation (Fig. 2e).

miR-204 Increases the Level of DNA Methylation of *TMPRSS2/ERG* Promoter

Following previous analysis results, we investigated *TMPRRS2/ERG* promoter methylation status. Transfection of miR-204 mimic resulted in promoter hyper-methylation (up to 90%), compared to control condition, where transfection with non-targeting RNAs resulted in CpG islands methylation of only 39.3% (Fig. 2f). This result suggested that miR-204 negatively regulates *TMPRSS2/ERG* expression through its promoter hyper-methylation.

Master Transcriptional Factors RUNX2 and ETS1 Are Positive Regulators of TMPRSS2/ERG Oncofusion

We have previously shown that transcriptional master regulators involved in prostate cancer progression and metastasis like *MYB*, *RUNX2*, and *ETS1* are regulated by miR-204 in *TMPRSS2/ERG* occurrence fashion [14]. Transitory siRNA mediated silencing of *MYB*, *RUNX2*, and *ETS1* in the VCaP cell line resulted in significant downregulation of the *TMPRSS2/ERG* transcripts (Fig. 2g) and a corresponding


A
B
linear
β-turn
linear
β-turn
surface accessibility
antigenicity
surface accessibility
antigenicity
C
linear
beta
surface
antigenicity
linear
surface
antigenicity
beta
linear
D
TMPRSS2
ERG
N
C
E
F

◀ **Fig. 1** TMPRSS2/ERG sequence variants and putative molecular model potentially sufficient for proximity ligation assay-based detection of fusion protein products. **a** MSA using MUSCLE algorithm (UGENE) of TMPRSS2/ERG sequenced and synthetic variants vs. native TMPRSS2 isoforms (*left*) and MSA of TMPRSS2/ERG vs. native ERG isoforms (*right*). Several consensus regions projected to show full sequences. *Top* overlaid histogram and nucleotide *colors* depict consensus level. **b** B cell putative epitopes predicted for full length sequences of TMPRSS2 (*left*) and ERG (*right*), showing protein sequence position (*x*) vs. score over cutoff value (*y*) shaded in yellow. Used algorithms are Bepipred Linear Epitope (linear), Chou and Fasman (beta turn), Emini (surface accessibility), Kolaskar and Tongaonkar (Antigenicity). Positions cutting off MSA consensus sequences are depicted in *blue* and *red*, accordingly. **c** MSA using MUSCLE algorithm of fusion, native and high probability epitope predicted amino acid sequences. (*Scale* UGENE overview). **d** Phyre-2 (*left*) and I-Tasser (*center*) models of patient sequenced TMPRSS2/ERG variant lacking ETS-DNA binding site. Crystallography model of native uninhibited DNA binding domain of ERG (PDB 4IRG) and dsDNA were aligned in the model. Phyre2 (*right*) model of synthetic TMPRSS2/ERG variant 1, produced out of genomic native genes sequences. Epitope mapping on TMPRSS2 of patient fusion: GSP.VGM.YG.YM or GSP ... G.Y (*medium blue*), epitope mapping of ERG: YMEEKHMPPPNM or EEKHMPPPNM (*hot pink*). Epitope mapping on synthetic fusion model—TMPRSS2: GSPP.IGP.Y.H (*medium blue*), ERG: VPAD ... RETP (*hot pink*). **e** Proximity ligation assay (PLA) for detection of TMPRSS2/EGR protein products. PLA amplification plots of VCaP protein lysate serial dilutions, generated by qPCR amplification of the proximity loop generated when antibody-oligo conjugates detecting TMPRSS2 and ERG were in close proximity. **f** Linear regression of averaged Ct ($n = 3$) vs. $\log_{10}$ lysate dilution. *Yellow* 95% confidence interval (CI)

decrease of its protein product expression (Fig. 2h, i), implying involvement of these transcription factors in *TMPRSS2/ERG* regulation. Additionally, wild-type *ERG* isoforms were also downregulated when *RUNX2* and *ETS1* were transitory silenced (Supporting Information Fig. S4).

Master Transcriptional Factors Maintain Decreased Level of DNA Methylation of *TMPRSS2/ERG* Promoter

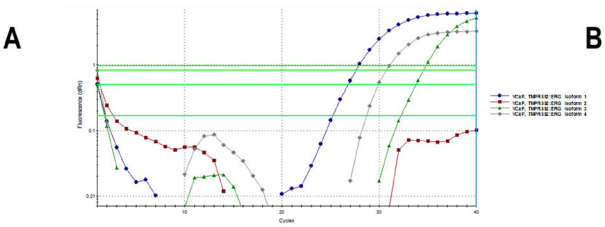
Trying to elucidate the mode of regulation transcription factors *RUNX2* and *ETS1* exert on *TMPRSS2/ERG*, we found that transitory silencing of *RUNX2* in VCaP cell line resulted in significant fusion gene promoter DNA hyper-methylation up to 91%, compared to only 36% methylation of CpG islands after negative control non-targeting RNAs transfection. Similar silencing of *ETS1* resulted in much lesser increase in DNA methylation of *TMPRSS2/ERG* promoter of only 45%, compared to intact 36% (Fig. 2j).

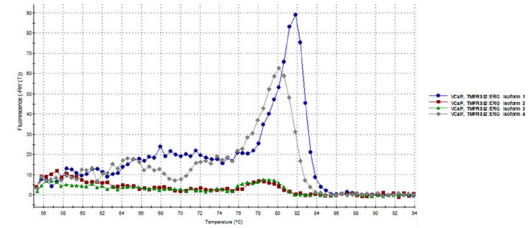
TMPRSS2/ERG Transcript and Protein Expression in Prostate Cancer Patients' Sera

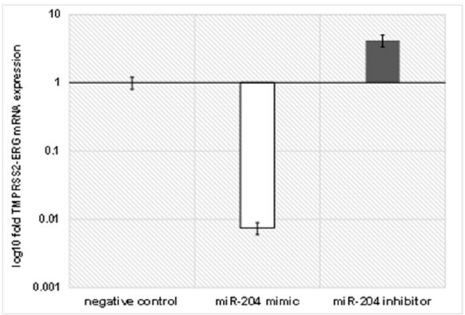
Sera from 35 patients with prostate cancer were investigated for TMPRSS2/ERG transcript and protein product expression. Sera collected from healthy men and women were used as negative controls. As a positive spike control, a VCaP cell line lysate was added to healthy donor serum. *TMPRSS2/ERG* mRNA was detected using specific primers previously described elsewhere in another group's study [36], able to detect four major variants of *TMPRSS2/ERG* rearrangement. A specific amplified product was detected up to 30 RT-qPCR cycle, and amplified product after 30 cycles were accepted as *non-specific phenomenon*, as described above [36]. There were specific melting and amplification curves for prostate cancer patients' sera and non-specific for sera from healthy donors (Fig. 3a, b). Dissociation curves had exhibited high peaks for patients' sera and very low peaks in healthy donor serum samples. Additionally, the transcript levels of two native ERG isotype mRNAs were examined in the same sera. In VCaP cells, two of four major variants described in study above were expressed—variant I (TMPRSS2-exon 1/ERG exon 4) and variant IV (TMPRSS2-exon 1/ERG exon 5). We further investigated TMPRSS2/ERG expression in patients' sera following variant 1 as most common, and the expression was assessed as relative expression compared to standard VCaP cell lysate concentration. The highest dissociation peaks were observed in VCaP cells and then in patients' sera and the lowest dissociation peaks were observed in woman sera (Fig. 3c, d). The protein TMPRRS2/ERG product was assessed using PLA described above in same prostate cancer patients' sera (Fig. 3e), where LNCaP cell lysate and healthy donors' sera were used as negative controls, while VCaP cell lysate was used as a positive spike control. By serial dilutions of the VCaP and LNCaP cell lysates, regression calibration curves to calculate the concentrations obtained from analysis of the sera were prepared. Data were expressed as number of TMPRSS2/ERG protein molecules in microliter serum. Cluster analysis of protein and transcript expression of fusion gene and ERG isoforms expression across patients with correlation linkage statistics suggested not all patients exhibiting fusion transcript had fusion protein products, while those having wild-type ERG transcripts expressed lacked fusion forms (Fig. 3F, Supporting Information Figs. S5, S6).

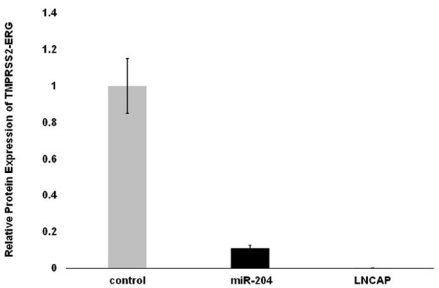
miR-204 Upregulates AR Expression

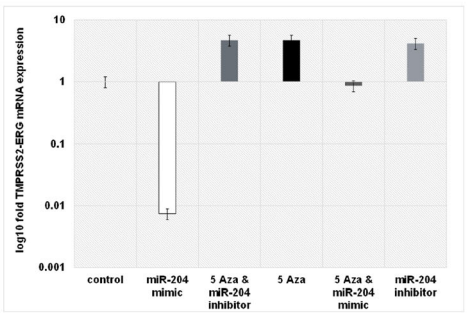
As TMPRSS2 is under androgen control and the rearrangement TMPRSS2/ERG is an important part of prostate cancer advancement, we investigated the role of miR-204 on androgen receptor expression in fusion-free (LNCaP) and fusion-harboring scenario (VCaP). *AR* transcript and protein expression were both increased after miR-204 artificial overexpression using synthetic mimic in both cell lines, while miR-204 suppression, using synthetic complementary inhibition, produced significant decrease in *AR* mRNA and protein levels in LNCaP cells, compared to non-targeting control transfected cells (Fig. 4a, c). Fusion harboring VCaP cells demonstrated dysregulated expression of AR upon miR-204 artificial inhibition (Fig. 4b, c).

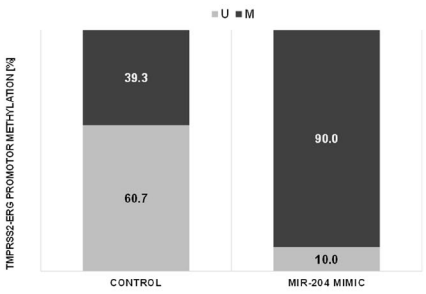


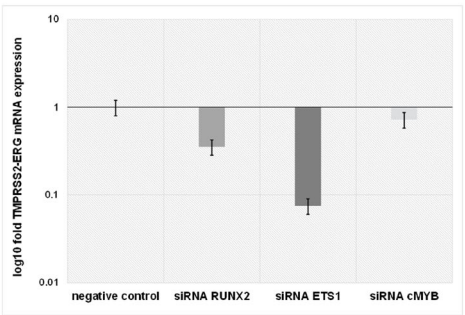


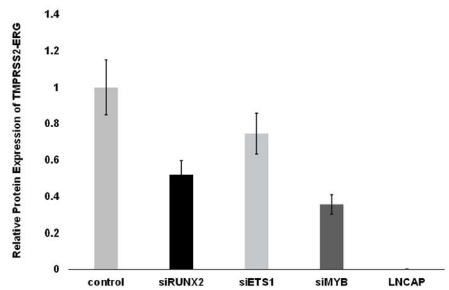

C
log10 fold TMPRSS2-ERG mRNA expression
10
1
0.1
0.01
0.001
negative control
miR-204 mimic
miR-204 inhibitor

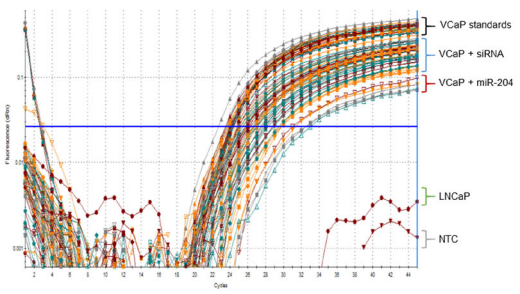

D
Relative Protein Expression of TMPRSS2-ERG
1.4
1.2
1
0.8
0.6
0.4
0.2
0
control
miR-204
LNCAP

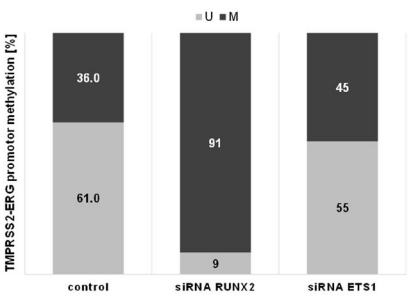

E
log10 fold TMPRSS2-ERG mRNA expression
10
1
0.1
0.01
0.001
control
miR-204 mimic
5 Aza & miR-204 inhibitor
5 Aza
5 Aza & miR-204 mimic
miR-204 inhibitor


F
U
M
TMPRSS2-ERG PROMOTOR METHYLATION [%]
39.3
60.7
90.0
10.0
CONTROL
MIR-204 MIMIC


G
log10 fold TMPRSS2-ERG mRNA expression
10
1
0.1
0.01
negative control
siRNA RUNX2
siRNA ETS1
siRNA cMYB


H
Relative Protein Expression of TMPRSS2-ERG
1.4
1.2
1
0.8
0.6
0.4
0.2
0
control
siRUNX2
siETS1
siMYB
LNCAP


I
VCaP standards
VCaP + siRNA
VCaP + miR-204
LNCaP
NTC


J
U
M
TMPRSS2-ERG promotor methylation [%]
36.0
61.0
91
9
45
55
control
siRNA RUNX2
siRNA ETS1

◀ **Fig. 2** miR-204 is a negative regulator of TMPRSS2/ERG oncofusion, modulating its promoter methylation and regulatory TFs. **a** Amplification plot of *TMPRSS2/ERG* four most frequent fusion variants (primers according Nguyen et al.) in VCaP cells detected using RT-qPCR. Only amplification until Ct 30 is considered specific. **b** Thermal dissociation curves of the same RT-qPCR. **c** VCaP cells were harvested for total RNA. qPCR was performed for *TMPRSS2/ERG* (variant 1). Cycle number was normalized to GAPDH and relative expression (Pfaffl, $2^{\Delta\Delta CT}$) calculated. qPCR data are represented as relative log10 fold change of miR-204 mimic (*white*) or miR-204 inhibitor (*dark gray*) normalized to non-targeting mimic transfected cells (*empty*). **d** TMPRSS2/ERG protein product relative expression (PLA) was assessed in protein cell lysates of miR-204 mimic transfected VCaP cells (*black*) or LNCaP cells (*empty*), normalized to non-targeting mimic transfected VCaP cells (*light gray*). **e** *TMPRSS2/ERG* (variant 1) relative $\log_{10}$ fold expression in VCaP cells after 5-AzaC treatment alone (*black*), miR-204 mimic (*white*), or inhibitor (*light gray*) transfection alone, or 5-AzaC treatment followed by miR-204 mimic (*darker gray*) or inhibitor (*dark gray*) transfection was assessed by qPCR. Non-targeting mimic transfected VCaP cells were used as control (*empty*). **f** Unmethylated (*light*) to methylated (*dark*) ratio (%) of CpG island in TMPRSS2/ERG promoter assessed using MS-PCR of genomic DNA from VCaP cells transfected using miR-204 mimic or non-targeting control. **f** *TMPRSS2/ERG* (variant 1) relative $\log_{10}$ fold expression in VCaP cells after 48 h siRNA *RUNX2* (*gray*), siRNA *ETS1* (*dark gray*), siRNA *cMYB* (*light gray*), or non-targeting control (*empty*) transfection. **h** TMPRSS2/ERG protein product relative expression (PLA) was assessed in protein cell lysates of siRNA RUNX2 (*black*), siRNA ETS1 (*light gray*), siRNA cMYB (*dark gray*), or non-targeting control transfected VCaP cells (*gray*). LNCaP cell lysate was used as negative control (*empty*). **i** TMPRSS2/ERG detection PLA amplification curves of protein cell lysates. **j** Unmethylated to methylated ratio (%) of CpG island in TMPRSS2/ERG promoter assessed using MS-PCR of genomic DNA from VCaP cells transfected using siRNA RUNX2, siRNA ETS1. or non-targeting control

DNA Methyltransferase Inhibition Exhibit Similar to miR-204 Effect on AR Expression

Total genome demethylation using 5-azaCytidine alone resulted in significant upregulation of the *AR* mRNA transcripts and protein in both LNCaP and VCaP cells (Fig. 4a, c), suggesting DNA methylation as regulatory mechanism. Exogenous miR-204 synthetic mimic transfection resulted in even stronger cumulative *AR* mRNA and protein upregulation in LNCaP cells. Similarly to miR-204 inhibitor transfection alone, the 5-AzaC-treated cells transfected with miR-204 inhibitor demonstrated decrease in *AR* mRNA expression in LNCaP cells (Fig. 4a, c).

Conversely, 5-AzaC treatment in fusion harboring VCaP cells resulted in decreased effects of miR-204 inhibition, and reverse effect of miR-204 mimic, compared to fusion-free cancer cells (Fig. 4b, c).

miR-204 Decreases the Level of DNA Methylation of AR Promoter

We further investigated whether the miR-204 effect was a result of AR promoter methylation (Supporting Information Fig. S7). Overexpression of miR-204 produced a decrease in DNA methylation of the CpG region near exon1 of AR that was small in LNCaP cells (4%) (Fig. 4d) and stronger in VCaP cells (20%) (Fig. 4e). This AR promoter demethylation corroborated with the expression data observed after miR-204 overexpression in these cell lines.

Master Transcriptional Factors Are Positive Regulators of AR

Since we found the transcription factors MYB, RUNX2, and ETS1 to be involved in TMPRSS2/ERG expression and promoter methylation, we investigated how this was coordinated with AR, known to be a major TMPRSS2 inductor. Silencing *MYB*, *RUNX2*, and *ETS1*, one at a time, we found that they were also required for AR transcript expression in both LNCaP and VCaP cell lines. Notably, in fusion harboring VCaP cells, the effect of *RUNX2* and c-*MYB* to sustain AR expression was reduced (Fig. 4f, g).

Master Transcriptional Factors Maintain Lower Levels of DNA Methylation of AR Promoter

Looking for eventual mechanism of RUNX2 and ETS1 action on AR expression, we found that their siRNA mediate silencing resulted in significant AR promoter hyper-methylation, suggesting that they regulated AR levels by keeping its promoter less methylated. Silencing *RUNX2* resulted in an AR promoter methylation increase from 20 to 57% in LNCaP and 30 to 47% in VCaP cells. Similarly, silencing *ETS1* resulted in an AR promoter methylation increase from 20 to 81% in LNCaP cells and 30 to 67% in VCaP cells (Fig. 4h, i).

High-Resolution Proteomics Revealed miR-204 Overexpression to Switch Differentially Several Sets of Proteins Involved in Epigenetic Regulation and Androgen Receptor Signaling in LNCaP Cells

Since we found that miR-204 was able to modulate the promoter methylation of an androgen receptor-controlled gene as TMPRSS2; hence, the fusion gene TMPRSS2/ERG as well, and also the androgen receptor promoter itself, we decided to investigate further the mechanics of miR-204 actions using high-resolution proteomics.

Androgen receptor signaling enabled prostate cancer cell line LNCaP was subject to miR-204 mimic transfection for 24 h and subsequent LTQ Orbitrap proteomics. Non-sense, non-coding RNAs were transfected to LNCaP cells used as negative controls. We found that miR-204 induced "reciprocal switch" of groups of co-expressed proteins (Supporting Information Fig. S8, Fig. 5a). The proteins were pre-selected to be significantly differentially expressed and subject to hierarchical clustering.

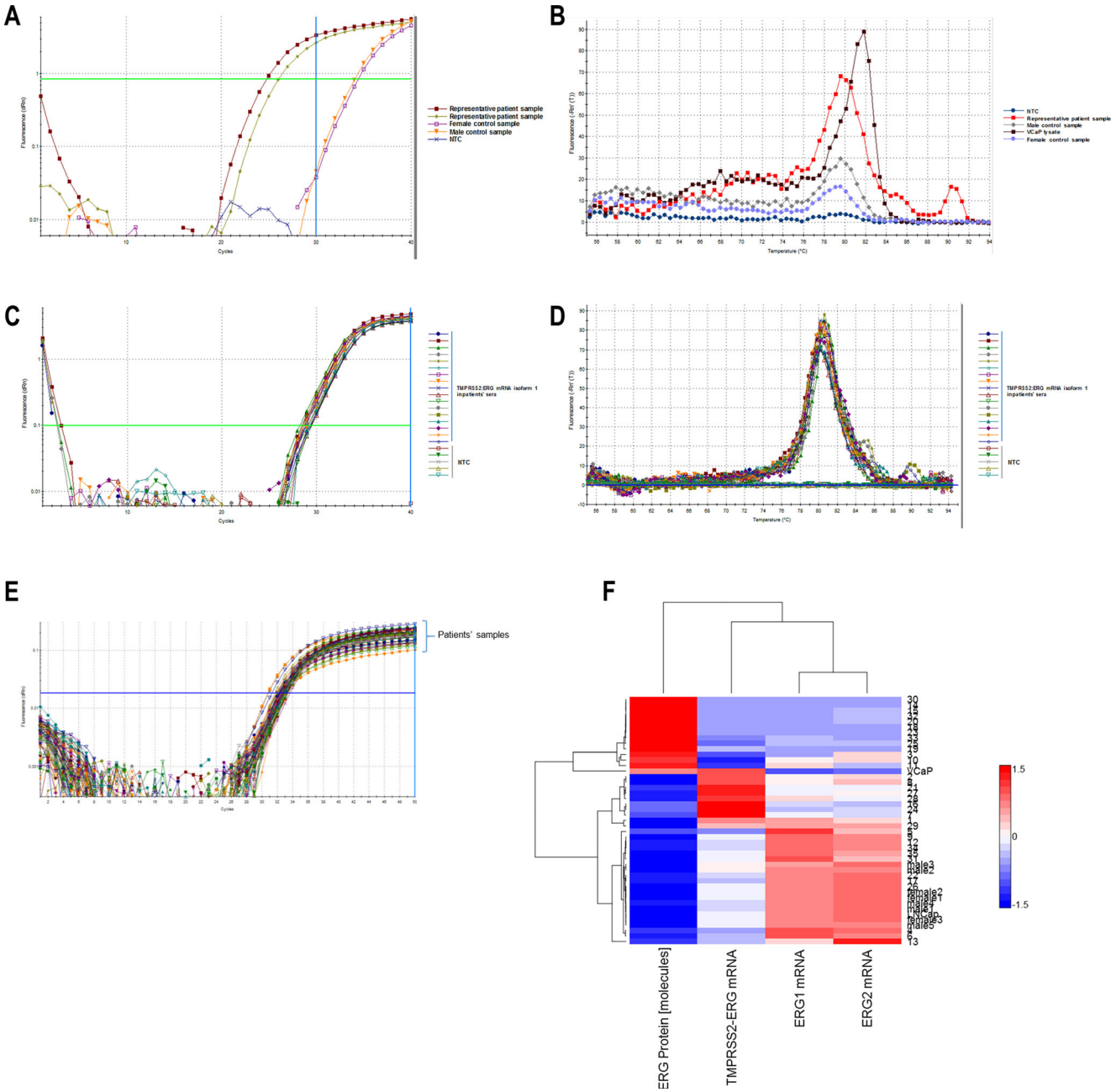


Fig. 3 TMPRSS2/ERG transcript and protein expression in prostate cancer patients' sera. **a** Amplification plot of TMPRSS2/ERG four most frequent fusion variants (primers according Nguyen et al.) in prostate cancer patients' and healthy male and female donors' sera, detected using RT-qPCR. Only amplification until Ct 30 is considered specific. **b** Thermal dissociation curves of the same RT-qPCR. **c** Amplification plot of TMPRSS2/ERG fusion I (TMPRSS2-exon 1:ERG exon 4) of different patients' sera. Non-template controls are also included. **d** Thermal dissociation curves of the same patients' sera RT-qPCR. **e** TMPRSS2/ERG detection PLA amplification curves of protein cell lysates of patients' sera. **f** Patients' samples hierarchical clustering using median linkage metrics (*rows*) vs. hierarchical clustering of TMPRSS2/ERG protein abundance (expressed as molecules), *TMPRSS2/ERG* fusion I transcript abundance (relative expression, pre-normalized to VCaP cell line basal one), wt *ERG* isoform 1 transcript abundance and wt *ERG* isoform 2 transcript abundance, all in sera. All data were normalized using median absolute deviation before clustering

We further analyzed for over-representation of proteins (iPathwayGuide) based on their place in signaling pathways and also based on gene set enrichment analysis and annotation enrichment of pathways (EGAN) (Supporting Information Figs. S9, S10).

We found using iPathwayGuide, 173 proteins to be significantly expressed (Supporting Information Table 1), 81 of which generally related to cancer. Among these genes, 54 were related to prostate cancer, 12 to androgen receptor, and 3 genes were related to DNA methylation or hypo-methylation.

Some of the most significantly overexpressed by miR-204 proteins encompassed DNA demethylation—methyl-CpG-binding domain protein 3 (MBD3), metastasis capability of androgen-independent prostate cancer cells-plectin (PLEC), and AR activity regulation related glycogen synthase kinase 3 beta (GSK3B). Some of the most significantly downregulated by miR-204 proteins involved in differential AR transactivation and co-activation were peroxiredoxin 2 (PRDX2), DNA-PKcs (PRKDC), and AR-dependent marker—FOHL1 (PSMA) (Fig. 5b).

Using Exploratory Gene Association Network (EGAN), we constructed a protein over-represented hypergraph that connected differentially overexpressed proteins including protein-protein interaction, common pathways co-expression to pathway annotation enrichment and "gene set enrichment analysis" (GSEA) (EGAN supports Broad at MIT developed GSEA metrics). Linkage to GSEA (probability score), gene ontology terms, PANTHER (Supporting Information Table 2) and KEGG pathway, specific gene sets that were defined in Molecular Signatures Database (MSigDB) on prostate cancer, miR-204 known targets and androgen receptor were incorporated in annotation over-representation analysis (Supporting Information Table 3), with part of the terms shown on the hypergraph (Fig. 5c).

Interestingly, only 1 out of 14 detected proteins annotated as "*hsa-miR-204 targets* having a 3′-UTR seed region UCCCUUU" (TargetScan predicted in MSigDB), was actually downregulated as expected—GANAB, while all other 13 proteins (DDX3Y, UGDH, CNOT1, RAB1A, SF3B1, NAA15, ERP44, UBA6, AP2A2, CMBL, NBEA, MPI, ZADH2) were upregulated, suggesting loss of miR-204 restriction control.

We found 3 out of 57 proteins annotated in "genes upregulated in PCa after 5-AzaC demethylation compound treatment" to be miR-204 regulated. We found also enrichment on miR-204 overexpression in GO processes terms "DNA methylation," "DNA methylation on cytosine," and "C-5 methylation of cytosine" (5-MeC enriched), represented by downregulation of DNA (cytosine-5)-methyltransferase 1 (DNMT1).

Analyzing GO terms, we found mapping with the term "methylation" and several histone regulation enzymes—*histone H3-R3 methylation*, *histone methyltransferase* (*H3-R3*), *histone arginine methylation*, and *tRNA methylation*, but no mapping to the closely related *histone methyltransferase* (*H3-K36*), *histone methyltransferase* (*H3-K4*), *histone methyltransferase* and (*H4-K20*). Thus, the additionally dysregulated by miR-204 epigenetic related proteins were mapped to methylation (TPMT, TRMT61A, RAB3B, MAT2A, AHCY, GAMT); histone H3-R3 methylation (PRMT5); and tRNA methylation (NSUN2, TRMT61A) (Supporting Information Fig. S11).

AR Signaling-Related Gene Sets Are positively Represented in LNCaP Cells in miR-204 Overexpressed Context

As some of the differentially regulated genes were tight to the AR signaling axis and some were upregulated by miR-204 while being DNA hyper-methylation downregulated in PCa, we further investigated how miR-204 affects AR regulation and downstream targets. We used EGAN-based annotation of significantly miR-204 regulated proteins to available MSigDB gene sets (MSigDB C3: transcription factor targets).

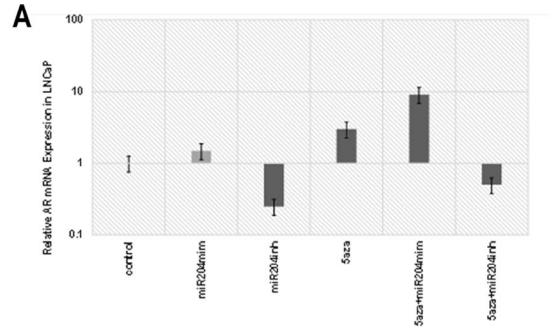
We found differentially regulated by miR-204 six proteins in "MSigDB C3: Regulation of Androgen receptor activity," six proteins in "MSigDB C3: Genes Up-regulated in PCa in response to androgen," four proteins in "MSigDB C3: AR- > Downregulated in PCa in response to Androgen," one of six proteins in "MSigDB C3: Androgen receptor activity Co-regulation Molecules" (Supporting Information Table 4). We further followed three different transcription start sites (TSS) that were targeted by the AR for its transcriptional downstream regulation and we found only three miR-204 affected proteins had common TSS—KIAA1468, NCDN, and SF3B1 ("MSigDB C3: transcription factor targets: AR- > promoter [−2 kb, 2 kb], TSS: GGTACANNRTGTTC T; TSS: NNNNNNRGNACRNNGTGTTCTNNNNNN; TSS: NNNGNRRGNACANNGTGTTCTNNNNNN") (Fig. 6).

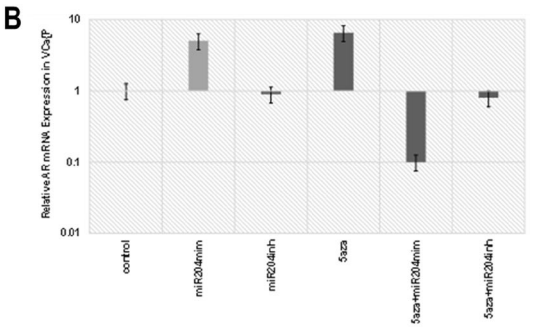
In general, miR-204 was established to switch several sets of genes thus affecting proteins involved in epigenetic regulation, androgen receptor signaling, cancer fusion gene products and tumor pro-survival signaling in LNCaP cells.

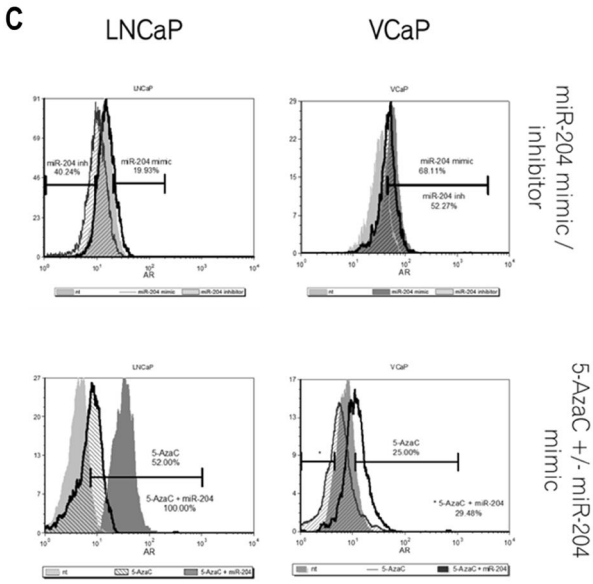
Discussion

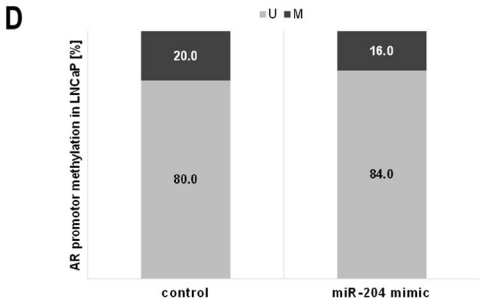
Hallmark of cancerogenesis is the loss of epithelial differentiation in favor of stem-like metastatic phenotype [37]. Genomic instability and resulting alterations are under constant survival selective pressure, preserving only those conferring perturbed cancer regulatory gene networks a steady state until the occurrence of next driver mutation. Since tissue and cell type differentiation programs are tightly regulated by epigenetic control mechanisms cancer de-differentiation phenomenon engages epigenetic reprogramming as well. Paramount for PCa is the loss of AR-dependent differentiation and the acquirement of androgen insensitivity.

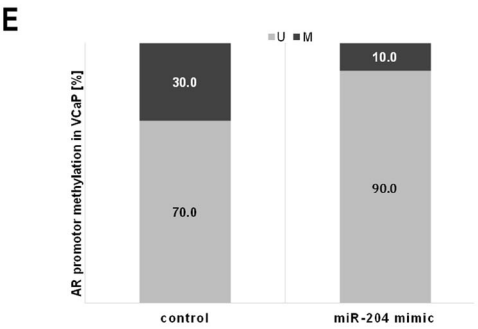
In this study, we found that the non-coding micro-RNA miR-204, dysregulated in PCa, exert regulatory action on AR and most common prostate cancer fusion product containing AR-dependent promotor TMPRSS2/ERG. miR-204 also reprogrammed genes under AR control, in some cases acting as tumor suppressor, but in other as oncomiR. Our main focus was on the regulatory axis AR-TMPRSS2/ERG. By

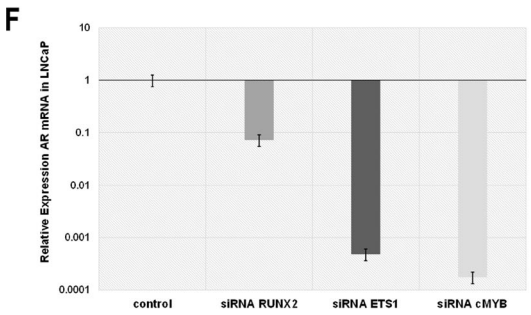

A
Relative AR mRNA Expression in LNCaP
100
10
1
0.1
control
miR204mim
miR204inh
5aza
5aza+miR204mim
5aza+miR204inh

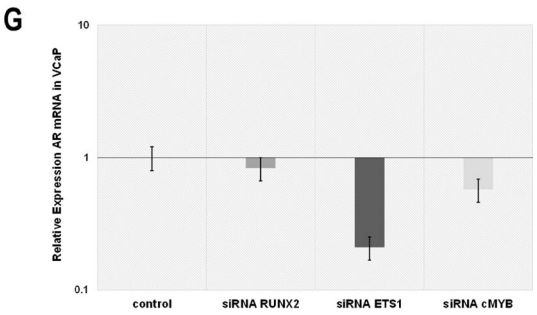

B
Relative AR mRNA Expression in VCaP
10
1
0.1
0.01
control
miR204mim
miR204inh
5aza
5aza+miR204mim
5aza+miR204inh

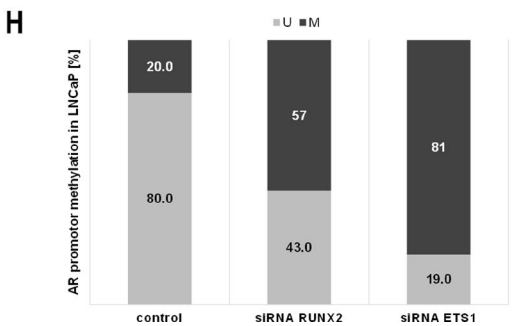

C
LNCaP
VCaP
miR-204 mimic / inhibitor
miR-204 inh 40.24%
miR-204 mimic 19.93%
miR-204 mimic 68.11%
miR-204 inh 52.27%
5-AzaC +/- miR-204 mimic
5-AzaC 52.00%
5-AzaC + miR-204 100.00%
5-AzaC 25.00%
5-AzaC + miR-204 29.48%
AR

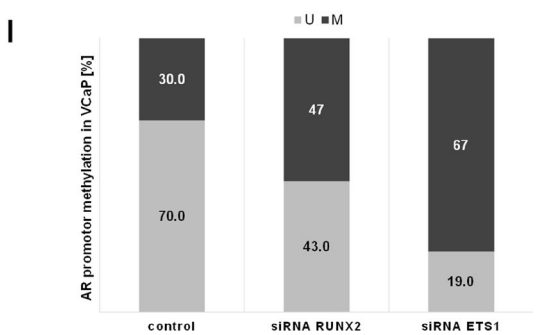

D
U M
AR promotor methylation in LNCaP [%]
20.0
80.0
16.0
84.0
control
miR-204 mimic


E
U M
AR promotor methylation in VCaP [%]
30.0
70.0
10.0
90.0
control
miR-204 mimic


F
Relative Expression AR mRNA in LNCaP
10
1
0.1
0.01
0.001
0.0001
control
siRNA RUNX2
siRNA ETS1
siRNA cMYB


G
Relative Expression AR mRNA in VCaP
10
1
0.1
control
siRNA RUNX2
siRNA ETS1
siRNA cMYB

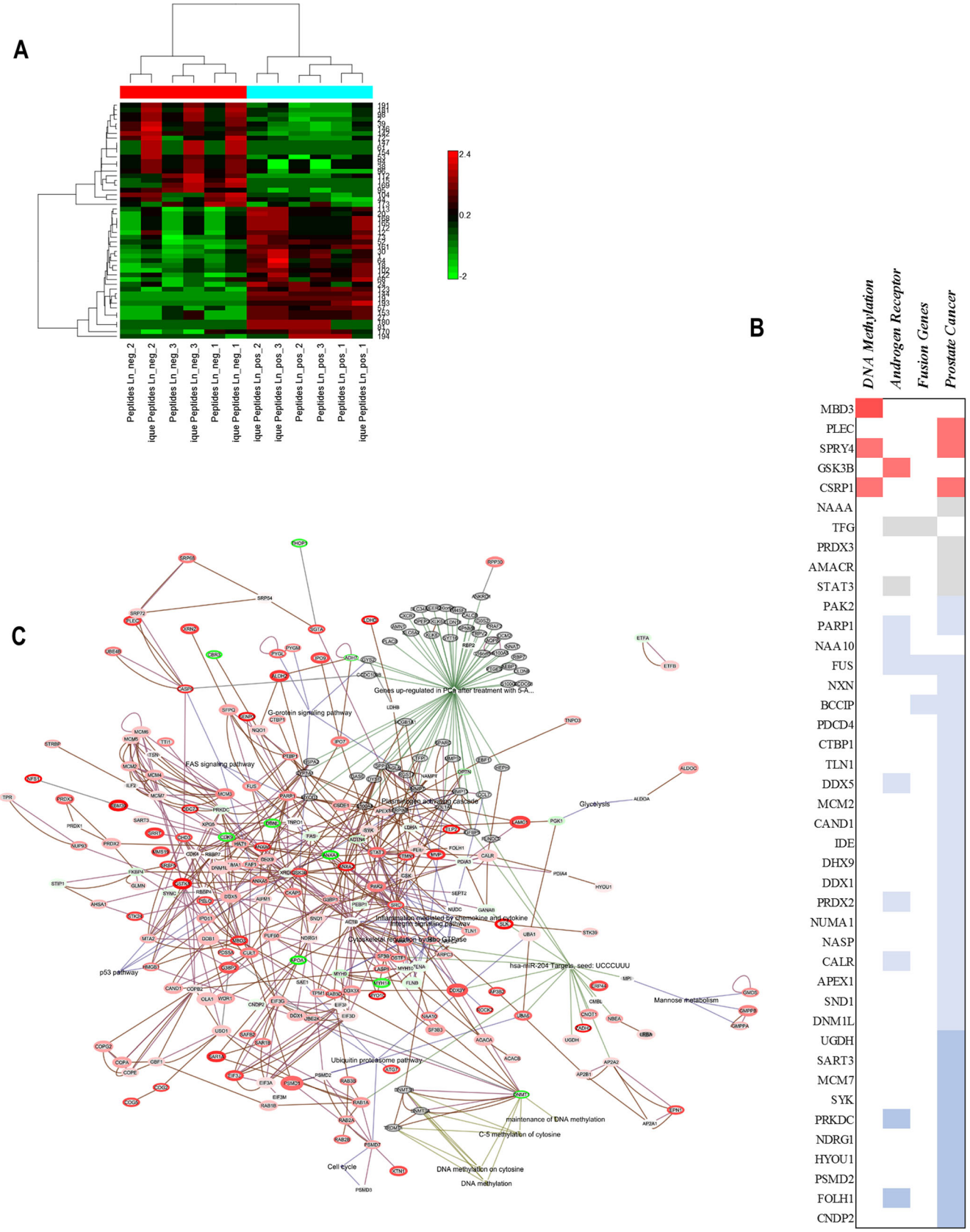

H
U M
AR promotor methylation in LNCaP [%]
20.0
80.0
57
43.0
81
19.0
control
siRNA RUNX2
siRNA ETS1


I
U M
AR promotor methylation in VCaP [%]
30.0
70.0
47
43.0
67
19.0
control
siRNA RUNX2
siRNA ETS1

◀ **Fig. 4** miR-204 upregulates AR expression, decreasing the level of DNA methylation of AR promoter, along with miR-204 modulated TFs. **a**, **b** *AR* relative $\log_{10}$ fold expression in LNCaP (**a**) and VCaP (**b**) cells after 5-AzaC treatment alone, miR-204 mimic or inhibitor transfection alone, or 5-AzaC treatment followed by miR-204 mimic or miR-204 inhibitor transfection was assessed by qPCR. Non-targeting mimic transfected LNCaP cells were used as control. **c** AR protein abundance in LNCaP and VCaP cells was evaluated using FCS. FCS overlay histograms of non-targeting control (*light gray shaded*), miR-204 mimic (*black line*) or miR-204 inhibitor (*light pattern*) transfected LNCaP cells are presented. FCS overlay histograms of non-targeting control (*light gray shaded*), miR-204 mimic (*black pattern*) or miR-204 inhibitor (*dark gray shaded*) transfected VCaP cells are presented. FCS overlay histograms of non-targeting control (*light gray shaded*), 5-AzaC treated (*light pattern*) or 5-AzaC treated and miR-204 transfected cells (*dark gray shaded*) transfected LNCaP cells are presented. FCS overlay histograms of non-targeting control (dark gray shaded), 5-AzaC treated (*black line*) or 5-AzaC treated (*black pattern*) transfected VCaP cells are presented. Histogram area difference is presented as percentage. **d**, **e** Unmethylated (light) to methylated (*dark*) ratio (%) of CpG island in *AR* promoter assessed using MS-PCR of genomic DNA from LNCaP (**d**) or VCAP (**e**) cells transfected using miR-204 mimic or non-targeting control. **f**, **g** *AR* relative $\log_{10}$ fold expression in LNCAP (F) and VCaP (G) cells after siRNA *RUNX2* (gray), siRNA *ETS1* (dark gray), siRNA *cMYB* (*light gray*) or non-targeting control (*empty*) transfection. **h**, **i** Unmethylated to methylated ratio (%) of CpG island in *AR* promoter assessed using MS-PCR of genomic DNA from LNCaP (**h**) or VCaP (**i**) cells transfected using siRNA RUNX2, siRNA ETS1, or non-targeting control

developing a method to detect the in-frame protein products of this most frequent fusion, we found that miR-204 suppress TMPRSS2/ERG mRNA and protein in both ERG fusion-negative (LNCaP) and ERG fusion-positive (VCaP) cells through altering its promotor DNA methylation. Recent studies have explained why AR positive PCa harbor TMPRSS2/ERG fusion [9, 38]. The ERG-containing fusion is stimulated by AR signaling, but at the same time, it directly disrupts AR differentiation program by repressing AR, and binding AR in its AREs, preventing activation of androgen-dependent differentiation target genes. The ERG fusion also stimulates its alternative wt ERG allele. As a consequence, ERG initiates repressive epigenetic program of de-differentiation and acquiring stem-like properties by activating PcG H3K27 methyltransferase EZH2 [9]. EZH2 further recruits DNMT1 to its repressive targets by DNA methylation [39]. DNMT1 has a role in the establishment and regulation of tissue-specific patterns of methylated cytosine residues [40, 41]. This effect is potentiated in ERG-positive cells by direct recruiting of arginine methyltransferase PRMT5 by ERG to AREs and subsequent AR methylation at 761, preventing differentiation and rendering AR transcriptionally hyperactive [42]. Overall, TMPRSS2/ERG drives PCa cells towards stem-like mesenchymal phenotype stimulated by the AR which it also represses. We found here that besides suppressing TMPRSS2/ERG expression, miR-204 controls not only its promotor DNA methylation but also suppresses the expression of PRMT5 and DNMT1. The suppression of PRMT2 and DNMT1 would suggest that miR-204 still acts as tumor suppressor miR in this context. On the other hand, miR-204 induced repression of TMPRSS2/ERG would suggest that miR-204 restricts TMPRSS2/ERG from overexpression, as otherwise ERG would suppress AR in ERG fusion-positive cells tremendously [9]. This could probably suppress the feedback loop between the AR and TMPRSS2/ERG. The expression of TMPRSS2/ERG requires some basal stimulation of its AR-dependent promotor by the AR. In fact, ERG does not suppress AR in fusion-negative cells [43], as AR hyper-stimulation is required for chromosomal looping to bring together TMPRSS2 and ERG and produce fusion [44].

As TMPRSS2/ERG fusion suppression is achieved by DNA hyper-methylation, this would imply that intermediate interactors are acting to control the binding of the DNMTs to either fusion promotor or to EZH2. At the same time, miR-204 upregulated AR through promotor/exon 1 DNA hypo-methylation, especially in ERG fusion-positive context. In this regard, here we found that miR-204 upregulated methyl-CpG binding domain protein 3 (MBD3). This protein recognizes methylated DNA [45] and it is required in establishing or maintaining transcriptionally repressed chromatin [40] by creating MBD2-MDB3 complex (with methyl-CpG binding domain protein 2 (MBD2)), recruiting further chromatin remodeling and histone deacetylase complexes like histone deacetylase 1 (HDAC1) and metastasis associated 1 family member 2 (MTA2), as well as demethylase DNMT1 [46]. Although associated with transcriptional repression, in PCa HDACs are also paradoxically required for activation of a substantial fraction of AR target genes including *TMPRSS2*—the ERG fusion transcriptional driver itself [47]. The histone deacetylase MBD3/MBD2/HDAC1 complex is thus promoted by miR-204 overexpression, supporting further the hypothesis that miR-204 potentiates TMPRSS2/ERG action in PCa, although TMPRSS2/ERG is downregulated simultaneously by promoter hyper-methylation. This notion is supported by the finding that ERG-positive prostate cancers are strongly HDAC1-positive [43]. The dualistic activity of miR-204 is further elucidated by analyzing its positive regulation of AR. Here, we found that miR-204 acts on AR expression in coordinated manner, by AR promotor hypo-methylation, significantly downregulating DNMT1 and upregulating SWI/SNF complex subunit (SMARCC1). Although miR-204 should exert tumor suppressor and epithelial differentiation properties by inducing AR, what actually happens is that during PCa advancement to castration-resistant mesenchymal phenotype, AR signaling is required for acquiring ERG- and other TFs fusion mutations that have AR promotors and are AR controlled. In PCa, AR binding is altered and AR signaling in reprogrammed by various TFs like ETS1, ERG, RUNX2, etc. and the role of modulatory miRs as miR-204 resides in their own "reprogramming" by exploiting their AR promoting activities. In this regard, miR-204 counters


A
2.4
0.2
-2
Peptides Ln_neg_2
Peptides Ln_neg_3
Peptides Ln_neg_1
Peptides Ln_pos_2
Peptides Ln_pos_3
Peptides Ln_pos_1
B
DNA Methylation
Androgen Receptor
Fusion Genes
Prostate Cancer
MBD3
PLEC
SPRY4
GSK3B
CSRP1
NAAA
TFG
PRDX3
AMACR
STAT3
PAK2
PARP1
NAA10
FUS
NXN
BCCIP
PDCD4
CTBP1
TLN1
DDX5
MCM2
CAND1
IDE
DHX9
DDX1
PRDX2
NUMA1
NASP
CALR
APEX1
SND1
DNM1L
UGDH
SART3
MCM7
SYK
PRKDC
NDRG1
HYOU1
PSMD2
FOLH1
CNDP2
C
Genes up-regulated in PCA after treatment with 5-A...
G-protein signaling pathway
FAS signaling pathway
Plasminogen activation cascade
Glycolysis
Inflammation mediated by chemokine and cytokine
Integrin signaling pathway
Cytoskeletal regulation by Rho GTPase
p53 pathway
hsa-miR-204 Targets, seed: UCCCUUU
Mannose metabolism
Ubiquitin proteasome pathway
maintenance of DNA methylation
C-5 methylation of cytosine
Cell cycle
DNA methylation on cytosine
DNA methylation

◀ **Fig. 5** High-resolution proteomics reveals miR-204 overexpression differentially switched several sets of proteins involved in epigenetic regulation and androgen receptor signaling in LNCaP cells. **a** Hierarchical clustering *(row)* of proteins differentially expressed in LNCaP after transfection with miR-204 mimic or non-targeting control. Hierarchical clustering (*columns*) of peptide spectral counts data among treatment replicates. **b** miR-204 overexpression modulated protein differential expression as analyzed using iPathwayGuide, combining gene expression information with pathway gene topological position. Proteins inclusion cutoffs were $p < 0.05$, minimum fold change ≥0.5. Proteins were classified as DNA methylation-related, AR-related, fusion genes (participate in gene fusions), and prostate cancer-related. *Color* represents Fold change: *dark red* between 1.5 and 2; *dark red* between 1.16 and 1.5; *gray* between 0.8 and 1.16; *blue* between 0.8 and 0.4; and *dark blue* below 0.45. **c** Over-represented hypergraph constructed by EGAN, connecting differentially overexpressed proteins based on protein-protein interaction, common pathways co-expression, annotation enrichment and "gene set enrichment analysis" (GSEA). Graph was annotated using GSEA (probability score), GO terms, PANTHER and KEGG pathway databases, and MSigDB prostate cancer-specific gene sets on miR-204 known targets and AR

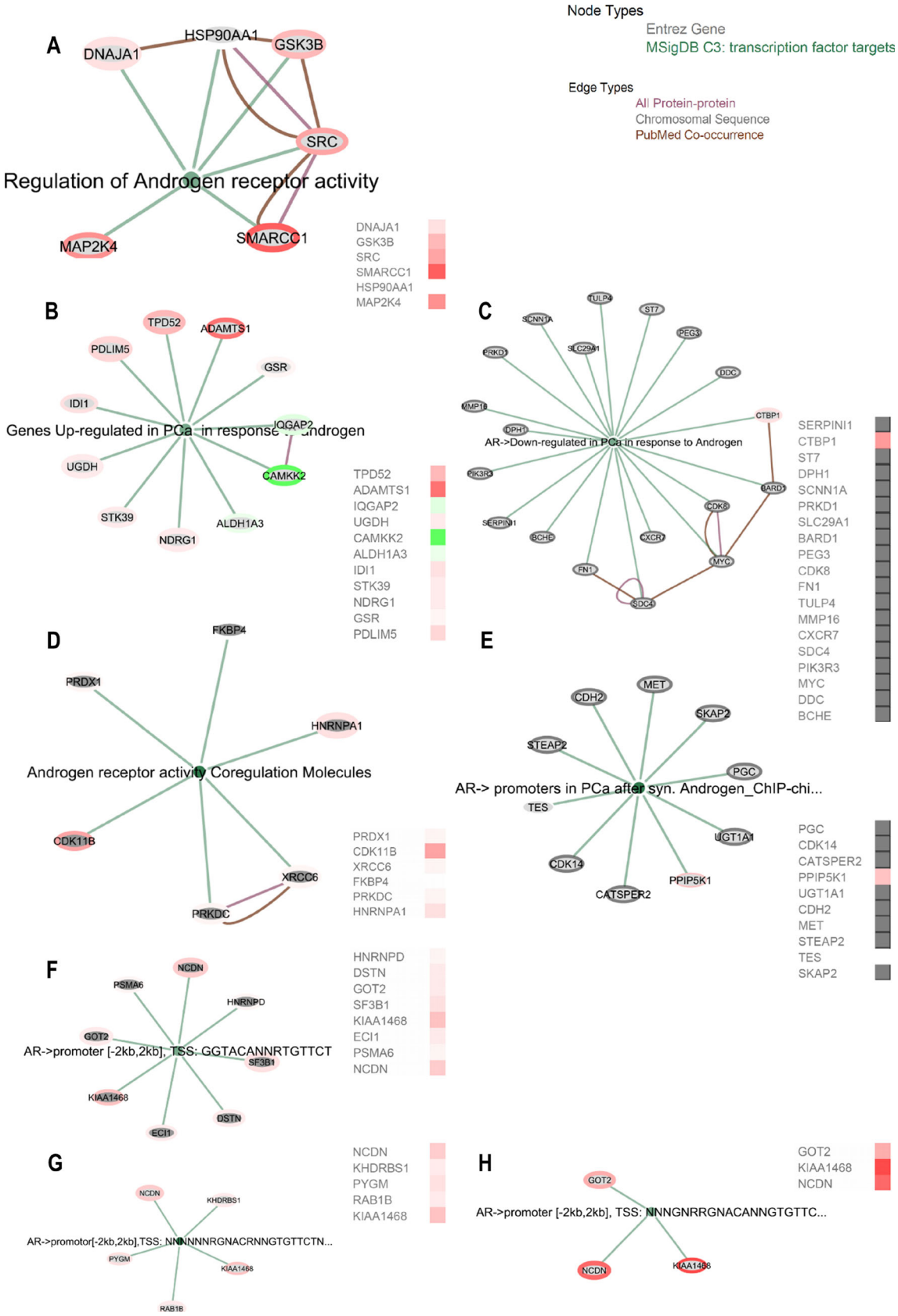
DNMT1 repressive activity towards AR to promote ERG fusion formation. DNMT1 is PCa overexpressed as E2F1/DNMT1 axis is activated during the emergency of castration-resistant PCa in vitro and in clinical samples [48], repressing AR via both DNA methylation-independent and -dependent fashion [48, 49]. Hence, miR-204 acts by cooperating DNMT1 downregulation with upregulation of chromatin remodeling SMARCC1 to upregulate the expression of AR. SMARCC1 is upregulated in PCa and is part of the intranuclear complex that enhances the AR transactivation altering the chromatin structure [50].

We found further clues that miR-204 has balancing role helping the transition from AR-responsive to castration-resistant phenotype maintaining by AR signaling regulation and interaction with AR binding and differentiation disrupting TFs. In TMPRSS2/ERG fusion-negative PCa cell lines, miR-204 attenuates cell proliferation and the expression of TFs ETS1, RUNX2, and cMYB, a phenomenon that is lost in TMPRSS2/ERG fusion harboring cell lines. *ETS1* transcription is increased in advanced prostate cancer, promoting castration-resistant phenotype and cancer cell motility induction [51], by directly binding AR in its AREs and preventing its program. Similarly, RUNX2 also directly interacts with AR, reprograming it. AR-RUNX2 protein complex binds the *SNAI2* (SLUG) enhancer region, resulting in an increased tumor invasiveness [52], while many genes that are androgen-dependent are silenced. RUNX2 is the master TF required for bone marrow metastasizing in PCa as it drives PCa cells to bone marrow stem cell niche [21]. The TF c-MYB known to promote cell proliferation and tumorigenesis is also found in PCa bone metastasis. C-MYB engages DNA-damage response and promotes DNA repair in castration-resistant PCa. It is transcriptionally activated by androgen deprivation or impairment of androgen receptor signaling and c-MYB signaling could even replace signaling previously exerted by the AR [20]. RUNX2 and ETS1 are also stimulated by c-MYB [14].

We found that in ERG fusion-negative cells, ETS1, RUNX2, and c-MYB are required for AR expression, while in ERG-negative cells, only ETS1 kept significantly this ability, probably due to its potentiating role on miR-204. The dualistic transition effect miR-204 is also evident here, as in TMPRSS2/ERG-negative cells miR-204 downregulated these TFs, while they both stimulate miR-204 expression, and AR expression, but in TMPRSS2/ERG-positive cells, miR-204 upregulates these TFs, while they lower their ability to upregulate AR.

Using LTQ Orbitrap proteomics, we found that even though miR-204 demonstrated some tumor suppressor abilities by downregulating genes overexpressed in PCa, while it additionally promoted oncogenes already upregulated with the advance of PCa. The dualistic activity of many such molecules is probably partially related to the dualistic role the AR signaling itself plays in the PCa progression [53]. Since AR plays a role in non-coding RNA maturation [54], there is possibility an underlying non-coding RNA duality associated with AR to be linked to AR co-regulatory axis duality. Thus, miR-204 played besides tumor suppressor, oncomiR actions as well, downregulating PCa tumor suppressors, and also by upregulating multiple cancer-related genes.

Additionally, to design the use in this study method for TMPRSS2/ERG protein products detection, we analyzed TMPRSS2/ERG sequence variants and their structural models and the probability of epitope preservation in case more than exon-1 of TMPRSS2 is present. Previously, a common TMPRSS2/ERG fusion variant has been shown to be detectible in the urine of men with prostate cancer and has been coupled with other molecular markers in urine-based cancer detection [55, 56]. We investigated TMPRSS2/ERG expression in advanced (stages III–IV) prostate cancer patients' sera as we expected that its availability would coincide with sufficient circulating tumor cells. TMPRSS2/ERG mRNA correlated positively in part of cases to wt ERGs in concordance to previously established positive feed forward loop between the fusion and native ERG [9, 43], the cases expression detectable protein form had only few such positive correlations. This could probably be due to the increased likelihood of detection of longer TMPRSS2 exon-containing forms, that could result from harsher rearrangements and hence loss of fusion feed forward loop. Another aspect of this lack of full correlation is probably related to the ambiguity in TMPRSS2/ERG observations in many pathology studies with either ERG expression [10] or Gleason score [57, 58]. Although lots of fusions are reported to be between exon-1 of TMPRSS2 and ERG, there are also in-frame functional TMPRSS2/ERG that contain up to exon-3 or even up to exon-5 of TMPRSS2 [6]. To our understanding, the more complex fusion products


Node Types
Entrez Gene
MSigDB C3: transcription factor targets
Edge Types
All Protein-protein
Chromosomal Sequence
PubMed Co-occurrence
A
DNAJA1
HSP90AA1
GSK3B
SRC
Regulation of Androgen receptor activity
MAP2K4
SMARCC1
DNAJA1
GSK3B
SRC
SMARCC1
HSP90AA1
MAP2K4
B
TPD52
ADAMTS1
PDLIM5
GSR
IDI1
IQGAP2
Genes Up-regulated in PCa in response to Androgen
UGDH
CAMKK2
STK39
ALDH1A3
NDRG1
TPD52
ADAMTS1
IQGAP2
UGDH
CAMKK2
ALDH1A3
IDI1
STK39
NDRG1
GSR
PDLIM5
C
TULP4
ST7
SCNN1A
PEG3
SLC29A1
PRKD1
DDC
MMP16
CTBP1
DPH1
AR->Down-regulated in PCa in response to Androgen
PIK3R3
BARD1
CDK8
SERPINI1
BCHE
CXCR7
MYC
FN1
SDC4
SERPINI1
CTBP1
ST7
DPH1
SCNN1A
PRKD1
SLC29A1
BARD1
PEG3
CDK8
FN1
TULP4
MMP16
CXCR7
SDC4
PIK3R3
MYC
DDC
BCHE
D
FKBP4
PRDX1
HNRNPA1
Androgen receptor activity Coregulation Molecules
CDK11B
XRCC6
PRKDC
PRDX1
CDK11B
XRCC6
FKBP4
PRKDC
HNRNPA1
E
MET
CDH2
SKAP2
STEAP2
PGC
AR-> promoters in PCa after syn. Androgen_ChIP-chi...
TES
UGT1A1
CDK14
PPIP5K1
CATSPER2
PGC
CDK14
CATSPER2
PPIP5K1
UGT1A1
CDH2
MET
STEAP2
TES
SKAP2
F
NCDN
PSMA6
HNRNPD
GOT2
AR->promoter [-2kb,2kb], TSS: GGTACANNRTGTTCT
SF3B1
KIAA1468
DSTN
ECI1
HNRNPD
DSTN
GOT2
SF3B1
KIAA1468
ECI1
PSMA6
NCDN
G
NCDN
KHDRBS1
AR->promotor[-2kb,2kb],TSS: NNNNNNRGNACRNNGTGTTCTN...
PYGM
KIAA1468
RAB1B
NCDN
KHDRBS1
PYGM
RAB1B
KIAA1468
H
GOT2
AR->promoter [-2kb,2kb], TSS: NNNGNRRGNACANNGTGTTC...
NCDN
KIAA1468
GOT2
KIAA1468
NCDN

◀ **Fig. 6** AR signaling related gene sets are positively represented in LNCaP cells after miR-204 overexpression. **a**–**d** EGAN-based annotation of significantly enriched proteins after miR-204 mimic transfection in MSigDB C3 gene sets: **a** "Regulation of Androgen receptor activity," **b** "Genes Up-regulated in PCa in response to androgen," **c** "AR->Downregulated in PCa in response to Androgen," **d** "Androgen receptor activity Co-regulation Molecules," **e** "AR-> promoters in PCa after synthetic Androgen – Chip-chip," **f**–**h** miR-204 modulated proteins enrichment in transcription start site (TSS) that were targeted by the AR for its transcriptional downstream regulation: "MSigDB C3: transcription factor targets: AR->promoter [-2kb, 2kb], TSS: GGTACANNRTGTTCT; TSS: NNNNNNRGNACRNNGTGTTCTNNNNNN; TSS: NNNGNRRGNACANNGTGTTCTNNNNNN"

containing peptide sequences from TMPRSS2 encoded in-frame, or at least preserving in-frame ERG, are the only functional fusion products that have to be considered, as no proper protein folding and transcriptional activity will be available otherwise. On the other hand, there are fusions that even lack ETS-DNA binding domain, but there is no clue if they are functional.

In summary, hsa-miR-204 provides a model of a balancer of the AR function and androgen controlled TMPRSS2/ERG fusion gene, during the phase of preserved AR sensitivity. Epithelial lineage-specific miR-204 promotes AR signaling required for ERG-fusion translocation, restricting differentiation perturbing ETS1, RUNX2, and MYB until the TMPRSS2/ERG fusion gene is acquired. TMPRSS2/ERG plays a critical role in cancer progression by disrupting lineage-specific differentiation of the prostate and potentiating the PRMT5 and EZH2-mediated de-differentiation program. At this point, miR-204 tumor suppressor activity is reprogrammed to oncomiR activity by changing the chromatin organization, epigenetic reprogramming of cell differentiation and histone and DNA methylation regulation, upregulation of other differentiation disrupting AR-interacting TFs like RUNX2, ETS1, and c-MYB. miR-204 also prevents TMPRSS2/ERG fusion from overexpression thus maintaining its regulatory loop with AR. At the point of castration resistance, miR-204 promotes c-MYB expression, the latter replacing lacking AR signaling and protecting PCa cells. This study thus provide means of understanding cancer ambiguity and identify new targets for preventing transition to cancer resistance phenotype.

Acknowledgements This study was funded by grants from The National Sciences Fund at Bulgarian Ministry of Education and Sciences (DMU 03/27 and DCOST 01/23) provided to Krassimira Todorova, Ph.D., D.Sc. and Soren Hayrabedyan, M.D., Ph.D., D.Sc.

The experimental work described herein was performed in laboratories of Institute of Biology and Immunology of Reproduction and Proteomics Core of Essex University, School of Biological Sciences. Method of detection of TMPRSS2/ERG protein products has patent pending procedure at Bulgarian National Patent Office (2014, K.T., S.H.).

Authors' Contributions K.T and S.H. designed the study, TMPRSS2/ERG protein detection method. K.T. did the cell-based experiments, FCS, gene silencing. and molecular biology studies. M.V.M. and G.M. did the LTQ Orbitrap proteomics and proteomics data analysis. S.H. did the bioinformatics analysis. M.M. and N.F. revised and refined the manuscript, providing constructive feedback.

Compliance with Ethical Standards

Conflict of Interest The authors declare no conflict of interest.

References

1. Wright ME, Tsai MJ, Aebersold R (2003) Androgen receptor represses the neuroendocrine transdifferentiation process in prostate cancer cells. Mol Endocrinol 17:1726–1737. doi:10.1210/me.2003-0031
2. Waltering KK, Helenius MA, Sahu B et al (2009) Increased expression of androgen receptor sensitizes prostate cancer cells to low levels of androgens. Cancer Res 69:8141–8149. doi:10.1158/0008-5472.CAN-09-0919
3. Hu R, Dunn TA, Wei S et al (2009) Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. Cancer Res 69:16–22. doi:10.1158/0008-5472.CAN-08-2764
4. Demichelis F, Fall K, Perner S et al (2007) TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. Oncogene 26:4596–4599. doi:10.1038/sj.onc.1210630
5. Tomlins S a, Rhodes DR, Perner S et al (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 310:644–648. doi:10.1126/science.1117679
6. Clark J, Merson S, Jhavar S et al (2007) Diversity of TMPRSS2-ERG fusion transcripts in the human prostate. Oncogene 26:2667–2673. doi:10.1038/sj.onc.1210070
7. Weier C, Haffner MC, Mosbruger T et al (2013) Nucleotide resolution analysis of TMPRSS2 and ERG rearrangements in prostate cancer. J Pathol 230:174–183. doi:10.1002/path.4186
8. Oberley LW, Oberley TD, Buettner GR (1980) Cell differentiation, aging and cancer: the possible roles of superoxide and superoxide dismutases. Med Hypotheses 6:249–268
9. Yu JJ, Yu JJ, Mani R-SS et al (2010) An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG Gene fusions in prostate cancer progression. Cancer Cell 17:443–454. doi:10.1016/j.ccr.2010.03.018
10. Mani R-S, Iyer MK, Cao Q et al (2011) TMPRSS2-ERG-mediated feed-forward regulation of wild-type ERG in human prostate cancers. Cancer Res 71:5387–5392. doi:10.1158/0008-5472.CAN-11-0876
11. Petrovics G, Liu A, Shaheduzzaman S et al (2005) Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. Oncogene 24:3847–3852. doi:10.1038/sj.onc.1210745
12. Börno ST, Fischer A, Kerick M et al (2012) Genome-wide DNA methylation events in TMPRSS2-ERG fusion-negative prostate cancers implicate an EZH2-dependent mechanism with miR-26a hypermethylation. Cancer Discov 2:1024–1035. doi:10.1158/2159-8290.CD-12-0041
13. Schwartzman J, Mongoue-Tchokote S, Gibbs A et al (2011) A DNA methylation microarray-based study identifies ERG as a gene commonly methylated in prostate cancer. Epigenetics 6:1248–1256. doi:10.4161/epi.6.10.17727
14. Todorova K, Metodiev MV, Metodieva G et al (2016) miR-204 is dysregulated in metastatic prostate cancer in vitro. Mol Carcinog 55:131–147. doi:10.1002/mc.22263
15. Sun Y, Koo S, White N et al (2004) Development of a micro-array to detect human and mouse microRNAs and characterization of

expression in human organs. Nucleic Acids Res 32:e188. doi:10.1093/nar/gnh186
16. Todorova K, Mincheff M, Hayrabedyan S et al (2013) Fundamental role of microRNAs in androgen-dependent male reproductive biology and prostate cancerogenesis. Am J Reprod Immunol. doi:10.1111/j.1600-0897.2012.01139.x
17. Wang L, Tang H, Thayanithy V et al (2009) Gene networks and microRNAs implicated in aggressive prostate cancer. Cancer Res 69:9490–9497. doi:10.1158/0008-5472.CAN-09-2183
18. Turner DP, Findlay VJ, Moussa O et al (2011) Mechanisms and functional consequences of PDEF protein expression loss during prostate cancer progression. Prostate 71:1723–1735. doi:10.1002/pros.21389
19. Smith AM, Findlay VJ, Bandurraga SG et al (2012) ETS1 transcriptional activity is increased in advanced prostate cancer and promotes the castrate-resistant phenotype. Carcinogenesis 33:572–580. doi:10.1093/carcin/bgs007
20. Li L, Chang W, Yang G et al (2014) Targeting poly(ADP-ribose) polymerase and the c-Myb-regulated DNA damage response pathway in castration-resistant prostate cancer. Sci Signal 7:ra47. doi:10.1126/scisignal.2005070
21. Akech J, Wixted JJ, Bedard K et al (2010) Runx2 association with progression of prostate cancer in patients: mechanisms mediating bone osteolysis and osteoblastic metastatic lesions. Oncogene 29:811–821. doi:10.1038/onc.2009.389
22. Todorova K, Zasheva D, Kanev K, Hayrabedyan S (2014) miR-204 shifts the epithelial to mesenchymal transition in concert with the transcription factors RUNX2 , ETS1 , and cMYB in prostate cancer cell line model. J Cancer Res 2014:1–14. doi:10.1155/2014/840906
23. Baniwal SK, Khalid O, Gabet Y et al (2010) Runx2 transcriptome of prostate cancer cells: insights into invasiveness and bone metastasis. Mol Cancer 9:258. doi:10.1186/1476-4598-9-258
24. Husseiny MI, Kuroda A, Kaye AN et al (2012) Development of a quantitative methylation-specific polymerase chain reaction method for monitoring Beta cell death in type 1 diabetes. PLoS One 7:e47942. doi:10.1371/journal.pone.0047942
25. Metodieva G, Nogueira-de-Souza NC, Greenwood C et al (2013) CD74-dependent deregulation of the tumor suppressor scribble in human epithelial and breast cancer cells. Neoplasia 15:660–668
26. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.P.B.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. doi:10.1038/nbt.1511
27. Liu H, Sadygov RG, Yates JR (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76:4193–4201. doi:10.1021/ac0498563
28. Yoav Benjamini YH (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300
29. Gopalan A, Leversha MA, Satagopan JM et al (2009) TMPRSS2-ERG gene fusion is not associated with outcome in patients treated by prostatectomy. Cancer Res 69:1400–1406. doi:10.1158/0008-5472.CAN-08-2467
30. Hu Y, Dobi A, Sreenath T et al (2008) Delineation of TMPRSS2-ERG splice variants in prostate cancer. Clin Cancer Res 14:4719–4725. doi:10.1158/1078-0432.CCR-08-0531
31. Chow A, Amemiya Y, Sugar L et al (2012) Whole-transcriptome analysis reveals established and novel associations with TMPRSS2:ERG fusion in prostate cancer. Anticancer Res 32:3629–3642
32. Attard G, Clark J, Ambroisine L et al (2008) Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. Oncogene 27:253–263. doi:10.1038/sj.onc.1210640
33. Mertz KD, Setlur SR, Dhanasekaran SM et al (2007) Molecular characterization of TMPRSS2-ERG gene fusion in the NCI-H660 prostate cancer cell line: a new perspective for an old model. Neoplasia 9:200–206. doi:10.1593/neo.07103
34. Furusato B, Gao C-L, Ravindranath L et al (2008) Mapping of TMPRSS2-ERG fusions in the context of multi-focal prostate cancer. Mod Pathol 21:67–75. doi:10.1038/modpathol.3800981
35. Markert EK, Mizuno H, Vazquez A, Levine AJ (2011) Molecular classification of prostate cancer using curated expression signatures. Proc Natl Acad Sci U S A 108:21276–21281. doi:10.1073/pnas.1117029108
36. Nguyen P-N, Violette P, Chan S et al (2011) A panel of TMPRSS2:ERG fusion transcript markers for urine-based prostate cancer detection with high specificity and sensitivity. Eur Urol 59:407–414. doi:10.1016/j.eururo.2010.11.026
37. Yu J, Yu J, Rhodes DR et al (2007) A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. Cancer Res 67:10657–10663. doi:10.1158/0008-5472.CAN-07-2498
38. Mounir Z, Korn JM, Westerling T et al (2016) ERG signaling in prostate cancer is driven through PRMT5-dependent methylation of the androgen receptor. Elife 5:1–19. doi:10.7554/eLife.13964
39. Viré E, Brenner C, Deplus R et al (2006) The polycomb group protein EZH2 directly controls DNA methylation. Nature 439:871–874. doi:10.1038/nature04431
40. Tatematsu KI, Yamazaki T, Ishikawa F (2000) MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. Genes Cells 5:677–688. doi:10.1046/j.1365-2443.2000.00359.x
41. Borowczyk E, Mohan KN, D'Aiuto L et al (2009) Identification of a region of the DNMT1 methyltransferase that regulates the maintenance of genomic imprints. Proc Natl Acad Sci U S A 106:20806–20811. doi:10.1073/pnas.0905668106
42. Mounir Z, Korn JM, Westerling T et al (2016) ERG signaling in prostate cancer is driven through PRMT5-dependent methylation of the androgen receptor. Elife. doi:10.7554/eLife.13964
43. Farooqi AA, Hou M-F, Chen C-C et al (2014) Androgen receptor and gene network: micromechanics reassemble the signaling machinery of TMPRSS2-ERG positive prostate cancer cells. Cancer Cell Int 14:34. doi:10.1186/1475-2867-14-34
44. Wu D, Zhang C, Shen Y et al (2011) Androgen receptor-driven chromatin looping in prostate cancer. Trends Endocrinol Metab 22:474–480. doi:10.1016/j.tem.2011.07.006
45. Zhang Y, Ng HH, Erdjument-Bromage H et al (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev 13:1924–1935
46. Saito M, Ishikawa F (2002) The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2. J Biol Chem 277:35434–35439. doi:10.1074/jbc.M203455200
47. Welsbie DS, Xu J, Chen Y et al (2009) Histone deacetylases are required for androgen receptor function in hormone-sensitive and castrate-resistant prostate cancer. Cancer Res 69:958–966. doi:10.1158/0008-5472.CAN-08-2216
48. Valdez CD, Kunju L, Daignault S et al (2013) The E2F1/DNMT1 axis is associated with the development of AR negative castration resistant prostate cancer. Prostate 73:1776–1785. doi:10.1002/pros.22715
49. Valdez CD, Davis JN, Odeh HM et al (2011) Repression of androgen receptor transcription through the E2F1/DNMT1 axis. PLoS One 6:e25187. doi:10.1371/journal.pone.0025187
50. Heebøll S, Borre M, Ottosen PD et al (2008) SMARCC1 expression is upregulated in prostate cancer and positively correlated with tumour recurrence and dedifferentiation. Histol Histopathol 23:1069–1076
51. Shaikhibrahim Z, Langer B, Lindstrot A et al (2011) Ets-1 is implicated in the regulation of androgen co-regulator FHL2 and reveals specificity for migration, but not invasion, of PC3 prostate cancer cells. Oncol Rep 25:1125–1129. doi:10.3892/or.2011.1156
52. Little GH, Baniwal SK, Adisetiyo H et al (2014) Differential effects of RUNX2 on the androgen receptor in prostate cancer: synergistic

stimulation of a gene set exemplified by SNAI2 and subsequent invasiveness. Cancer Res 74:2857–2868. doi:10.1158/0008-5472.CAN-13-2003

53. Mills IG (2014) Maintaining and reprogramming genomic androgen receptor activity in prostate cancer. Nat Rev Cancer 14:187–198. doi:10.1038/nrc3678
54. Narayanan R, Jiang J, Gusev Y et al (2010) MicroRNAs are mediators of androgen action in prostate and muscle. PLoS One 5: e13637. doi:10.1371/journal.pone.0013637
55. Laxman B, Tomlins SA, Mehra R et al (2006) Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. Neoplasia 8:885–888. doi:10.1593/neo.06625
56. Laxman B, Morris DS, Yu J et al (2008) A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res 68:645–649. doi:10.1158/0008-5472.CAN-07-3224
57. Fine SW, Gopalan A, Leversha MA et al (2010) TMPRSS2-ERG gene fusion is associated with low Gleason scores and not with high-grade morphological features. Mod Pathol 23:1325–1333. doi:10.1038/modpathol.2010.120
58. Hägglöf C, Hammarsten P, Strömvall K et al (2014) TMPRSS2-ERG expression predicts prostate cancer survival and associates with stromal biomarkers. PLoS One 9:e86824. doi:10.1371/journal.pone.0086824