VILNIUS UNIVERSITY

KRISTINA DANIŪNAITĖ

DIAGNOSTIC AND PROGNOSTIC DNA METHYLATION BIOMARKERS OF PROSTATE CANCER

Doctoral dissertation

Biomedical Sciences, Biology (01 B)

Vilnius, 2017

The doctoral dissertation was prepared at the Institute of Biosciences (former Faculty of Natural Sciences), Life Sciences Center, Vilnius University, during 2012–2016.

Scientific supervisor:

Prof. Sonata Jarmalaitė, PhD (Vilnius University, Biomedical Sciences, Biology – 01 B)

VILNIAUS UNIVERSITETAS

KRISTINA DANIŪNAITĖ

PROSTATOS VĖŽIO DIAGNOSTINIAI IR PROGNOSTINIAI DNR METILINIMO ŽYMENYS

Daktaro disertacija

Biomedicinos mokslai, Biologija (01 B)

Vilnius, 2017

Disertacija rengta Vilniaus universiteto Gyvybės mokslų centro Biomokslų institute (buvusiame Gamtos mokslų fakultete) 2012–2016 m.

Mokslinė vadovė:

Prof. dr. Sonata Jarmalaitė (Vilniaus universitetas, biomedicinos mokslai, biologija – 01 B)

CONTENTS

ABBREVIATIONS

ADAMTS12	*ADAM metallopeptidase with thrombospondin type 1 motif 12* gene
ADT	Androgen deprivation therapy
APC	*Adenomatous polyposis coli* gene
AR	Androgen receptor
ARE	Androgen-responsive element
ASAP	Atypical small acinar proliferation
BCR	Biochemical disease recurrence
CCDC181	*Coiled-coil domain containing 181* gene
CD44	*CD44 molecule (Indian blood group)* gene
CDKN2A	*Cyclin dependent kinase inhibitor 2A* gene
CHD1	C*hromodomain helicase DNA-binding protein 1* gene
CGI	Cytosine-guanine dinucleotide-rich island, CpG island
CI	Confidence intervals
COPZ2	*Coatomer protein complex subunit zeta 2* gene (*MIR152* host gene)
CpG	Cytosine-guanine dinucleotide
Cq	Cycle of quantification
CRPC	Castration-resistant prostate cancer
DAPK1	*Death-associated protein kinase 1* gene
DNMT1	*DNA methyltransferase 1* gene
DRE	Digital rectal examination
EAU	European Association of Urology
EC	Endogenous control
EF	Epigenetic regulation-associated factor
EMT	Epithelial-mesenchymal transition
EPAS1	*Endothelial PAS domain protein 1* gene
ERG	*ETS transcription factor-related* gene
ESTRO	European Society for Radiotherapy & Oncology
ETS	E26 transformation-specific family of transcription factors
ETV1	*ETS variant 1* gene
ETV4	*ETS variant 4* gene
EZH2	*Enhancer of zeste homolog 2* gene
FC	Absolute fold change value
FDR	Benjamini-Hochberg false discovery rate
FFPE	Formalin-fixed paraffin-embedded sample preparation procedure
FILIP1L	*Filamin A interacting protein 1 like* gene
FOXA (FOXA1)	*Forkhead box A1* gene (protein)
GIPR	*Gastric inhibitory polypeptide receptor* gene (*MIR642A* and *MIR642B* host gene)
GO	Gene Ontology
GSEA	Gene set enrichment analysis
GSTP1	*Glutathione S-transferase pi 1* gene
H	Kruskal-Wallis's H parameter
HM27	Illumina HumanMethylation27 platform
HM450	Illumina HumanMethylation450K platform
hnRNA	Heterogeneous nuclear RNA
HPRT1	*Hypoxanthine phosphoribosyltransferase 1* gene
HR	Hazard ratio
IDH1	*Cytosolic isocitrate dehydrogenase 1* gene
IP-DNA	Immunoprecipitated methylated DNA
KCTD8	*Potassium channel tetramerization domain containing 8* gene
KDM	Histone lysine demethylase
KDM1A (KDM1A)	*Lysine demethylase 1A* gene (protein)
KDM5B (KDM5B)	*Lysine demethylase 5B* gene (protein)
KLK3	*Kallikrein related peptidase 3* gene
MAGI2	*Membrane associated guanylate kinase, WW and PDZ domain containing 2* gene
MC	Methylated control/ reference human DNA
mCRPC	Metastatic castration resistant prostate cancer

MGMT	*O^6-methylguanine-DNA-methyltransferase* gene
MIR137HG	*MIR137 host* gene
MIR155HG	*MIR155 host* gene
MIR31HG	*MIR31 host* gene
miRNA	Small noncoding regulatory RNA (microRNA)
MSP	Methylation-specific PCR
NAALAD2	*N-acetylated alpha-linked acidic dipeptidase 2* gene
NCP	National Center of Pathology (Vilnius, Lithuania)
ncRNA	Noncoding RNA
NEK9	*NIMA related kinase 9* gene
NGS	Next-generation sequencing
NKX3-1	*NK3 homeobox 1* gene
NPT	Noncancerous prostate tissue
NTC	No-template control
p14	*CDKN2A* transcript variant *p14* (*$p14^{ARF}$*)
p16	*CDKN2A* transcript variant *p16* (*$p16^{INK4a}$*)
PCa	Prostate cancer/ tumor
PCA3	*Prostate cancer associated 3* gene
PI3K	Phosphoinositide 3-kinase
PIA	Proliferative inflammatory atrophy
PIN	High-grade intraepithelial neoplasia (also HGPIN)
PMR	Percentage of methylated reference DNA, representing relative DNA methylation level
PRAD	Prostate cancer dataset of The Cancer Genome Atlas project
PRKCB	*Protein kinase C beta* gene
PSA	Prostate-specific antigen
pT	Pathological tumor stage according to Tumor, Node, Metastasis staging system
PTEN	*Phosphatase and tensin homolog* gene
QMSP	Quantitative methylation-specific PCR
qPCR	Quantitative polymerase chain reaction
RARB	*Retinoic acid receptor β* gene
RASSF1	*RAS association domain family member 1* gene
RB1	*Retinoblastoma 1* (*RB transcriptional corepressor 1*) gene
Ref-DNA	Untreated reference DNA
RIN	RNA integrity number
RP	Radical prostatectomy
R_P	Pearson's correlation coefficient
R_S	Spearman's correlation coefficient
RT	Reverse transcription
SD	Standard deviation
SE	Standard error of mean
SIOG	International Society of Geriatric Oncology
SPOP	*Speckle type BTB/POZ protein* gene
TCGA	The Cancer Genome Atlas
TMPRSS2	*Transmembrane protease, serine 2* gene
TMPRSS2-ERG	*Transmembrane protease, serine 2* and *ETS transcription factor-related* gene fusion
TNM	Tumor, Node, Metastasis staging system
TP53	*Tumor protein 53* gene
TSG	Tumor suppressor gene
TSS	Transcription start site
UC	Unmethylated control human DNA
UTR	Untranslated region
UV	Ultraviolet electromagnetic radiation
$Z_{ad.}$	Mann-Whitney's Z adjusted parameter
ZMIZ1	*Zinc finger MIZ-type containing 1* gene

INTRODUCTION

Prostate cancer (PCa) is one of the most prevalent malignancies among men with high mortality rates worldwide [1]. During the last several decades, there have been numerous advances in basic research on PCa initiation and progression, as well as novel clinical advances that have improved patients' outcome [2].

Recently, genome-scale analyses have provided novel insight into the epigenomic landscape of PCa, including DNA methylation and microRNA (miRNA) expression profiles [3-8]. Despite the validation of previously known alterations, newly discovered genetic and epigenetic features were proposed as measures of PCa aggressiveness, as well as tools for diagnosis [3,5,8,9]. However, most of these genome-wide studies have focused on comparing tumors with normal tissues rather than stratifying by cancer aggressiveness, i.e. the likelihood of spreading rapidly outside the prostate. Therefore, further analysis at a global scale is needed to increase the knowledge about the genetic and epigenetic differences between indolent and aggressive PCa.

During the last decade, the increased understanding of genetic alterations in PCa has encouraged the development of molecular biomarker systems in order to facilitate both the diagnosis of the disease and the selection of the most effective treatment scheme, as well as to avoid unnecessary clinical procedures for the patient. However, many of such tests assess similar molecular features and, thus, occupy only some overlapping clinical niches, whereas the full phenotypic spectrum of this highly heterogeneous disease might not be properly covered [10]. At the molecular level, the diversity of PCa has been proved by numerous studies reporting various PCa biomarkers, only a small part of which have been further validated in independent cohorts [3,5,8]. Therefore, considering the varying PCa incidence and mortality rates worldwide, there is a current need for thorough molecular profiling of PCa in order to get a better insight of ubiquitous genetic features that could be exploited for PCa diagnostics and prognosis.

DNA methylation, as a mechanism for down-regulation of gene expression, is an early event in tumor development and has been extensively studied in PCa. However, despite few recent epigenomic investigations, most of the previously published studies used candidate gene approach and/ or were small scale (reviewed in [11]). In PCa, methylation at promoter regions has been implicated in the silencing of over 100 different genes among which *glutathione S-transferase pi 1* gene (*GSTP1*) is the most prominent (reviewed in [11]), whereas investigation of most of the other genes resulted in ambiguous data on their potential clinical value. Aberrant DNA methylation is a promising source of PCa biomarkers, however, thorough screening and consecutive validation in clinical samples are needed aiming to identify the most perspective biomarkers in practice.

Aim and tasks

The **aim** of the present study was to identify differentially methylated genes in PCa and to determine their diagnostic and prognostic value for potential application in clinical practice.

In order to achieve the objective, the following **tasks** were carried out:

1. To identify promoter methylation frequencies of a set of tumor suppressor genes (*RARB*, *GSTP1*, *RASSF1*, *MGMT*, *DAPK1*, $p16^{INK4a}$/*CDKN2A*, and $p14^{ARF}$/*CDKN2A*) in PCa tissues and to determine their diagnostic and prognostic value.

2. To evaluate the possibility to detect promoter methylation of the genes *RARB*, *GSTP1*, and *RASSF1* in urine of PCa patients and to determine their diagnostic and prognostic value.

3. To investigate genome-wide DNA methylation profile of tumors and paired noncancerous PCa tissues in order to identify tumor-specific and biochemical disease recurrence (BCR)-specific epigenomic aberrations and altered molecular pathways.

4. To validate promoter methylation differences of a set of 10 protein-coding genes (*PRKCB*, *CCDC181*, *ADAMTS12*, *ZMIZ1*, *FILIP1L*, *NAALAD2*, *KCTD8*, *EPAS1*, *NEK9*, and *CD44*) selected from global DNA methylation profiling data, to determine their diagnostic and prognostic value, and suitability for noninvasive testing in urine.

5. To validate promoter methylation differences of a set of 5 miRNA host genes (*MIR155HG/BIC*, *COPZ2*, *MIR137HG*, *MIR31HG*, and *GIPR*) selected from global DNA methylation profiling data and to determine their diagnostic and prognostic value.

6. To analyze gene expression levels of selected miRNA targets (*DNMT1*, *KDM1A*, and *KDM5B*) and to evaluate associations with miRNA host gene promoter methylation and *TMPRSS2-ERG* expression.

Scientific novelty and practical value of the study

In the present study, for the first time, aberrant DNA methylation was analyzed in the Lithuanian cohort of PCa cases using both genome-wide and candidate-gene approaches. Using clinical samples (tissues and urine), this study increased the molecular understanding of localized PCa at the epigenetic level. Tumor suppressor genes (TSGs), widely studied in various tumor types, have been evaluated for promoter methylation in the Lithuanian PCa cohort and associated with clinical-pathological characteristics. As these genes are commonly included in molecular test systems that are currently under development, the results obtained in the present study may assist in the selection of the most suitable biomarkers in a particular clinical context.

Genome-wide microarray-based DNA methylation profiling, performed in a Lithuanian cohort for the first time, led to identification of a set of novel putative biomarkers with diagnostic and/ or prognostic value that might even surpass currently available tests. Promoter methylation of these newly identified genes was also demonstrated to be detectable in urine of PCa patients, which shows its potential application for noninvasive testing. Besides,

methylation of *RAS association domain family member 1* gene (*RASSF1*) was the first ever reported prognostic PCa biomarker in urine.

Frequently observed methylation differences at miRNA host gene loci in microarray data encouraged more thorough analysis of such genes. To date, the diagnostic and prognostic potential of miRNAs has been mostly analyzed at the miRNA expression level only. In the present study, miRNA host gene methylation status was shown to have independent prognostic value for BCR-free survival. Besides, mir-155 host gene methylation in PCa was associated with miRNA silencing for the first time. This is also the first study to associate the upregulated *lysine demethylase 5B* gene (*KDM5B*) expression with the methylated status of mir-155 and mir-137 host gene promoters.

The results are innovative, since there is a limited number of PCa biomarkers proven to be potentially useful in the clinics so far, whereas miRNA host gene methylation studies have been mostly limited to PCa cell lines. Besides, the role of miRNA host gene (*MIR137HG* in particular) methylation was proposed as an additional regulatory layer in the formation of *transmembrane protease, serine 2* (*TMPRSS2*) and *ETS transcription factor-related* (*ERG*) gene fusion (*TMPRSS2-ERG*). Although additional validation is required, the obtained results seem to be rather promising. Protein-coding and miRNA host genes with newly identified DNA methylation differences in PCa might be attractive targets for epigenetic therapeutics.

Statements to be defended

1. Promoter methylation of a combination of genes *RARB*, *GSTP1*, and *RASSF1* is a sufficiently sensitive and moderately specific molecular tool for PCa detection in tissues and urine of PCa patients.

2. *RASSF1* methylation is a significant predictor of BCR-free survival of PCa patients and, in combination with tumor stage pT, could be used as a noninvasive biomarker for follow-up of patients diagnosed with Gleason score 6 tumor.

3. Changes in methylation of protein-coding and miRNA host genes are abundant in PCa and affect specific biological and molecular processes associated with tumor development and progression.

4. Promoter methylation of a combination of protein-coding genes *PRKCB*, *CCDC181*, and *ADAMTS12* is highly sensitive and specific for PCa and might be used as a tool for diagnostics.

5. Promoter methylation of genes *PRKCB*, *CCDC181*, and *NAALAD2* in PCa tissues might be utilized as independent predictors of BCR-free survival and may increase the prognostic potential of pathological parameters.

6. Methylation of mir-155, mir-152, and mir-137 host gene promoters is specific to PCa and could be utilized for PCa diagnostics.

7. Promoter methylation of mir-155, mir-152, and mir-31 host genes have independent prognostic value for BCR-free survival of PCa patients and may increase the prognostic potential of pathological parameters.

8. Methylation of mir-137 host gene might be a causal alteration leading to the formation of *TMPRSS2-ERG* fusion in PCa through the increased *KDM1A* expression.

1. LITERATURE OVERVIEW

1.1. Prostate cancer

Prostate cancer (PCa) is a complex, multifactorial disease that continues to be a significant factor in mortality around the world. Together with lung, colorectal, and breast cancer, PCa is one of the most common types of cancer. In 2012, more than 1.1 million new PCa cases were recorded worldwide, accounting for ~15% of new cancer cases in men (World Cancer Research Fund International; http://www.wcrf.org). It is the most common non-cutaneous malignancy among men and the second leading cause of cancer death in men worldwide (Fig. 1.1) [12]. In Europe, PCa incidence rates are among the highest and are expected to increase within the next decades (European Cancer Observatory; https://www.iarc.fr).

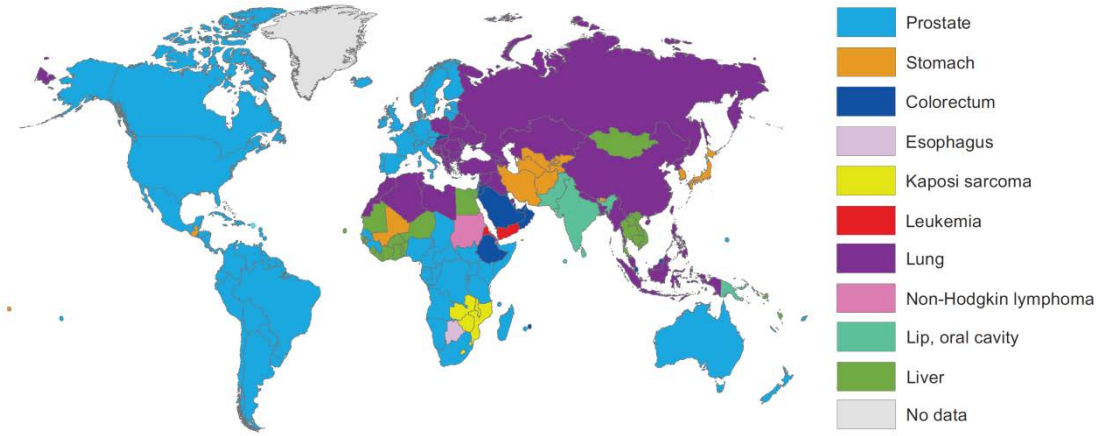


Figure 1.1. Most commonly diagnosed cancers in men in 2012 (adapted from [13]).

In Lithuania, PCa is the most common type of cancer among men and accounted for 29% of new cases diagnosed with cancer in 2012 (Cancer Registry, National Cancer Institute; http://www.nvi.lt). In 2006, the prevention program for early diagnosis of PCa was started, which led to a vast increase in numbers of newly detected PCa cases during the following year (Fig. 1.2) and, thus, made a significant impact on PCa diagnosis.

While many PCa lesions are indolent and remain localized, a subset shows an aggressive course with rapid development of metastases, which is fatal within a short time following diagnosis [14,15]. This reflects the underlying

heterogeneity of PCa and raises the need not only to accurately diagnose the disease, but also to determine its further course and select the most proper treatment strategy.

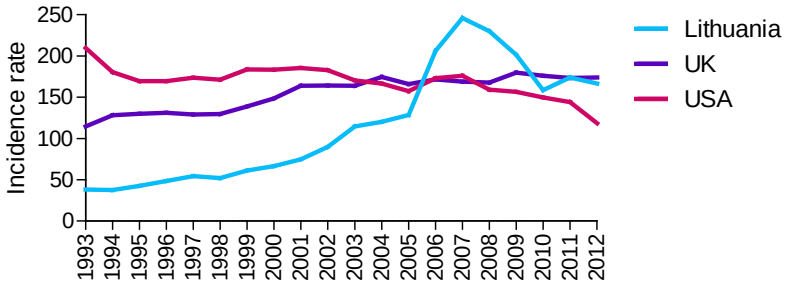


Figure 1.2. Trends of prostate cancer incidence rates in Lithuania, UK, and USA. Incidence rates are provided age-standardized per 100 000 men (European, UK, and USA age standards for Lithuanian, UK, and USA populations, respectively). Data compiled from Cancer Registry, National Cancer Institute (Lithuania; http://www.nvi.lt), European Cancer Observatory (http://eco.iarc.fr), Cancer Research UK (https://www.cancerresearchuk.org), and American Cancer Society (http://www.cancer.org).

1.1.1. Prostate gland: pathology and carcinogenesis

The human prostate is a walnut-sized gland (around 4×2×3 cm), composed of epithelial acini arranged in a fibromuscular stromal network, and weighs about 20 g. It is located in the male pelvis beneath the base of the urinary bladder and in front of the rectum. Approximately 70% of the gland consists of glandular tissue. According to a widely accepted zonal compartment system developed by McNeal *et al.*, the prostate is divided into four anatomically and clinically distinct zones ([16] and references therein). The glandular tissue is subdivided into three glandular zones – central, peripheral, and transitional – and a fourth nonglandular region, the anterior fibromuscular stroma (Fig. 1.3A). Most of the tumors (70%) arise in the peripheral zone surrounding the distal urethra, while the transitional zone accounts for 20% of cases (reviewed in [17]).

At the histological level, there are at least three distinct cell types in prostate epithelium: luminal secretory, basal, and rare neuroendocrine cells (Fig. 1.3B). Epithelial cells form prostatic glands that are surrounded by smooth muscle cells and fibroblasts [18]. Luminal cells are predominant, they

form a continuous layer of polarized columnar epithelium, produce prostatic secretory proteins, and express characteristic markers such as cytokeratins 8 and 18, as well as high levels of androgen receptor (AR). Basal cells are beneath the luminal epithelium, and express p63 and cytokeratins 5 and 14, while AR expression is at low or undetectable levels. Neuroendocrine cells are AR-negative and express endocrine markers such as chromogranin A and synaptophysin. The function of these cells is largely unclear (reviewed in [2]).

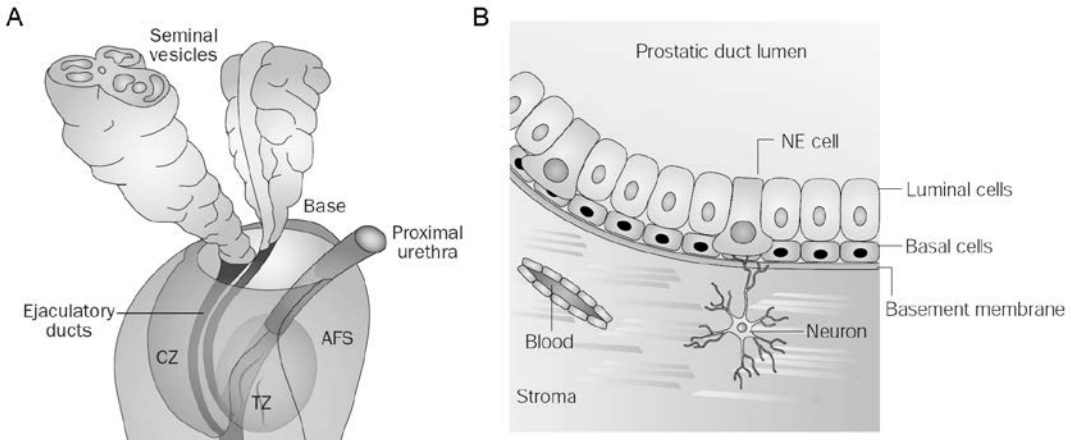


Figure 1.3. The anatomy and histology of the prostate gland. A – the anterior oblique view of the prostate; B – histology of the gland. CZ – central zone, TZ – transitional zone, PZ – peripheral zone, AFS – fibromuscular stroma, NE – neuroendocrine cell (images adapted from [16,19]).

As PCa is characterized by luminal cell expansion and absence of basal cells, luminal cells are commonly proposed to be the cells of origin for PCa [2]. Proliferation of basal cells in PCa is very rare. However, the subject is controversial as recent studies have reported both luminal and basal epithelial cells as PCa precursors in model systems [20-22]. Besides, gene expression profiling data do not support molecular subtypes of PCa based on luminal and basal differentiation as in breast or bladder cancer ([23] and references therein).

PCa is considered as a multifocal malignancy due to the common presence of multiple primary tumors (found in 60-90% of cases) that are histologically independent and often genetically distinct [17,24-26]. As discussed in [2], the

heterogeneity of PCa is potentially relevant for the understanding the difference between latent and clinical disease. In general, PCa is a slowly developing malignancy, which may remain indolent for decades.

Prostate carcinogenesis is a gradual process with a spectrum of premalignant states of different morphological, histochemical properties and invasiveness potential (Fig. 1.4) [27-30]. PCa development begins when a secreting prostate gland cell begins to proliferate uncontrollably. It is thought that many cancerous lesions give rise only to latent PCa which does not develop into clinically detectable disease as these foci might remain under active suppression sufficient to maintain them in subclinical state [2]. However, the aggressive forms show rapid growth and dissemination, which leads to lethality. Disparate tumor foci can progress at different rates depending on obtained genetic alterations that determine biological aggressiveness ([17,31] and references therein). Metastatic PCa may develop from individual clones which have gained a selective advantage during the carcinogenesis [31]. It has been reported that circulating tumor cells, which are able to initiate metastases, mostly originate from distinct foci which can be down to 0.2 mm^3 in size [32]. Recently, whole-genome sequencing of multiple metastatic lesions from the same patients revealed a common clonal origin supported by the presence of 40-90% of the total number of mutations [33].

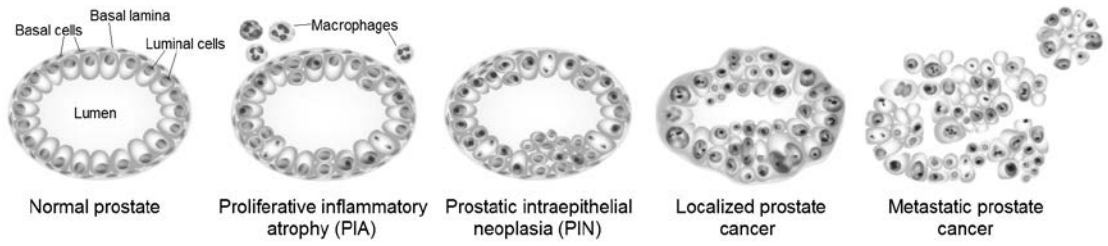


Figure 1.4. Stages of prostate cancer development. The role of proliferative inflammatory atrophy (PIA) as a stage in prostate carcinogenesis is controversial (adapted from [27]).

High-grade prostatic intraepithelial neoplasia (PIN or HGPIN) is the only widely accepted precursor of PCa (Fig. 1.4), which is characterized by progressive abnormalities of phenotype and genotype that are intermediate

between benign prostatic epithelium and cancer [28]. PIN is clinically important due to its high predictive value as a marker for PCa, especially when multifocal or observed in association with atypical small acinar proliferation (ASAP). As well as PCa, the majority of PINs develop in the peripheral zone and are multifocal. The strongest argument for implicating PIN as a precursor of PCa comes from studies showing a significantly increased prevalence and extent of PINs in tumor-containing prostates as compared to tumor-free prostates [34]. PIN and PCa are morphometrically and phenotypically similar. Besides, PIN and PCa also have similar genetic alterations (e.g. allelic loss at certain loci, *GSTP1* methylation, *TMPRSS2-ERG* gene fusion) [35,36]. PIN is most often found in areas which are in continuity with PCa ([28] and references therein). Moreover, increasing rates of angiogenesis in PIN progression are further evidence that PIN is precancerous [37]. PCa develops in most patients with PIN within 5-10 years [28,38].

Other suggested PCa precursors include proliferative inflammatory atrophy (PIA) and atypical adenomatous hyperplasia (AAH), also known as adenosis. PIA has been proposed as an intermediate step between normal state and PIN since it is quite commonly found to be merged with PIN (Fig. 1.4; [29] and references therein). It is a proliferative regenerative change caused by cellular injury, which is followed by inflammation-induced regeneration. As regenerating cells are at increased risk of mutation, this predisposes them to cancerous initiation [28]. Up to date, it is not clear whether PIA is an intermediate stage between normal prostatic epithelium and PIN, a direct precursor of PCa (bypassing PIN), or simply a coexisting lesion, as discussed in [28,29].

AAH is a small glandular proliferation found only incidentally during histological examination and may be misdiagnosed as adenocarcinoma, especially on needle biopsy, as AAH and PIA are both described as proliferations of small, pale glands. Most AAH foci are microscopic and the majority are discovered in the transitional zone. Genetic abnormalities do not

discriminate between AAH and PCa, however, morphological linkage between the two is weak to absent [39]. Therefore, the role of AAH as a precursor of PCa is disputable.

Cellular proliferation in the transitional zone might also result in benign prostatic hyperplasia (BPH) [40]. BPH is a chronic disease that exclusively affects the transitional zone, however, it does not lead to cancer. BPH and PCa are phenotypically distinct in the fact that BPH can show proliferation of the basal epithelial cell layer, whereas cancer involves pronounced expansion of luminal cells [41]. Therefore, BPH cases are often utilized as an independent control or comparison group in various molecular PCa studies. In addition, recent transcriptomic data indicate that PCa and BPH have different gene expression profiles [41].

Cell morphological changes leading to a histologically abnormal appearance of prostate tissues are preceded by a phase during which molecular alterations occur in a complete absence of any cytological or histological change [42]. Relatively large areas of a tissue/ organ are affected by a carcinogenic insult resulting in the development of multifocal independent premalignant foci and molecular lesions that precede histological change [42]. The concept of expanding field in carcinogenesis, called field cancerization or field effect, was first introduced by Slaughter *et al.* and was based on observations that cancer developed in multifocal areas and that abnormal tissue surrounded the tumor [43]. This suggested that precancerous cells were present at histologically normal field [43,44].

Extensive molecular analyses have been conducted on histologically normal prostate tissues that are tumor-adjacent or distant from cancer, commonly utilizing them as a control group when compared to cancerous lesions as it is assumed that cancer-related alterations are absent [45]. However, multiple studies showed that gene and/ or protein expression profile of tumor-adjacent normal tissue differs from normal tissue which is not

associated with malignancy [46-48]. DNA methylation-based alterations have been also reported for various genes in histologically normal prostate tissues [36,49-54] and more recently confirmed by genome-wide methylation studies [3,7,53]. Nevertheless, as compared to other tumors, relatively little is known about field cancerization in PCa. However, the presence of genetic and epigenetic field effects in PCa is widely accepted, which supports the notion that histologically normal tissues, especially those adjacent to tumor, might not be the most ideal controls for molecular studies [48]. On the other hand, in clinical practice, detection of PCa-specific molecular alterations in normal-appearing prostate tissues could potentially be used to increase the diagnostic sensitivity or justify the need of repeat biopsies in cases with previous negative biopsies [52,55-58].

1.1.2. Prostate cancer diagnosis, staging, and prognosis

PCa is usually suspected on the basis of digital rectal examination (DRE) and/ or an increased prostate-specific antigen (PSA) level. In order to confirm the presence of the tumor, men must undergo biopsy [59]. PCa can be detected by DRE when the tumor volume is ≥0.2 mL, therefore, as an independent variable, PSA is a better predictor than DRE or transrectal ultrasound (TRUS) [59].

The introduction of a blood PSA test for routine use has made a great impact on early diagnosis of PCa and, at present, it is the first-line serum biomarker used for PCa detection [60]. However, the lack of a cut-off value associated with high sensitivity and specificity of this test also increased the risk of overdiagnosis, i.e. identification of small localized low grade tumors, many of which presumably represent indolent or latent forms of PCa that may not develop into symptomatic disease during the lifetime [2,61,62]. Overdiagnosis by PSA testing, which in turn leads to unnecessary biopsies, is estimated from 5% to 45% for men at 50-69 years of age [63]. Besides, men might have PCa despite low PSA levels, e.g. the prevalence of PCa was 6.6% among men with PSA level ≤0.5 ng/μL and increased up to 26.9% in the PSA

range 3.1-4.0 ng/μL [64]. Another inherent limitation of PSA testing is its moderate specificity in the setting of PCa screening, as other noncancerous conditions (e.g. BPH, prostatitis) may result in elevated PSA level [65].

If there are clinical indications for the repeat biopsy after a previously negative one, additional information may be gained by currently available molecular tests, some of which are included in the Guidelines on Prostate Cancer by European Association of Urology (EAU), European Society for Radiotherapy & Oncology (ESTRO), and International Society of Geriatric Oncology (SIOG) [59]. The use of most of such diagnostic tests is suggested before making the decision for repeat biopsy as this procedure might be preventable if test results come negative (Table 1.1).

Table 1.1. Currently available molecular tests for prostate cancer.

Test	Biomarker(s)	Test substrate	Suggested application	Test type	Reference
Prostate Health Index (PHI) test*	Total, free, intact PSA	Serum	After negative biopsy	Diagnostic	[66,67]
4Kscore Test*	Total, free, intact PSA, and hK2	Serum, plasma	After negative biopsy	Diagnostic	[68]
ConfirmMDx*	Methylation of *RASSF1*, *GSTP1*, *APC*	Biopsy	After negative biopsy	Diagnostic	[56]
Progensa*	Expression of *PCA3*	Urine	After negative biopsy	Diagnostic	[57]
Prostate Core Mitomic Test	Large-scale mitochondrial DNA deletions	Biopsy	After negative biopsy	Diagnostic	[69]
ProMark	Quantification of 8 proteins	Biopsy	Active surveillance	Prognostic	[70]
Prolaris*	Expression of 31 cell-cycle related genes	Biopsy	Watchful waiting	Prognostic	[71]
OncotypeDX Prostate*	Expression of 12 cancer-related genes	Biopsy	Active surveillance	Prognostic	[72]
Decipher	Array-based expression of 22 genes	Biopsy	Post-prostatectomy decision making	Prognostic	[73]
CELLSEARCH CTC	Count of circulating tumor cells	Blood	Monitoring of CRPC patients	Prognostic	[74]
ExoDx Prostate (IntelliScore)	Exosomal RNA (*ERG*, *PCA3*, and *SPDEF* as EC)	Urine	Identification of high-grade PCa at the time of biopsy/ surgery	Prognostic	[75]

CRPC – castration-resistant prostate cancer, EC – endogenous control, PCa – prostate cancer.
*Included in the EAU-ESTRO-SIOG Guidelines on Prostate Cancer [59].

When PCa is confirmed, the treatment strategy must be appointed. The treatment decision is based on tumor classification systems, which combine patients with similar clinical outcome. The Tumor, Node, Metastasis staging system (TNM) for PCa was first adopted in 1992 by the American Joint Committee on Cancer and the International Union Against Cancer [76]. The TNM system is based on three major parameters: T (tumor) is used to evaluate the size and extent of the primary tumor, N (node) shows the degree of spread to regional lymph nodes, and M (metastasis) describes the presence of distant metastasis. According to the current EAU-ESTRO-SIOG Guidelines on Prostate Cancer [59], the 2009 TNM classification is currently used (Table S1). Additionally, PCa can be further grouped into clinical stages according to TNM parameters (Table 1.2).

Table 1.2. Prostate cancer stages based on Tumor, Node, Metastasis (TNM) parameters.

Stage	T	N	M
I	1 or 2a	0	0
II	2b or 2c	0	0
III	3	0	0
IV	4	0	0
	Any T	1	0
	Any T	Any N	1

Currently, there are no definitive markers that could differentiate PCa into indolent and aggressive better than the histopathological Gleason grading system using biopsy tissues. It was developed by Donald Gleason between 1966-1974 [77]. Gleason score describes a prostate tumor according to its microscopic appearance. Grades 1 to 5 are assigned to the histologic patterns of the tumor, while adding the grades of the most and second most common patterns results in Gleason score ranging from 2 to 10. The primary and secondary grades are reported in addition to the Gleason score. A global Gleason score is given for multiple tumors, but a separate tumor focus, especially with a higher Gleason score, is usually also reported.

Over the years, the Gleason system has been revised in accordance with the advances in PCa diagnosis and therapy. Treatment decisions using a single

Gleason score of 7 fail to recognize that tumors evaluated with 3 + 4 and 4 + 3 are very different in their prognosis. Furthermore, in practice the lowest Gleason score assigned to cancer is 6, which often leads to an incorrect assumption on the part of patients that their cancer is in the middle of the scale [78]. To overcome these and other deficiencies of the system, the major revisions have been made in 2005 [79] and more recently in 2014 [80], when the grade groups according to International Society of Urological Pathology (ISUP) were introduced for the use together with Gleason score (Table 1.3).

Table 1.3. International Society of Urological Pathology (ISUP) 2014 grade groups according to Gleason score (adapted from [80]).

ISUP grade group	Gleason score
1	≤6
2	7 (3 + 4)
3	7 (4 + 3)
4	8 (4 + 4 or 3 + 5 or 5 + 3)
5	9 or 10 (4 + 5 or 5 + 4 or 5 + 5)

Accurate differentiation between aggressive and indolent PCa is critically important to reduce overtreatment so that patients with life-threatening disease will undergo proper treatment while those with indolent disease will be spared the radical therapy [10]. Presently, the Gleason system with 2005 modifications is the strongest prognostic factor for clinical course of PCa and response to treatment. Gleason grade 4 or 5, particularly if present in >5% of the PCa volume, is a prognostic indicator of cancer progression [59].

Other factors associated with PCa prognosis include tumor volume, location, and surgical margin status. Tumor volume is one of the most controversial parameters that is sometimes used for prognostic purposes as tumors tend to have irregular shapes with infiltrative growth patterns. Tumor volume is variously determined as tumor percentage or tumor dimension, and there is no consensus that this parameter by itself correlates with adverse findings in surgical material [81]. On the other hand, the location of the tumor in the transitional zone is associated with a decreased risk of disease progression [81]. Surgical margin status is an independent risk factor: although

evidence is considered insufficient, some indications may be provided regarding tumor multifocality [59,82].

Clinical observations also suggest that the increased number of neuroendocrine-like cells in PCa is associated with poor prognosis, however, the data are inconclusive [83]. It has been hypothesized that these cells arise by neuroendocrine differentiation or transdifferentiation process from luminal cells [84]. They do not proliferate and lack the expression of PSA and AR [85], thus, these cells are very resistant to treatment. Besides, neuroendocrine-like cells are known to express relatively high levels of various anti-apoptotic genes and secret a number of growth factors and peptide hormones, supporting the growth of surrouding tumor cells [85]. In general, neuroendocrine-like cells may contribute to poor PCa prognosis not only by surviving therapies, but also by stimulating the proliferation of existing cancer cells.

As PSA provides poor differentiation of PCa prognosis [65] and so far there are no other molecular biomarkers with high prognostic performance confirmed by prospective multicenter studies, the current inability to distinguish aggressive from indolent or latent PCa at diagnosis remains one of the major clinical challenges [2,61]. Although there are several available prognostic biomarker tests (Table 1.1), recommendations regarding their routine application are still being awaited [10,59].

1.1.3. Prostate cancer progression and treatment

As already mentioned, the clinical course of PCa is heterogeneous, ranging from indolent tumors requiring no therapy during the patient's lifetime to highly aggressive PCa developing into a metastatic disease. While localized cancer is usually treated with radical prostatectomy (RP) or radiotherapy, these regimens are followed or substituted with androgen deprivation therapy (ADT) and chemotherapy in cases of advanced PCa [59]. Generally, aggressive tumors should be identified and treated early, while avoiding overtreatment of clinically insignificant cases.

According to the annual report of American Cancer Society, the probability of developing invasive PCa during 50-59 years interval is 2.1% (1 in 48) for men, who are free of cancer at the beginning of this age decade, and roughly doubles with each next decade reaching 10.0% (1 in 10) for men in their 70s [12]. Most of the times, patients become long-term survivals and, as such, develop multiple medical needs over time, some of which are ageing-related. This leads to poor symptom control and low quality of life [86].

Men with localized PCa can have very different prognoses and, thus, a wide selection of treatment options are available. In contrast to other types of tumors [87], PCa lacks distinct histopathological subtypes that would differ in disease prognosis and response to treatment. Therefore, treatment options are mainly based on above-mentioned clinical-pathological parameters, according to which PCa is divided into risk groups (Table 1.4), with regard to symptoms and patient's quality of life. Depending on the severity of PCa, current treatment options include active surveillance, RP, hormone therapy, brachytherapy, high-intensity focal ultrasound, cryotherapy, radiotherapy, chemotherapy, bone-directed treatment, and others [23,59,88].

Table 1.4. Risk groups for biochemical disease recurrence of localized and locally advanced prostate cancer (PCa; adapted from [59]).

Group		Definition
Localized PCa	Low-risk	PSA <10 ng/µL and Gleason score <7 and T1-2a
	Intermediate-risk	PSA 10-20 ng/µL or Gleason score 7 or T2b
	High-risk	PSA >20 ng/µL or Gleason score >7 or T2c
Locally advanced PCa		Any PSA, any Gleason score, T3-4 or N1

T and N are parameters of Tumor, Node, Metastasis staging system (TNM), PSA – prostate-specific antigen.

Active surveillance is offered to patients with assumed low-risk PCa aiming to avoid unnecessary and potentially harmful treatment and achieve the proper timing for curative treatment, which is prompted by predefined thresholds [23,59]. Studies show that men with low-risk PCa can safely avoid treatment with a risk of death from PCa being as low as 1% at 10 years (discussed in [23]). RP is a common treatment used to try curing localized PCa completely, while aiming to avoid urinary incontinence and erectile

dysfunction, whenever possible. It involves removal of the entire prostate, together with resection of both seminal vesicles and surrounding tissue [59].

Despite the current technical progress in curative procedures, such as RP, there is a risk of cancer recurrence after treatment. BCR, described by a rising PSA level after treatment, indicates treatment failure and precedes metastatic progression. Therefore, after BCR has been confirmed, it is important to determine if the recurrence has developed at local or distant sites, while some patients with BCR after RP do not develop clinically evident disease recurrence at all [59]. Generally, many patients in the high-risk group (Table 1.4) likely have micrometastases (and, therefore, a higher risk of dying from PCa) at the time of PSA increase, while those in the low-risk group likely have a slow-growing local recurrence only [89]. Besides, the rate of PSA increase is also indicative of the type of disease recurrence: rapidly increasing PSA usually indicates distant metastases, whereas slow or later increment most likely shows local recurrence [90]. Therefore, prognostic biomarkers that would let identify aggressive PCa at its early stages are under investigation.

As the growth and survival of normal prostate and prostate cancer cells is dependent on androgen, the frontline treatment of patients having advanced PCa includes ADT. This treatment modality was first accomplished via surgical castration (discussed in [91]). Now it is usually achieved by medical castration, i.e. by either suppressing the secretion of testicular androgens or inhibiting the action of circulating androgens using anti-androgens [59]. Castration reduces the serum testosterone to very low levels. However, the progression of advanced PCa to castration-resistant prostate cancer (CRPC) in 1-2 years is inevitable and ultimately fatal [92]. The treatment of CRPC is based on expensive systemic therapy with cytostatic agents and next-generation targeted therapy (e.g. abiraterone acetate, enzalutamide), but only a part of CRPC patients respond to this treatment positively [92]. Furthermore, the benefits of AR-directed therapies are usually short-lived and resistance occurs invariably, leading to an incurable disease.

CRPC is characterized by increased levels of PSA despite low levels of testosterone. Until recently it has been thought to be a hormone refractory state, however, it is currently recognized that AR function is almost never lost (discussed in [93]). Development of CRPC and its further progression is inevitably linked to various molecular alterations in AR signaling pathway (e.g. AR gene amplification, increased expression of AR and/ or its regulators, gain of function mutations, post-translational mechanisms of AR activation, alternative slicing, etc.) [94]. Besides, although androgens are mainly produced by testes, extragonadal sources (like adrenal, intraprostatic, and intratumoral) are clinically important, because they are likely to be sufficient to support CRPC progression even though the secretion from the testes is suppressed [91].

PCa deaths are typically the result of metastatic CRPC (mCRPC), with the median survival of ≤2 years of men with mCRPC [91,95,96]. In Lithuania, about 500-600 men die of PCa each year (Cancer Registry, National Cancer Institute; http://www.nvi.lt). BCR after RP precedes clinically detectable metastases by 7-8 years. However, despite new improvements in imaging techniques PCa recurrence is still difficult to detect early, when salvage treatment is most beneficial [97,98]. PCa almost invariably metastasizes to bone (90%), while other common sites are lung (46%), liver (25%), pleura (21%), and adrenals (13%) [99]. Bone metastasis are a major cause of reduced quality of life and death of patients with mCRPC. Despite the recent insight on molecular factors in PCa that might be associated with bone tropism, the key mechanisms responsible for the dissemination of PCa cells to bone remain unresolved (reviewed in [100]).

Although several clinical parameters may help to predict the course of PCa and assist in treatment selection, once BCR is diagnosed the prompt initiation of salvage therapy is critically important. Elevated PSA indicates disease relapse only after its actual occurrence, which means that PCa has already developed into the stage eventually leading to death. Therefore, prognostic

biomarkers that could lead the way for personalized treatment decisions are needed.

1.1.4. Prostate cancer etiology

High PCa incidence rates and relatively long patients' survival contrast with the lack of evident etiology. The heterogeneity of the PCa, as well as the complex structure of the gland itself, has imposed difficulties into searching for an etiological origin of the malignancy [30]. The only established risk factors are age, African ethnicity/ race, a family history of PCa, and certain inherited genetic factors. Advancing age and ethnicity are the strongest nonmodifiable factors which are considered to play a pivotal role in PCa development [12,88].

Age. Age at diagnosis of cancer is a well-recognized prognostic factor. PCa is identifiable in some very young men and at an increasing rate with age, suggesting that PCa is usually a slowly developing disease with a long preclinical phase [101]. Younger men (≤50 years old) are more likely to have better survival compared with older men [14]. This has been explained by better tolerance for aggressive therapy, earlier recognition of the disease, as well as attendant lower stage and/ or grade of PCa [102]. However, among men with high-grade and locally advanced PCa, the youngest men have been shown to have a poor prognosis compared with older men [14].

Ethnicity. The risk of PCa is ~70% higher in men with African ancestry than in Caucasians, while native Chinese, Japanese and other east south-east Asian men have the lowest reported rates in the world [12,103]. Interestingly, Caribbean men of African descent have the highest documented PCa incidence and mortality rates in the world, while tumor stage and grade at diagnosis are highest among men in sub-Saharan Africa [12,104,105]. Reasons for the disparities in PCa diagnosis and treatment are complex and multifactorial, including socioeconomic, psychosocial, cultural, and educational factors, which affect access to care and treatment options. Furthermore, genetic factors and the anatomical location of a tumor within the prostate are considered to

contribute to worse PCa outcomes among men of African ancestry [104]. This might be partly explained by the fact that African and African-American men usually have higher PSA levels at the time of diagnosis, which might be associated with genetic variations. In particular, specific single nucleotide polymorphisms (SNPs) in *kallikrein related peptidase 3* (*KLK3*) gene, encoding PSA, are more common among African-Americans and, therefore, might account for increased serum PSA levels [106].

Family history. Next to age and ancestry, the family history is considered as one of the most important risk factors for PCa. Genetic studies suggest that familial predisposition might be responsible for up to 40-50% of PCa cases (reviewed in [107]). Familial PCa was first described in the 1950s [108]. Since then various epidemiological studies have reported the familial clustering of PCa, however, the magnitude of the estimated risk various rather significantly. In general, the risk of PCa increases from 2.5-fold with a single affected first-degree relative (i.e. father, son, or brother) to 5-fold with ≥2 affected first degree relatives [109,110]. The risk is also higher when a proband is younger [109,111]. PCa is commonly considered as one of the most heritable forms of cancer [112]. The *hereditary prostate cancer 1 susceptibility* locus (*HPC1*) in *ribonuclease L* gene (*RNASEL*) on chromosome 1q24-25 has been the first one linked to families with a high risk of PCa. Soon after, other loci have been discovered on various chromosomes (chr2, chr6, chr8, chr10, chr13, chr17, chrX, and others) [113,114].

Genetic susceptibility. As all cancers, PCa is a genetic disease that can be causes by both germline and somatic genetic alterations. Genetic susceptibility and epigenetic changes may predispose individuals to PCa. Mutations and/ or abnormal gene expression associated with classic genes such as *tumor protein 53* (*TP53*), *phosphatase and tensin homolog* (*PTEN*), *breast cancer 1* and *2* (*BRCA1* and *BRCA2*) genes might contribute to prostate carcinogenesis [115,116]. The completion of the Human Genome Project and the subsequent International Haplotype Map Project led to the identification of >40 PCa

susceptibility loci which explained up to 25% of PCa risk [117]. Increased PCa risk-associated conditions also include Lynch syndrome with characteristic *DNA mismatch repair gene* (*MMR*) mutation [118].

Various other modifiable environmental and occupational risk factors have also been implicated, but mechanisms underlying the link of these factors and PCa development remain unclear. Several studies suggest that obesity and smoking are not associated with the overall risk of developing PCa, but may increase the risk of developing the aggressive form of the disease [12]. According to a recent umbrella review by Markozannes *et al.*, associations between diet, body size, physical activity and the risk of PCa did not present strong evidence [15]. Interestingly, only the association of height with the risk of PCa was highly suggestive, i.e. a 4% higher risk of PCa was observed per 5 cm of greater height. Associations with body-mass index (BMI), weight, alcohol intake, dietary calcium and PCa incidence and/ or mortality were reported as moderately suggestive, while intake of dairy foods, waist circumference, waist to hip ration, and some other factors were supported by only weak evidence [15]. There is also some indication that firefighters can be exposed to chemicals which might increase their risk of PCa [119]. Other studied risk factors include sexually transmitted infections and vasectomy, however, no firm conclusions have been reached so far [120,121].

At present, due to the lack of conclusive data no definite recommendations could be provided for PCa prevention [59].

1.2. Genetic features of prostate cancer

Over the past years significant efforts have been invested in elucidating the molecular basis of PCa. In several recent studies, multiple recurrent genetic rearrangements have been identified in PCa, including various mutations, copy number changes, gene fusions, and other alterations [4,9,35,115,122]. A large number of studies have pointed towards AR as a central player in prostate carcinogenesis [123]. Presently, the widespread use of genome-wide profiling,

primarily aiming to determine molecular PCa subtypes, has significantly increased the understanding of genetic and epigenetic features of this malignancy.

1.2.1. Role of androgen receptor in prostate cancer

The role of the AR in the development and progression of prostate cancer has led to increasing interest in this nuclear receptor. The AR gene (Xq11-12) encodes a cytoplasmic steroid receptor, a member of the nuclear receptor superfamily, which acts as a ligand-dependent hormone-inducible transcription factor. Ligand binding causes AR translocation to the nucleus, where AR dimers interact with target genes via androgen-responsive elements (AREs) [124]. Until the widespread use of whole-genome technologies, only a few androgen-regulated genes have been identified, mostly limited to *KLK3* (*PSA*), *kallikrein related peptidase 2* (*KLK2*), and *NK3 homeobox 1* genes (*NKX3-1*) [125]. Increased serum levels of the important biomarker PSA suggest that AR activity is elevated in PCa patients. Recent gene expression profiling studies have identified >1000 androgen-responsive genes, of which over 200 genes are considered as a core set and include *prostate cancer associated 3* (non-protein coding; *PCA3*), *folate hydrolase 1* (*FOLH1*; also known as *prostate-specific membrane antigen*, *PSMA*), *transmembrane protease, serine 2* (*TMPRSS2*), *ETS transcription factor-related* (*ERG*), *dual specificity phosphatase 1* (*DUSP1*), *microseminoprotein β* (*MSMB*), *sorbitol dehydrogenase* (*SORD*), *cysteine rich secretory protein 3* (*CRISP3*), *a prostate-specific gene with cell growth-promoting function* (non-protein coding; *PCGEM1*), and other commonly studied genes ([126,127] and references therein).

Previous studies have shown that AR binds primarily to enhancers rather than promoter regions and that 86-96% of AR-binding sites are at nonpromoter regions. Furthermore, loops of intervening DNA are formed when AR-bound enhancers are interacting with promoters of AR-regulated genes (discussed in [128]). Chromatin looping has been also reported to be responsible for the

formation of androgen-responsive gene rearrangements, such as *TMPRSS2-ERG* fusion, a feature of PCa [129,130].

The amplitude of AR-mediated signaling is modulated by AR coregulators, which may both enhance (coactivators) and reduce (corepressors) AR transcription activity. More than 170 proteins are currently known as AR coregulators, whose role has been assessed in PCa. The most studied AR coregulators include forkhead box A1 (FOXA1), GATA binding protein 2 (GATA2), and erythroblast transformation-specific proto-oncogene 1 (ETS1), of which FOXA1 is considered to have an essential role in AR transcriptional activity [124].

Histone-modifying enzymes compose probably the major group of AR coregulators (discussed in [123]). Studies have shown that histone lysine methylation status contributes significantly to AR signaling which correlates with an increased expression of several histone lysine demethylases (KDMs), such as lysine demethylase 1A (KDM1A/LSD1) and 4C (KDM4C/JMJD2C) in prostate cancer [131-133]. Most of KDMs physically interact with AR to facilitate its recruitment to AREs and subsequent activation of the downstream target genes [132]. AREs are particularly enriched for histone 3 lysine 4 mono- or di-methylation (H3K4me1 or H3K4me2), generally associated with transcriptional enhancers [134].

The heterogeneity of prostate cancer suggests that there could be multiple initiating events leading to inactivation of tumor suppressors and/or activation of tumor promoters/oncogenes that could at some point of disease progression cross-talk with AR (reviewed in [135]).

1.2.2. Common genetic alterations and affected pathways

Prostate cancer is characterized by a variety of somatic mutations. Despite the genes and pathways that are deregulated in various cancer types, there are some distinctive genetic aberrations observed in PCa.

Unlike other cancers, only few recurrent and/ or actionable protein-coding mutations are known in PCa, however, a large proportion of tumors harbors gene fusions. One prominent example is the AR-dependent up-regulation of members of the E26 transformation-specific (ETS) family of transcription factors by gene fusions between the AR-regulated *TMPRSS2* gene promoter and the coding region of the ETS family members, *ERG* and *ETS variant 1* (*ETV1*) genes, which have been observed in ~50% of all PCa cases [129,136]. These fusions confer androgen responsiveness to ETS transcription factors, which leads to cell-cycle progression [137]. The overexpression of *ERG* might be sufficient to initiate the development of neoplastic lesions in the prostate. Other rearrangements are less frequent [138]. Recent studies have provided significant insight into cancer-associated molecular mechanisms in *TMPRSS2-ERG*-positive tumors, however, prostate carcinogenesis in fusion-negative cases remains unclear. Borno *et al*. showed that fusion-positive PCa contains more differentially methylated regions than fusion-negative tumors, with a significant enrichment in methylation of homeobox (HOX) gene promoters [139].

The phosphoinositide 3-kinase (PI3K) signaling pathway is activated by various lesions and affects cell proliferation, invasion, and survival. This pathway is altered in 25-70% of PCa. The tumor suppressor gene *PTEN*, the most important negative regulator of PI3K pathway, is one of the most commonly mutated genes in PCa. As *PTEN* mutations are more frequently identified in metastases than primary PCa, it is suggested that this gene has a critical role in PCa progression [140]. Like *PTEN*, the *PH domain and leucine rich repeat protein phosphatase 1* gene (*PHLPP1*), suggested to play a redundant role with *PTEN*, is recurrently deleted in PCa and tightly correlates with *TP53* deletion [141]. Disruptive rearrangements may also inactivate *membrane associated guanylate kinase, WW and PDZ domain containing 2* (*MAGI2*) gene, encoding a tumor suppressor protein interacting with PTEN [9].

Another commonly mutated gene in PCa is *speckle type BTB/POZ protein* (*SPOP*) gene, which encodes a substrate binding unit of a cullin-based E3 ubiquitin ligase. Mutations in this gene are the most common point mutations in PCa, with the frequency of 6-15%, and are found exclusively in the substrate-binding cleft of SPOP, indicating altered substrate binding in PCa [142,143].

Mutations in several genes involved in chromatin modifications have been identified in PCa, including *lysine demethylase 6A* (*KDM6A*), *lysine methyltransferases 2C* (*KMT2C*) *and 2D* (*KMT2D*), also known as *myeloid/lymphoid or mixed-lineage leukemia 3* (*MLL3*) and *2* (*MLL2*), genes, respectively [142,144]. Deletions involving the *chromodomain helicase DNA-binding protein 1* gene (*CHD1*) are associated with ETS fusion-negative PCa and occur at 10-25% frequency in both primary and metastatic PCa. Such tumors have been shown to contain significant increase in genomic rearrangements [145]. Overexpression of *enhancer of zeste homolog 2* gene (*EZH2*) in PCa is associated with aggressive course of the disease by acting through the activation of AR and other transcription factors [146].

Among the genes involved in cell cycle control, mutations of *retinoblastoma 1* gene (*RB1*) and *TP53* have been commonly observed in PCa. While *TP53* mutations are recurrently detected in both localized and advanced PCa, *RB1* is more commonly mutated (or deleted) in CRPC cases [122,144]. Other signaling pathways shown to be involved in prostate cancer initiation and progression include the MAPK/ERK and RAS/RAF pathways, although their role in PCa is less well elucidated (discussed in [115]) and, thus, requires further investigations, as well as numerous other low-frequency alterations.

1.2.3. Genomic profile of prostate cancer in “-omics” era

Recent years have broadened the understanding of the significance of genetic alterations in PCa. In the first reported whole-genome PCa sequencing by Berger *et al.*, a median of 3866 putative somatic base mutations per tumor

were identified [9]. The mutation frequency in PCa (0.33-0.94 somatic mutations per Mb) is lower than in lung (3.8/Mb) or breast (~1/Mb) cancer and, therefore, does not explain the frequently observed deregulation of many genes [4,147].

Whole-genome data, obtained by Berger *et al.* [9], also indicated that complex genomic rearrangements may have a critical role driving prostate carcinogenesis. Moreover, multiple rearrangements in genes that are spatially localized together was noted [9], suggesting that genome-wide approaches are required to unmask the full spectrum of mechanisms responsible for PCa development and progression.

In a recent study, Fraser *et al.* [143] not only confirmed frequent gene fusions in PCa, but also showed that large genome-scale rearrangements in localized intermediate-risk tumors are much more common than initially thought. A clustered-mutation phenomenon, called kataegis, commonly occurring within 10 kbp, was detected in 23% of tumors and was preferentially found in PCa samples with *CHD1* or *SPOP* mutations [143]. Chromosome shattering, or chromotripsis, another type of genomic rearrangement where all or part of a chromosome undergoes massive DNA breakage, partial deletion, and random rejoining of retained fragments, was identified in 20% of tumors. The prevalence of such genomic rearrangements in localized PCa highlighted the differences from the mutational profile of mCRPC, where aberrations in single genes are observed recurrently [143,148].

The whole-genome analysis approach enabled to define the sequence of molecular changes in PCa development. As reported by Baca *et al.*, *TMPRSS2-ERG* fusion, mutations in *SPOP*, *FOXA1*, and *NKX3-1* are early events in prostate carcinogenesis, which are followed by *CHD1* and *TP53* alterations and then by inactivation of *PTEN* [145]. This supports the notion that loss of *PTEN* function triggers the development of aggressive PCa [140].

A comprehensive molecular analysis of 333 primary prostate carcinomas, as a part of The Cancer Genome Atlas (TCGA) project (https://cancergenome.nih.gov) [4], revealed a molecular taxonomy in which 74% of prostate tumors fell into one of seven subtypes defined by specific genetic alterations, most of which have been reported previously. Four of the subtypes were characterized by cancer-driving gene fusions: *ERG* fusions (46%), *ETV1* (8%), *ETS variant 4* (*ETV4*; 4%), and *Fli-1 proto-oncogene ETS transcription factor* (*FLI1*; 1%) fusion or overexpression. The other three were defined by cancer-driving mutations in *SPOP* (11%), *FOXA1* (3%), and *cytosolic isocitrate dehydrogenase 1* gene (*IDH1*; 1%). Gene fusion-driven and mutation-driven PCa demonstrated different genomic, as well as epigenomic, profiles. The rest 26% of PCa cases could not be categorized into a molecular subtype, all together confirming that PCa is a highly heterogeneous cancer type [4].

Taken together, the recent genome-wide studies not only delivered robust insights into PCa genome, but also provided clinically actionable data, which could be applied for making treatment decisions as well as for the development of novel therapeutic agents. Due to the heterogeneous nature of PCa, continuity of genome-wide research would be key to the future improvements in the personalization of PCa management.

1.3. Aberrant DNA methylation and prostate cancer

Epigenetic modifications are defined as reversible biochemical changes affecting gene expression without altering the primary DNA sequence. DNA methylation at the 5’ carbon of cytosine (5-mC) in cytosine-guanine dinucleotides (CpGs) is the most intensively studied epigenetic mechanism for control of gene expression, which is nearly ubiquitous in multicellular organisms and essential for the normal development in mammals. CpGs in the genome are distributed unevenly: more than half of the genes contain short CpG-rich regions known as CpG islands (CGIs), while the rest of the genome is depleted for CpGs [149]. CGIs span the transcription start site (TSS) of roughly half of the human genes and mostly represent genes which are actively

expressed or poised for transcription [150]. Methylated DNA is recognized by proteins containing methyl-CpG binding domains (e.g. MBD1, MBD2, MBD4, MeCP2) or C2H2 zinc fingers (e.g. ZBTB33/Kaiso, ZBTB4, ZBTB38), which determine transcriptional repression of genes with promoter methylation [150].

CGI hypermethylation is frequent in cancer and is often associated with the silencing of tumor suppressor genes (TSG) and downstream signaling pathways [151]. However, regions with CpG content undergo hypomethylation, resulting in an overall decrease of methylation level in cancer cells. As an exception, CpG-poor distal enhancers that are unmethylated in normal cells often gain methylation in cancer (Fig. 1.5).

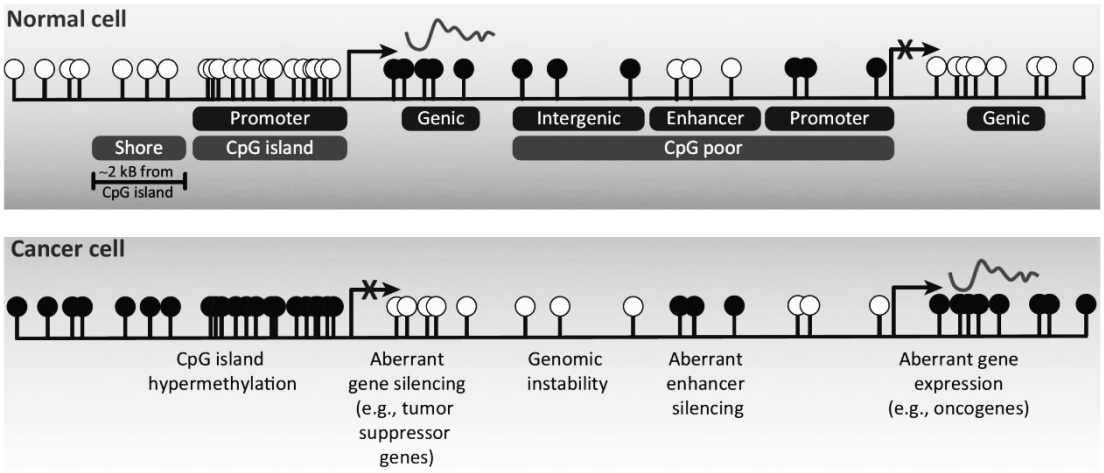


Figure 1.5. A schematic representation of major DNA methylation changes that occur in cancer cells (adapted from [151]). Black circle – methylated CpG, white circle – unmethylated CpG, curved line – expressed transcript.

Aberrant promoter hypermethylation is the most frequent and best-characterized epigenetic hallmark in human malignancies, including PCa [152]. During prostate carcinogenesis epigenetic changes in TSGs occur earlier than genetic aberrations and are more consistent among tumors than mutations [153,154]. This is suggested to be a key part of PCa development and progression [11]. Due to the relatively low rate of somatic mutations in PCa [122] and the reported overexpression of DNA methyltransferases [7], gene promoter methylation, as a mechanism for downregulating their expression, has been extensively studied in PCa.

Regarding the implementation possibilities in clinical practice, DNA methylation has several advantages over other commonly used biomarkers. In contrast to mRNA and many proteins, DNA is much more stable both *in vivo* and *ex vivo*, and can withstand harsh conditions for prolonged periods. Moreover, it can be amplified for increased sensitivity, thus, allowing measurements on limited amounts of test samples. Applicability to a wide variety of clinical samples (biopsy specimens, body fluids) and relatively simple, as well as inexpensive, analysis techniques, such as quantitative PCR, make DNA methylation changes an attractive source to search for PCa biomarkers for use in clinical practice.

1.3.1. Methylation of tumor suppressor genes in prostate cancer

The majority of previous DNA methylation studies of PCa have focused on the analysis of candidate genes. These studies reported over a hundred genes that are hypermethylated in PCa as compared to benign prostate samples. Many of these hypermethylated genes are protein-coding and regulate cell cycle, apoptosis, DNA repair, hormonal response, and invasion/ metastasis (Table 1.5). However, because of high inter-individual variations, only some of them are currently recognized as putative diagnostic biomarkers of PCa, most notably *glutathione S-transferase pi 1* (*GSTP1*), *RAS association domain family member 1* (*RASSF1*), and several other genes, while evident prognostic DNA methylation biomarkers are scarce.

GSTP1 is the most commonly studied TSG in PCa, particularly with regard to diagnostic application [155]. It is involved in detoxification process and elimination of potentially genotoxic foreign compounds by conjugating glutathione to them. Therefore, it protects cells from DNA damage. Due to consistently frequent promoter methylation in PCa (up to 95%) and only rare presence of this alteration in benign prostate tissues, numerous studies have reported *GSTP1* as a biomarker that could be used for PCa detection (Table 1.5). Besides, methylated *GSTP1* promoter has been commonly detected in samples from patients diagnosed with PIN [156], indicating that this epigenetic

alteration is an early event in PCa development. Significant correlations between *GSTP1* methylation and clinical factors of poor prognosis were also identified [157,158], however, evidence regarding the prognostic value of *GSTP1* methylation are so far conflicting [159-164].

Table 1.5. Promoter methylation frequencies of commonly studied genes by targeted approach in primary prostate cancer.

Gene symbol	Functional category	Methylation frequency*	References
GSTP1	DNA repair	36-95%	[157,161,164-171]
APC	Invasion/ metastasis	27-100%	[157,161,162,164-166,168,169,172,173]
RASSF1	Cell cycle control, apoptosis, DNA repair	53-99%	[157,164-166,170, 174-176]
RARB	Hormonal response	40-95%	[162,165-167,170,177]
MGMT	DNA damage response/ repair	0-30%	[157,164,165,167,173]
p16/ CDKN2A	Cell cycle control	3-77%	[157,164-166,178]
EDNRB	Cell cycle control, cell adhesion	49-100%	[162,164,170,173]
TIMP3	Invasion/ metastasis	0-97%	[157,164,167,173]
CD44	Cell cycle control, cell adhesion	19-72%	[170,171,173,179]
CDH1	Cell adhesion, invasion, metastasis	0-61%	[164-166,171]
p14/ CDKN2A	Cell cycle control	0-35%	[157,164,178]
DAPK1	Apoptosis	0-36%	[164,165,167]
PTGS2	Cell cycle control	65-88%	[162,164,169]
TIG1	Hormonal response	53-96%	[162,172,177]

*Only studies that analyzed ≥40 prostate cancer samples are included.

RASSF1 is another frequently studied TSG in PCa with diverse functions, including regulation of apoptosis, proliferation, microtubule dynamics during mitotic progression, association with DNA repair proteins, and others (discussed in [180]). Methylation of CGIs within *RASSF1* promoter is frequently observed in PCa (Table 1.5) and was reported to be associated with decreased gene expression [180]. In PCa, *RASSF1* methylation is associated with aggressive course of the disease [158,165], advanced pT [157], and Gleason score [181], although more recent analysis did not show significant correlation with the latter [175]. Data on associations between *RASSF1* methylation and PSA levels are also contradictory [157,158,165,175]. According to the recent meta-analysis by Pan *et al.*, *RASSF1* methylation in PCa was 14.7 times more frequent than in control cases and was associated with cancer risk when analyzed in tissues, as well as in urine or blood [182].

Besides, in this meta-analysis, methylation of *RASSF1* was significantly associated with Gleason score, but not with pT or PSA level [182]. Although several studies have reported association of *RASSF1* methylation with PCa progression (using various clinical endpoints), the data are inconclusive, probably due to limited sample sizes in most of the studies [159,160,164,175].

Retinoic acid receptor β gene (*RARB*) encodes a nuclear receptor, which is an important mediator in regulating cell growth and differentiation. It is expressed in most human tissues. Frequent hypermethylation of this gene has been observed in PCa (Table 1.5), but rarely in normal or BPH tissues [157,165,167]. Similarly to *GSTP1* and *RASSF1*, promoter methylation of *RARB* has been also detected in PIN samples, making this gene one of the most suitable biomarkers for early PCa detection [167,183]. In a recent study by Moritz *et al.*, high methylation levels of *RARB* were associated with BCR and lymph node involvement [163]. On the contrary, earlier studies could not detect a significant association between *RARB* methylation and disease progression [160,162,184].

CD44 molecule (Indian blood group) gene (*CD44*) encodes an integral membrane glycoprotein involved in cell adhesion and migration. Inactivation of this gene is thought to occur later in PCa development and, thus, is considered to be important for the disease progression and metastasis [179]. Aberrant methylation of this gene has been observed in PCa at various frequencies (Table 1.5) and was also found to be predictive of BCR [171,184]. However, another independent study did not confirm this [173]. Therefore, the importance of *CD44* methylation in prostate carcinogenesis remains uncertain.

Other noteworthy genes include *adenomatous polyposis coli* (*APC*), *O^6-methylguanine-DNA-methyltransferase* (*MGMT*), *death-associated protein kinase 1* (*DAPK1*), and *cyclin dependent kinase inhibitor 2A* (*CDKN2A*) genes (Table 1.5). *APC* encodes a multifunctional protein that acts as a tumor suppressor which is involved in cell migration, adhesion, and mitosis. *APC*

methylation has been reported at early PCa stages and the frequency increased with disease progression [158,161]. However, its methylation was also detected in BPH, which might indicate the lack of specificity of this biomarker [162]. Methylation of *MGMT*, coding a DNA repair protein, was reported in various tumors, however, in PCa its diagnostic and/ or prognostic value is elusive [157,165,173]. *DAPK1* is an important regulator of apoptosis, but the reported methylation frequencies in PCa are also moderate (Table 1.5). *CDKN2A* locus encoding several transcript variants, which differ in their first exon and, thus, have different promoters, namely $p16^{INK4a}$ (*p16*) and $p14^{ARF}$ (*p14*), can be aberrantly expressed in PCa [185]. However, data regarding the methylation frequencies of the two promoters are contradictory and, despite a few small-scale outlier studies, methylation rates at these loci are low (Table 1.5). Nevertheless, a recent meta-analysis indicated that methylation of *p16* was a risk factor for PCa [186].

The extensive studies of the diagnostic potential of promoter methylation in PCa have led to the development of the first commercial epigenetic test "ConfirmMDx for Prostate Cancer" (MDxHealth, Irvine, CA, USA; Table 1.1). It addresses false-negative biopsy concerns and evaluates methylation of the genes *GSTP1*, *RASSF1*, and *APC* [155], exploiting the previously discussed concept of field effect in PCa. After the clinical validation in European [55] and U.S. cohorts [56], which showed a negative predictive value of 88%, this assay has already been included in the EAU-ESTRO-SIOG Guidelines on Prostate Cancer 2016 as a tool for gaining additional information before making the decision for a repeat biopsy [59].

In summary, although a lot of genes have been studied for promoter methylation in PCa, most of them have insufficient sensitivity and/ or specificity. Therefore, after initial validation steps, only a few appear to retain their diagnostic and/ or prognostic potential. Even though the performance parameters can be increased adding other genes into a multiple-gene assay, the "ConfirmMDx for Prostate Cancer" test is the only fully developed and

commercially available DNA methylation-based assay for PCa so far. This stimulates to search for yet unidentified biomarkers with excelling performance.

1.3.2. Prostate cancer methylome and identification of novel DNA methylation biomarkers

Over the most recent years, it has been shown that the epigenome of cancer cells is altered not less than at the same extend as their genome [3,9,187]. Until recently, DNA methylation analyses were carried out on particular genetic loci by assays utilizing traditional PCR or sequencing of genomic DNA, previously modified by bisulfite treatment or enriched by immunoprecipitation. The advent of high-density microarray and next-generation sequencing (NGS) technologies to the market has enabled genome-wide analysis of DNA and histone modifications and led to new insights into epigenomic profiles of PCa.

Current epigenome-wide analyses, that were initially directed towards discovery of yet unknown biomarkers [5,8], revealed distinctive DNA methylation profile of PCa [3,4]. Specific patterns of promoter DNA methylation have been described around transcription start sites, where methylation was observed at promoters regardless of the presence/ absence of a CpG island [3]. Importantly, promoter-associated CGI methylation significantly increased in accordance to PCa progression [3,188], although similar numbers of methylation events in PCa and benign adjacent tissues were observed within intronic and intergenic loci [3]. Moreover, differentially methylated regions specific to aggressive PCa were identified to be proximal to the genes with distinct functions from regions shared by indolent and aggressive disease [189].

Furthermore, TCGA project revealed a diversity of methylation changes that defined four epigenetically distinct subtypes of primary PCa. In particular, *ERG* fusion-positive tumors were subdivided into two subtypes according to methylation, and epigenetic patterns largely differed from those of *ETV1* or

ETV4 fusion-positive tumors [4]. However, another study reported no obvious tumor subtypes according to the methylation profile only [7]. In TCGA data, tumors with mutant *SPOP* and *FOXA1* exhibited similar epigenetic profiles, whereas *IDH1* mutant tumors were distinguishable for strongly elevated genome-wide DNA methylation and possessed the highest number of silenced genes [4]. PCa assessment at a genome-wide scale has not only confirmed the previously proposed key genes for molecular subtyping of PCa [129,142], but has also shown the diversity among the subtypes. Moreover, it also led to identification of novel epigenetic alterations further elucidating the extent of PCa heterogeneity and providing the resource for continued investigation.

Although using rigorous selection criteria, epigenome-wide studies have produced numerous novel candidate PCa biomarkers. Many of these genes possess functions that are not clearly related, if known at all, with neither normal prostate development nor carcinogenesis. Therefore, such putative biomarkers not only require thorough validation, but also insights into their potential mode of action are desirable.

In a recent study, Mahapatra *et al.* analyzed methylation profiles in four different PCa groups (198 RP samples in total) aiming to identify biomarkers 1) for early detection, 2) associated with recurrence, 3) distinguishing increasing PSA without or with metastases, and 4) associated with metastatic spread, which altogether resulted in over 200 genes, and only a few were further validated (Table 1.6) [8].

In another study by Kim *et al.*, the list of 2481 cancer-specific methylation events was narrowed down to 3 genes, of which only one has not been reported previously [3]. Kobayashi *et al.* found 56 genes, most of which also had not been previously identified, that were better identifiers of PCa than *GSTP1* (Table 1.6) [7]. Other recent studies identified *EFEMP1*, *FZD1*, *CYB5R2*, and many other genes as novel methylation biomarkers of PCa development and/ or progression [190,191].

Table 1.6. Several novel putative DNA methylation biomarkers of prostate cancer identified by genome-wide methylation analysis and validated in at least one cohort.

Gene symbol	Gene name	Group comparison	Technology	Reference
HAAO	*3-Hydroxyanthranilate 3,4-dioxygenase*	PCa *vs.* matched normal tissues	HM27	[8]
CRIP1	*Cysteine rich protein 1*	Recurrent *vs.* nonrecurrent		
FLNC	*Filamin C*			
RASGRF2	*RAS protein specific guanine nucleotide releasing factor 2*	Recurrent *vs.* nonrecurrent; clinical recurrence *vs.* BCR; systemic *vs.* local recurrence		
PHLDA3	*Pleckstrin homology like domain family A member 3*	Clinical recurrence *vs.* BCR		
BCL11B	*B-cell CLL/lymphoma 11B*	Systemic *vs.* local recurrence		
WFDC2	*WAP four-disulfide core domain 2*	PCa *vs.* benign tissues	NGS	[3]
CAPG	*Capping protein (actin filament), gelsolin-like*	Normal v*s.* PCa *vs.* CRPC	RRBS	[188]
RARRES	*Retinoic acid receptor responder 2*			
FILIP1L	*Filamin A interacting protein 1 like*			
CGRRF1	*Cell growth regulator with ring finger domain 1*	PCa *vs.* adjacent normal *vs.* healthy donors	Microarrays	[53]
ZNF296	*Zinc finger protein 296*	PCa *vs.* adjacent normal	HM27	[7]
MCAM	*Melanoma cell adhesion molecule*			
HOXD3	*Homeobox D3*	PCa *vs.* reference DNA	Microarrays	[5]
BMP7	*Bone morphogenetic protein 7*			

BCR – biochemical recurrence, CRPC – castration-resistant prostate cancer, HM27/ HM450 – Illumina Infinium HumanMethylation27/ 450K BeadChip, respectively, NGS – next-generation sequencing, PCa – prostate cancer, RRBS – reduced representation bisulfite sequencing.

The rapid advancement of genome-wide analysis technologies resulted in identification of a large pool of methylated genes in PCa and, thus, suggested novel potential biomarkers for PCa diagnosis and prognosis. Although this significantly increased the knowledge of epigenetic changes in PCa in general, most of the recently published studies used relatively small PCa sample cohorts for validating the diagnostic and/ or prognostic potential of the identified targets or lack this analysis step at all. Therefore, speeding up the route of novel epigenetic biomarkers from bench to bedside should be the major focus in future studies.

1.3.3. DNA methylation biomarkers in urine

PCa-derived methylated DNA is easily detectable in body fluids, such as urine [192-196], blood [197,198], serum [194,199], plasma [195,200], ejaculate [194], or prostate secretions [201]. This potentially allows for development of noninvasive or minimally invasive molecular tests, which are expected to replace or at least to augment the use of invasive biopsy. Liquid biopsy could be scheduled more frequently, which is especially important during PCa treatment through providing timely evidence of disease recurrence or resistance (discussed in [202]). DNA methylation in body fluids from early stage PCa patients may better reflect all tumor foci unlike tissue biopsy, which poorly accounts for PCa heterogeneity. Regarding prostate anatomy and the common tumor localization in its peripheral zone, which surrounds distal urethra, urine is the most suitable body fluid for liquid biopsy as it is easily obtainable and biomarkers are less diluted than in serum or plasma [196].

Methylated DNA in voided urine from PCa patients was first analyzed by Cairns *et al.* [193]. Although the authors detected *GSTP1* methylation only in 27% of cases, the assay was highly specific as no altered methylation was detected in BPH [193]. Later studies reported *GSTP1* methylation in 30-83% of urine samples, with specificity for PCa reaching 98% [194-196,203,204]. As discussed by Truong *et al.*, the varying sensitivity is most probably affected by the variability in the process of urine collection: although the reported methylation frequencies were low in the earliest studies, the sensitivity was improved to 75% by prostatic massage [205]. According to the meta-analysis by Wu *et al.*, the pooled specificity of *GSTP1* methylation in urine was around 90% regardless of the methylation analysis method, however, the pooled sensitivity ranged from 45% to 75% depending on the method used, although no particular trend regarding the method was identified [206].

In order to overcome limitations of a single-gene testing, methylation of various gene panels have been evaluate for PCa detection urine. Hoque *et al.* examined a nine gene panel, of which *GSTP1*, *MGMT*, *p16*, and *p14* showed a

combined sensitivity of 87%, while the specificity of the assay reached 100% [204]. However, analyzing a panel of genes *GSTP1*, *RASSF2*, *TFAP2E*, and *HIST1H4K*, Payne *et al.* did not identify any gene combination that could outperform a single biomarker [196].

In the study by Roupret *et al.*, the high sensitivity (86%) and specificity (89%) was observed analyzing a combination of top four most commonly investigated genes in PCa, i.e. *GSTP1*, *RARB*, *RASSF1*, and *APC* [203]. In the multicenter study by Vener *et al.*, the performance of a panel of genes *GSTP1*, *RARB*, and *APC* was evaluated in urine under different storage and shipment conditions (234 samples in total) [192]. The sensitivity (55% and 53%) and specificity (80% and 76%) values were comparable in both cohorts, indicating that DNA methylation assays in urine are rather robust [192]. The same combination of genes, i.e. *GSTP1*, *RARB*, and *APC*, was further analyzed in the following study by Baden *et al.* and revealed that such an assay (referred to as ProCaM™) might aid in the biopsy decision-making process in conjuction with current screening algorithms [207]. Specifically, the assay might improve the diagnostic accuracy in high-risk men with PSA in the range 2.0-4.0 ng/mL and in low-risk men with PSA in the range 4.1-10.0 ng/mL [207].

In contrast to a relatively large number of studies investigating diagnostic biomarkers, only a few have reported the prognostic potential of DNA methylation biomarkers in body fluids of PCa patients. Despite the reported *GSTP1*, protocadherin 17 (*PCDH17*), and protocadherin 10 (*PCDH10*) methylation in serum [208-210], to date only two studies have identified prognostic PCa biomarkers in urine. In the most recent study by Jatkoe *et al.*, genes *GSTP1* and *APC* were analyzed in urine of 665 men, undergoing prostate needle biopsy, and accurately identified (negative predictive value 100%) men with the low-risk PCa (i.e. without adverse pathology) who could be candidates for active surveillance [211]. In another study, Zhao *et al.* found that a four gene panel (*GSTP1*, *APC*, *CRIP3*, and *HOXD8*) was able to predict

reclassification of PCa patients (all with Gleason score 6 tumor) that were on active surveillance and, thus, disease progression over time [212].

Although major progress has been made in developing noninvasive DNA methylation-based tests for PCa diagnosis, overcoming low sensitivity remains the main issue, even using combinations of various genes (discussed in [205]). Thus, to date none of them are comercially available. Nevertheless, the specificity reaching 100% is the major advantage of such biomarkers, posing a minimal risk of false positives, while a reliable urine-based biomarker for PCa detection and/or prognosis is yet to be identified.

1.4. MiRNA genes, regulation of their expression, and functional importance

MiRNAs are approximately 20-22 nt single-stranded noncoding RNAs (ncRNAs) that regulate physiological and pathological processes at the post-transcriptional level [213,214]. Traditionally, miRNAs are thought to repress the expression of a protein-coding gene by binding the target gene's complementary sequence at the 3' untranslated region (3'-UTR) with the miRNA's seed region (from 2^{nd} to 7^{th} or 8^{th} consecutive nt from the 5'-end of the mature miRNA [213]. More recently, miRNAs have also been shown to downregulate or upregulate their targets by binding to 5'-UTR, promoter, and coding sequences (discussed in [215]). During the last decade, numerous studies have reported differences in miRNA expression profile that were associated with PCa development and progression (reviewed in [216]), however, the mechanisms of their regulation have been less much addressed.

1.4.1. Genomic location of miRNA genes

Based on their genomic location, miRNAs are classified as intragenic (mostly intronic) or intergenic. In mammalian genomes, the latter are located in poorly annotated regions and are transcribed from their own promoters as primary transcripts (pri-miRNAs) [217], whereas roughly about half or more miRNAs are hosted inside protein-coding genes or genes encoding for other

ncRNA classes like small nucleolar RNAs (snoRNAs), long intergenic ncRNAs (lincRNAs), or other unspecified ncRNA classes (Fig. 1.6) [218,219]. Sometimes a miRNA gene can have a "mixed" location, i.e. it can be located either in an exon or an intron of the same or different (overlapping) host gene transcripts [218]. According to the canonical model of intronic miRNA biogenesis, the intron of the miRNA-harboring host gene is considered to be the pri-miRNA for the intronic miRNAs, and the completely spliced heterogeneous nuclear RNA (hnRNA, also termed pre-mRNA) becomes the protein-coding mRNA [220,221]. It has been shown that miRNAs are processed from intronic regions of hnRNAs before the catalysis of splicing [220,222].

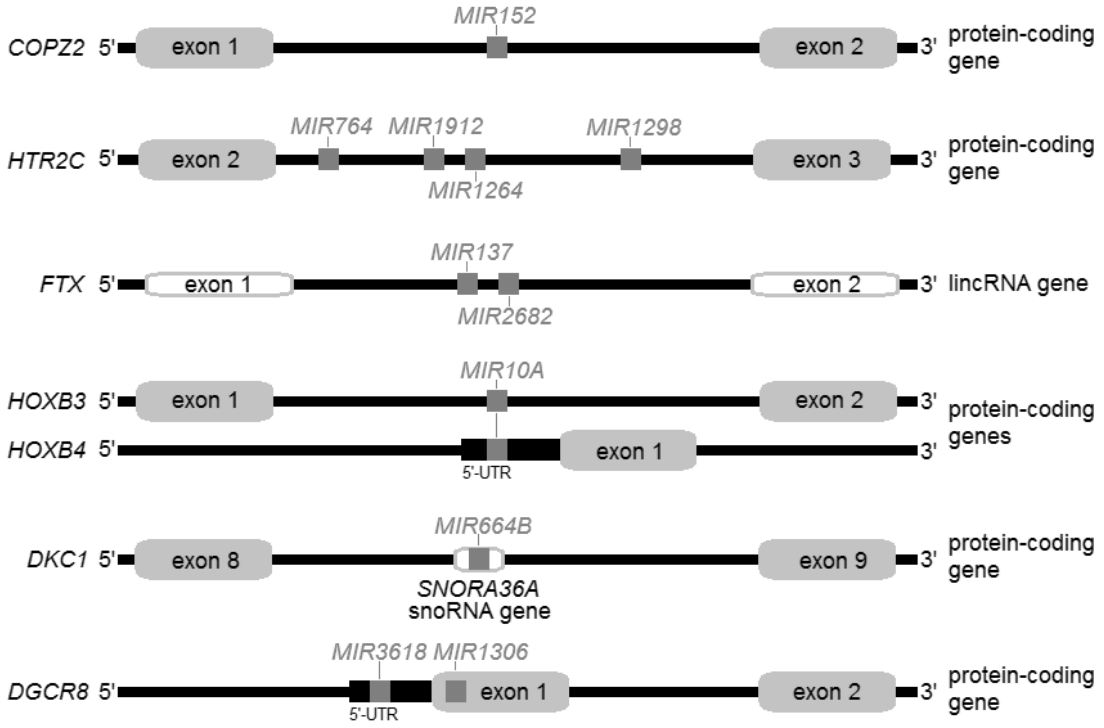


Figure 1.6. Schematic representation (not scaled) of co-location of miRNA genes and their host genes (adapted from [219]). 5'-UTR – 5' untranslated region, lincRNA – long intergenic noncoding RNA, snoRNA – small nucleolar RNA.

Intragenic miRNAs were found distributed among all human chromosomes, although several (chr14, chr19, and chrX) comprised less miRNAs than others [219]. Interestingly, miRNAs are also located within genes encoding for components of the miRNA biogenesis machinery (e.g. *DGCR8*, *DICER1*, and *SND1*) [219].

MiRNA genes are more frequently located in short host genes than expected by chance, what is hypothesized as a favorable evolutionary feature due to the gene's interaction with the pri-miRNA splicing mechanism [223]. On the other hand, the miRNA host introns ("mirtrons") are very biased at the 5'-end of the mRNA, indicating that intronic miRNAs are biased at the 5'-end of the host gene. As it is well known that most of the first introns in genes are much larger than others and certain genes have enhancer and silencer sequences within the first introns, it is proposed that miRNAs located in the first introns are more likely to have a strong functional role [222,223].

Some of human miRNAs (roughly 10%) are located in clusters. According to miRBase release v21 (http://www.mirbase.org), there are 34 clusters consisting of ≥3 miRNAs (inter-miRNA distance <5000 bp) in the human genome (assembly version GRCh38). As miRNA clusters can overlap with a single host gene and, thus, a particular host gene can contain multiple miRNA genes, the total number of host genes is lower than the number of intragenic miRNAs [219]. Clustered miRNAs are generally thought to have a common primary transcript and, therefore, are co-regulated and co-expressed [218,219,224].

1.4.2. Regulation of miRNA expression

While intergenic miRNAs have their own unique promoters, expression of intronic miRNAs has been shown to be in concordance with the expression of their respective host genes, especially if the miRNA gene and its host gene share the same strand orientation [213,219,225-227]. Moreover, evolutionary conserved miRNA genes tend to be co-expressed with their host genes, even though the non-conserved miRNAs dominate in the human genome [228]. Besides, several studies have shown a functional link between miRNAs and their host genes [225,229]. Interestingly, genes that are highly co-expressed with their miRNA gene were shown to be more likely predicted as miRNA targets [230]. The possibility of existence of independently transcribed human intronic miRNAs have been also proposed based both on bioinformatics tools

and experimental evidence ([221] and references therein). However, since most intronic miRNA genes do not have their own promoters and other regulatory elements, they may have a stronger tendency to be located towards the 5‘-end of their host genes, and efficiently utilize the regulatory signals within the 5‘-UTR of the host genes [222,223]. Additionally, it has been shown that most of the predicted regulatory regions (>94%) of intronic miRNA genes overlapped with their host gene promoters [231].

1.4.3. Epigenetic silencing of miRNA genes by DNA methylation

Recently, epigenetic silencing of microRNAs (miRNAs) has emerged as an important mechanism of oncogenesis [232]. According to a recent study, 81.2% of protein-coding miRNA host genes, with a miRNA gene at their 5’ end, were located ~500 bp downstream of CpG islands [233]. Besides, miRNA expression is more commonly downregulated than upregulated in human cancers. Taken together, it seems likely that DNA methylation might be the underlying mechanism for the frequent down-regulation of miRNAs.

Decreased expression of some miRNAs is currently known to be in association with CGI hypermethylation. MiRNAs showing aberrant gene methylation in cancer include miR-124, miR-137, miR-31, miR-34, miR-9, and others (discussed in [232]). In PCa, methylation of mir-130b~310b cluster, mir-193b, mir-1258, mir-200c, mir-141, and several others has been observed [234-236]. Besides, as recently reported by Torres-Ferreira *et al.*, miRNA gene methylation is also detectable in urine of PCa patients [235]. The analysis revealed that methylation levels of mir-193b, mir-129-2, and mir-34b/c, both in tissue and urine, might be potential PCa biomarkers for diagnosis and prognosis [235]. Furthermore, promoter methylation assay of mir-129-2 and mir-663a in urine successfully distinguished patients with urothelial carcinoma from those with other urogenital carcinomas, including PCa [237].

It has been also noted that epigenetic regulation might be responsible for downregulated expression of co-located miRNAs together with their host

genes [229]. In previous studies, several miRNA genes together with their host genes cells (e.g. *MIR152* & *COPZ2*; *MIR191* & *DALRD3*; *MIR126* & *EGFL7*) have been shown to be repressed either by CpG island methylation or/ and histone modification in the promoter region [226,229,238]. According to Godnic *et al.*, about half of epigenetically silenced miRNA genes are located within the 5'-UTR or in the first intron or exon of their host genes, which again suggests the possibility of shared promoter regions [219].

From a clinical point of view, miRNA gene and/ or host gene methylation could be a useful biomarker for PCa detection and prediction of its further progression.

1.4.4. Functional importance of miRNA deregulation in cancer

A particular miRNA can interact with multiple mRNAs and, therefore, it has a potential to regulate a large network of targets. At present, 2588 mature human miRNAs (1881 precursors) have been described (miRBase release 21; http://www.mirbase.org), while tremendous efforts to identify their numerous targets are ongoing.

Depending on the targets being regulated, miRNAs act as tumor suppressors (miR-137, miR-152, miR-31), oncogenic miRNAs (miR-21, miR-9), or have mixed roles (miR-125b, miR-155) [239]. Down-regulation of a specific miRNA might result in the increased expression of multiple target genes or reactivation of previously silenced sequences. Such up-regulation of genes coding epigenetic factors, i.e. epigenetic modulators and other proteins that are critical for enzymatic complexes, leads to global changes in the epigenome and, consequently, tumor development [240,241]. Overexpression of DNA methyltransferase 1 (*DNMT1*) has been linked to silencing of various tumor suppressor genes in cancer cells by their promoter methylation [232]. Abnormal histone methylation pattern associated with alterations in expression of histone modifying genes, such as histone lysine demethylases (KDMs), is also widely recognized [240].

Restoration of tumor-suppressive miRNAs, as well as inactivation of tumor-promoting miRNAs, offers huge prospects in cancer treatment and drug development. Therefore, the role of various epigenetic mechanisms, such as miRNA silencing by methylation, responsible for aberrant expression of these epigenetic factors are being investigated. However, recent data show that the underlying regulatory mechanisms might be much more complicated than anticipated. In the study by Godnic *et al.*, each of the residing miRNAs was predicted to target genes which also host other miRNA genes [219]. Moreover, if a particular miRNA targets a gene from the miRNA processing machinery, this might indicate a negative regulatory loop, possibly associated with the disrupted processing of miRNAs [219].

Currently, the epigenetic inactivation of tumor-suppressive miRNAs is already recognized as one of the hallmarks of cancer and exciting developments of the use of such biomarkers are expected in the future [242]. It is anticipated that studies focusing on relationship between epigenetic regulation and miRNAs might lead not only to the discovery of novel biomarkers, but also to the identification of potential therapeutic targets.

2. STUDY COHORTS AND METHODS

2.1. Patients and samples

Approval to conduct biomedical research (2007-11-23 Nr. 50; with extension 2011-09-07 Nr. 6B-11) was obtained from Lithuanian Bioethics Committee and all patients gave informed consent for participation. The present study was part of the large-scale PCa biomarker research conducted according to standardized protocols of sample collection and processing as reported previously [24,136,243,244].

In the present study, 311 PCa and 35 BPH patients were involved in total. Prostate tissue samples were collected from RP material of 248 PCa patients and from open prostatectomy samples of 17 BPH patients, who underwent surgery at the Urology Center of Vilnius University Hospital Santaros Clinics (former Santariskiu Clinics) from 2008 to 2014. Tissues were sampled and tumor cellularity was evaluated by expert pathologist at National Center of Pathology (NCP). Noncancerous prostate tissues (NPT) collected from 72 PCa patients were also available for the investigation, including 35 tumor-adjacent NPTs. Urine (~30 ml) was collected during prostatectomy from 253 PCa and 32 BPH patients. None of the patients had preoperatively received radiotherapy, chemotherapy, or hormone therapy.

Gleason score was determined according to International Society of Urological Pathology (ISUP) 2005 guidelines and ISUP grade groups were formed according to ISUP 2014 recommendations [80,81]. In agreement with EAU-ESTRO-SIOG Guidelines on Prostate Cancer 2016, BCR was defined by two consecutive PSA values of >0.2 ng/mL and rising [59]. Follow-up data for all the cases involved in the study were updated in March 2016. Full follow-up data were available for 90.0% (280/311) of patients with a mean follow-up time of 3.8±1.9 years. For the methylation analysis in formalin-fixed paraffin-embedded (FFPE) samples and urine implemented during the first stage of the investigation, the follow-up data revised in July 2013 was used.

Clinical-pathological and molecular (*TMPRSS2-ERG* fusion) characteristics of the study subgroups are detailed in Table S2. The identification of *TMPRSS2-ERG* fusion transcript status was reported previously [136,243].

2.2. Sample preparation for nucleic acid extraction

Up to 60 mg of fresh-frozen tissue samples were submerged in liquid nitrogen and mechanically homogenized into powder using cryoPREP™ CP02 Impactor with tissueTUBE TT1 (Covaris, Woburn, MA, USA; Fig. 2.1A and B). FFPE tissue blocks of ~4 mm^3 (Fig. 2.1C) were cut with a blade into small pieces and deparaffinized three times with 500 µL of 100% xylene (Carl Roth, Karlsruhe, Germany) for 10 min at 55 °C followed by treatment with ≥96% ethanol (Carl Roth) for 10 min at 55 °C (repeated three times in total). Then samples were incubated at 37 °C until ethanol was fully removed.

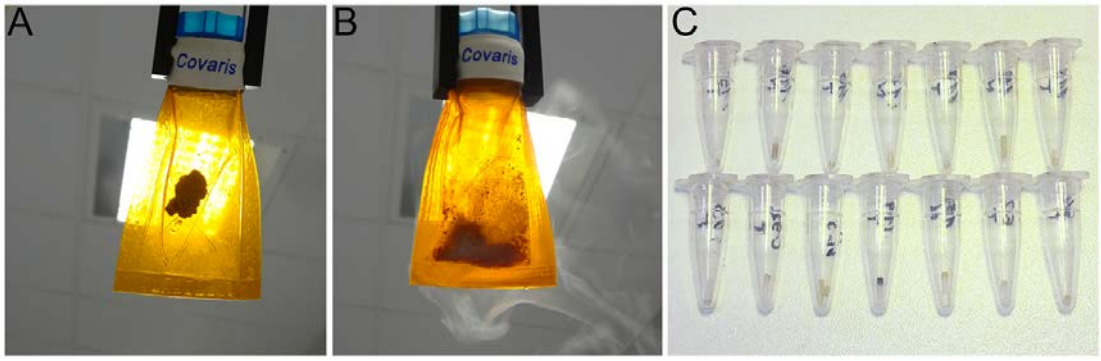


Figure 2.1. Prostate tissue samples used for molecular analysis. Fresh-frozen prostate tissue sample in a specialized bag-type tube is shown in front of a light source prior (A) and immediately after homogenization (B). Traces of nitrogen vapor are visible. C – formalin-fixed paraffin-embedded (FFPE) prostate tissue blocks (in 0.2 mL PCR microtubes).

Urine samples were centrifuged at 1000 rpm for 15 min at 4 °C (Hettich® Universal 320R Centrifuge, DJB Labcare, Buckinghamshire, United Kingdom) at NCP, then sediments were washed twice with 1× PBS and resuspended in the same buffer for storage at -70 °C until use. Immediately prior to lysis step, about 1.5-2.0 mL of thawed urine sediment samples were washed with 1× PBS one more time and concentrated to ≤100 µL of total sample volume.

2.3. DNA extraction

For the isolation of genomic DNA, up to 30 mg of homogenized tissue powder, all material of the deparaffinized tissue blocks, or total volume of prepared urine sediment samples were used. Samples were treated with proteinase K (Thermo Scientific™, Thermo Fisher Scientific, Vilnius, Lithuania) in 500 µL of lysis buffer for tissue (50 mM Tris-HCl pH 8.5, 1 mM EDTA, 0.5% Tween-20; all from Carl Roth) or for urine sediment samples (10 mM Tris-HCl pH 8.0, 1% SDS, from Carl Roth; 75 mM NaCl, from Sigma-Aldrich, St. Louis, MO, USA) for up to 18 hours at 55 °C, and DNA was extracted following the standard phenol-chloroform purification and ethanol precipitation.

The concentration and purity of the extracted DNA were evaluated using the NanoDrop™ 2000 spectrophotometer (Thermo Scientific™, Thermo Fisher Scientific, Wilmington, DE, USA). Additionally, DNA integrity of the samples selected for genome-wide methylation profiling was checked electrophoretically (Fig. 2.2A).

2.4. Bisulfite conversion

For targeted DNA methylation analysis, 400 ng of extracted DNA were modified with bisulfite manually or using EZ DNA Methylation™ Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions, except that the initial incubation of samples was performed at 42 °C for 15 min.

According to the manual protocol, genomic DNA was denatured with 3 M NaOH (Sigma-Aldrich) for 15 min at 37 °C followed by bisulfite modification with 2.3 M $Na_2S_2O_5$ (Sigma-Aldrich) and 10 mM hydroquinone (Sigma-Aldrich) for 15 h at 50 °C. After the incubation DNA was purified using Wizard® DNA Clean-up System (Promega, Madison, WI, USA) and desulfonated with 3 M NaOH for 15 min at 37 °C. Modified DNA was

precipitated with 70% ethanol, resuspended in 40 µL of sterile water and immediately used for the analyses or stored at -20 °C.

2.5. Genome-wide DNA methylation profiling

2.5.1. DNA methylation microarrays

Global DNA methylation profiling of 9 PCa and NPT pairs from the same patients was performed using two-color Human DNA Methylation 1×244K (HD) Microarrays (Gene Expression Omnibus, GEO, accession identifier GSE89243; https://www.ncbi.nlm.nih.gov/geo/) according to the manufacturer's protocol G4170-90012 v2.1 (Agilent Technologies, Santa Clara, CA, USA). Briefly, 5 µg of purified DNA in 1× PBS was sonicated into fragments of 150-900 bp in size using Bioruptor® system with a cooling element (Diagenode, Liege, Belgium) applying high intensity mode (Fig. 2.2A-C).

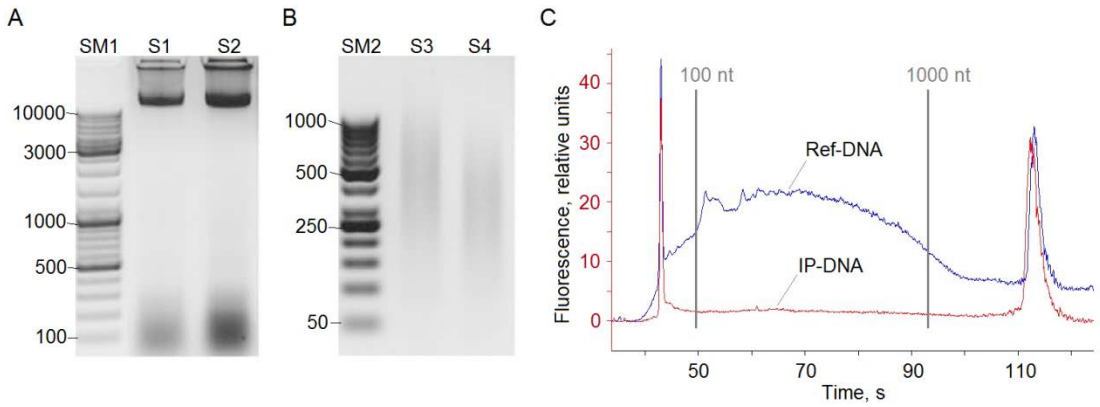


Figure 2.2. Genomic DNA analysis by means of electrophoresis. A – genomic DNA integrity analysis in 1.5% agarose gel after DNA purification; B – sonicated DNA analysis in 3% agarose gel; C – immunoprecipitated methylated DNA (IP-DNA) and reference DNA (Ref-DNA) analysis after purification using 2100 Bioanalyzer system with Agilent High Sensitivity DNA analysis kit (Agilent Technologies). Agarose gels were stained with ethidium bromide (Carl Roth). S1-4 – prostate cancer DNA samples, SM1 – DNA size marker GeneRuler DNA Ladder Mix, ready-to-use (#SM0333, Thermo Scientific™), SM2 – DNA size marker GeneRuler™ 50 bp DNA Ladder (#SM0373, Thermo Scientific™).

For the immunoprecipitation of the methylated DNA (IP-DNA), 80% of the sonicated sample was mixed with 50 µL of the prepared 5-methylcytosine (5-mC) monoclonal antibody 33D3 (Diagenode) and magnetic beads

(DynaBeads Pan Mouse IgG, Invitrogen™, Thermo Fisher Scientific, Carlsbad, CA, USA) mixture and 250 µL of 2× IP buffer, and gently mixed for 18 hours at speed 40 in Stuart® tube rotator SB3 (Cole-Parmer, Vernon Hills, IL, USA) at 4°C. IP-DNA and untreated reference DNA (Ref-DNA) were purified with phenol-chloroform (Carl Roth) using MaXtract High Density gel-filled tubes (Qiagen, Valencia, CA, USA).

After immunoprecipitation IP-DNA and Ref-DNA labeling was performed using SureTag DNA Labeling kit followed by purification according to manufacturer's protocol (Agilent Technologies). Sample yield and specific activity of the dyes were determined spectrophotometrically. Cy5- and Cy3-labeled IP-DNA and Ref-DNA samples (Fig. 2.3A) were combined in a single mixture per sample and hybridized onto Human DNA Methylation 1×244K microarrays, design ID 023795 (Agilent Technologies), for 40 hours at 67°C in a rotating hybridization oven (Agilent Technologies). After hybridization microarrays were washed immediately in a buffer system prepared in advance (Fig. 2.3B) and scanned with Agilent G4900DA SureScan microarray scanner (Agilent Technologies; Fig. 2.3C).

Feature Extraction software v10.7.3.1 (Agilent Technologies) was used for data extraction from microarray TIFF images (Fig. 2.3D). The data provided by the software were further subjected to quality control, pre-processing, and differential analysis.

2.5.2. Microarray data processing and analysis

All microarray datasets were normalized using the same procedure starting from the raw data. Saturated, non-uniform, and outlier probe signals were treated as compromised and removed from the analysis. Normalized log ration (Cy5/Cy3) representing IP-DNA/Ref-DNA was used for further calculations. Probe annotations, based on the microarray design identifier, were uploaded from SureDesign platform (https://earray.chem.agilent.com/suredesign) and updated using UCSC Genome Browser (https://genome.ucsc.edu) [245,246]

according to the human genome assembly version GRCh38. Probes undetected in ≥30% of all samples were filtered out followed by additional group comparison-specific filtering leaving only probes detected in 100% of samples in at least 1 of 2 groups to be compared.

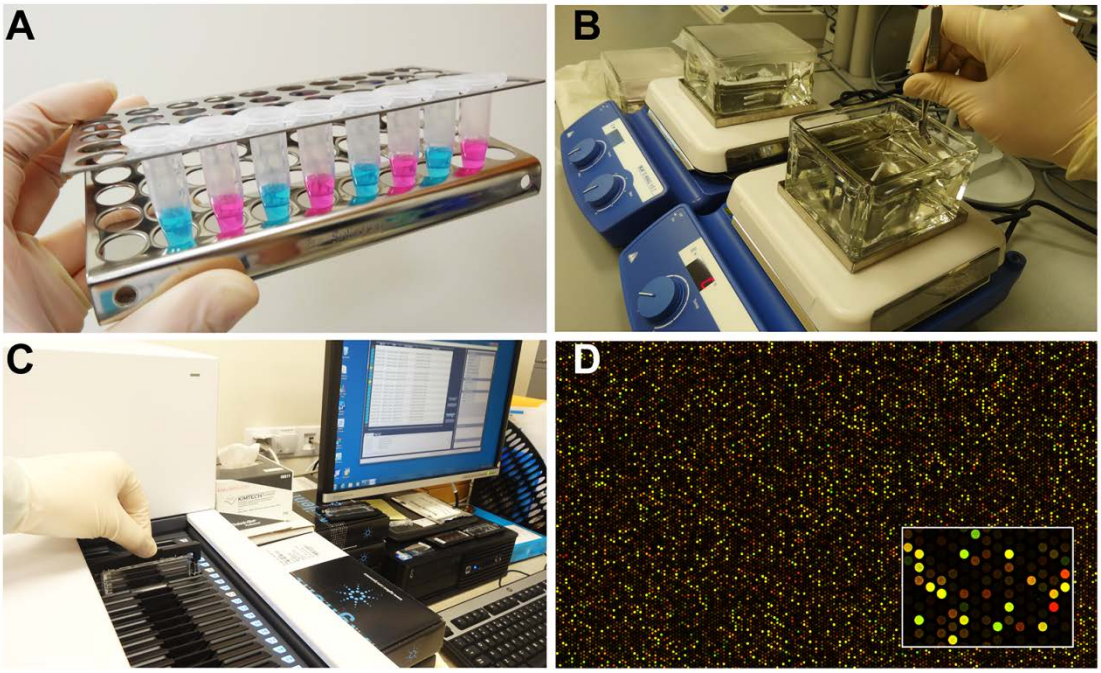


Figure 2.3. Several steps of DNA methylation analysis by means of microarrays. A – immunoprecipitated methylated DNA and reference DNA samples after labeling followed by purification; B – microarray washing in the buffer system after hybridization; C – microarray scanning with Agilent G4900DA SureScan scanner system (Agilent Technologies); D – a magnified fragment of a microarray TIFF image file (the diameter of a dot on a microarray is 30 μm).

The same microarray dataset was utilized for the analysis of miRNA host gene-associated loci. MiRBase release 21 (http://www.mirbase.org) was used for obtaining miRNA annotations which were attributed to microarray probes [247,248]. Randomly selected annotations (~1%) associated with miRNA host genes were checked manually. In this analysis, probes not associated with miRNAs were removed and the rest were filtered for being detected in >30% of the samples.

Fold change (FC) values were estimated and paired (where applicable) or unpaired t-tests were used. Initially, a stringent and a non-stringent filtering approaches were applied and two different P-values were defined, accordingly.

As the stringent filtering, i.e. absolute FC ≥ 1.5 and corrected P-value <0.0500, did not yield any significant probes, the non-stringent filtering was used in the analyses: methylation level according to a probe was classified as different comparing particular groups, if the absolute FC value was ≥1.2 or ≥1.5 for different group comparisons and non-adjusted P-value was <0.0500. Calculations and data visualization were performed with GeneSpring™ GX v13.1.1 software (Agilent Technologies).

2.5.3. Pathway analysis of differentially methylated genes

Gene Ontology (GO) [249,250] analysis was performed with The Biological Networks Gene Ontology tool v3.0.3 (BiNGO) [251], as a plug-in for the open source software platform Cytoscape v3.2.1 (provided by National Institute of General Medical Sciences, Bethesda, MD, USA) [252]. Only GO terms belonging to Biological Process category were included in the analysis. Official gene symbols (approved by HUGO Gene Nomenclature Committee) were used for input (http://www.genenames.org) [253]. GO annotations for the differentially methylated genes were retrieved using BiNGO and then tested for significance using hypergeometric test with Benjamini-Hochberg false discovery rate (FDR) correction (q-value <0.01). The obtained results were then visualized with Cytoscape as networks of enriched GO terms.

The gene set enrichment analysis (GSEA) for the identified differentially methylated genes between groups was performed using publicly available online GSEA tool and Molecular Signatures Database (MSigDB, v5.2; http://software.broadinstitute.org/gsea) [254], both maintained by Broad Institute (Cambridge, MA, USA). Hallmark genes sets (50 in total), which summarize and represent specific well-defined biological states or processes and display coherent expression, were utilized for GSEA [255]. FDR q-value with the cut-off <0.05 was used for multiple testing correction. GSEA data was visualized using Qlucore Omics Explorer v3.0, trial license (Qlucore, Lund, Sweden).

2.6. Targeted methylation analysis by methylation-specific PCR

Bisulfite-modified DNA served as a template for methylation-specific PCR (MSP) with primers specific for methylated or unmethylated DNA (Table S3). Primers were selected from publications or designed with Methyl Primer Express® Software v1.0 (Applied Biosystems™, Thermo Fisher Scientific, Carlsbad, CA, USA) and ordered from Metabion (Martinsried, Germany). For the genes selected from the microarray analysis, primers were designed to overlap with the location of the microarray probes of interest. The contents of the reaction mixture are provided in Table 2.1. Reaction conditions were optimized prior to the study and included 35-39 cycles with primer annealing step at 55-67°C for 45 s (Table S3).

Table 2.1. Composition of the methylation-specific PCR mixture per well.

Reaction component	Final concentration/ amount	
	Maxima® Hot Start Taq DNA Polymerase kit (Thermo Scientific™)	AmpliTaq Gold® DNA Polymerase kit (Applied Biosystems™)
10× PCR buffer	1×	1×
25 mM $MgCl_2$	2.5 mM	2.5 mM
16 mM dNTP mix	0.4 mM each	0.4 mM each
100 µM forward & reverse primers	1 µM each	0.5 µM each
Hot Start Taq DNA Polymerase	0.05 U/µL	0.05 U/µL
360 GC Enhancer	–	1 µL
Bisulfite-converted template DNA	10-20 ng	10-20 ng
Sterile H_2O	Up to 25 µL	Up to 25 µL

Bisulfite-modified leukocyte DNA from healthy donors served as a negative control for methylated DNA and CpG methyltransferase-treated (Thermo Scientific™, Thermo Fisher Scientific, Vilnius, Lithuania) bisulfite-modified leukocyte DNA served as a positive control. No-template controls (NTCs) were included in each MSP assay. Amplified products were analyzed in 3% agarose gel prepared using 1× TAE buffer (Thermo Scientific™, Thermo Fisher Scientific, Vilnius, Lithuania) and stained with ethidium bromide (Carl Roth; Fig. 2.4).

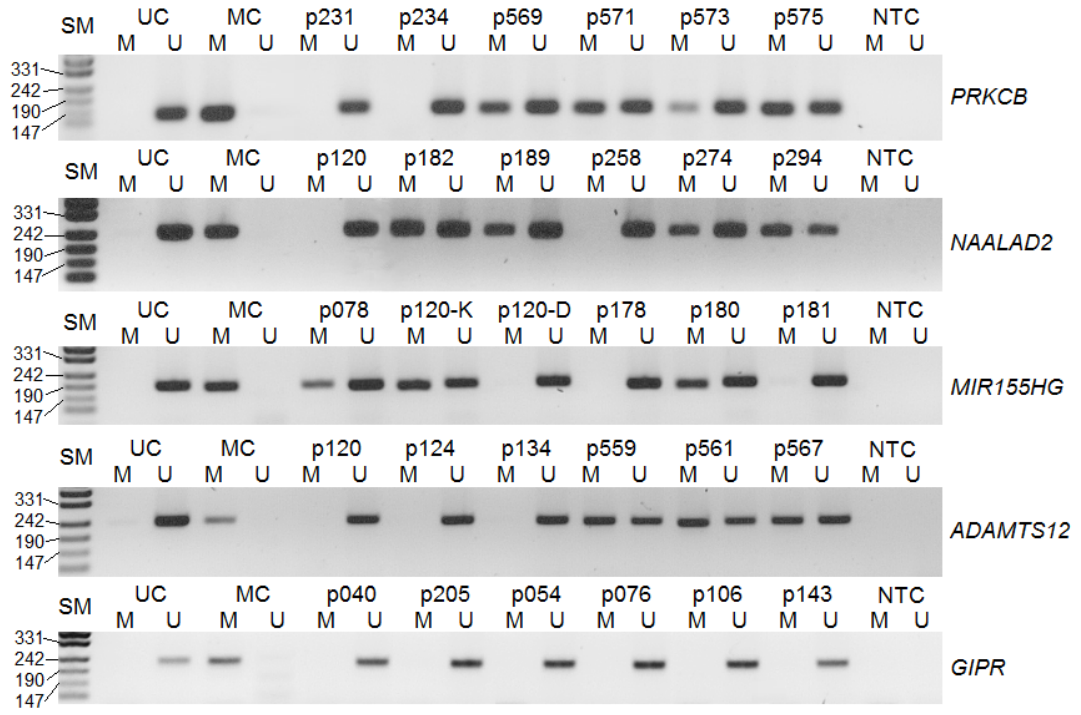


Figure 2.4. Examples of promoter methylation analysis of genes *PRKCB*, *NAALAD2*, *MIR155HG*, *ADAMTS12*, and *GIPR* by means of MSP in various prostate samples (p000). SM – DNA size marker pUC19 DNA/MspI (HpaII) (Thermo Scientific™), fragments lengths are provided in bp, UC/ MC – unmethylated/ methylated controls, NTC – no-template control, M/ U – amplification with primers specific for methylated/ unmethylated DNA.

2.7. Quantitative methylation-specific PCR

Quantitative MSP (QMSP) primers and hydrolysis probes for genes *PRKCB*, *CCDC181*, and *ADAMTS12* were designed using either Methyl Primer Express® Software v1.0 (Applied Biosystems™) or MethPrimer software v1.0 (http://www.urogene.org/methprimer/index.html) [256], while others (for *RARB*, *GSTP1*, *RASSF1*, and *ACTB*) were selected from publications (Table S4). Endogenous control gene (*ACTB*) was included in each assay to normalize for DNA input.

QMSP was performed in separate wells in triplicates for each set of primers. For the methylation analysis of genes *RARB*, *GSTP1*, and *RASSF1*, Maxima™ Probe qPCR Master Mix (2X), ROX Solution provided (Thermo Scientific™) was used, while for genes *PRKCB*, *CCDC181*, and *ADAMTS12* – TaqMan® Universal Master Mix II (Applied Biosystems™). The contents of reaction mixtures are provided in Table 2.2. All assays were carried out on

Mx3005P qPCR System (Agilent Technologies) under the following conditions: 95°C for 10 min followed by 50 cycles of 95°C for 15 s and 60°C for 1 min.

Table 2.2. Composition of the quantitative methylation-specific PCR mixture per well.

Reaction component	Final concentration/ amount	
	Maxima™ Probe qPCR Master Mix (2X), ROX Solution provided (Thermo Scientific™)	TaqMan® Universal Master Mix II no UNG (Applied Biosystems™)
2× Master Mix	1×	1×
10 µM forward & reverse primers	0.3 µM each	0.3 µM each
100 µM probe	50 nM	50 nM
5 µM ROX	30 nM	[included in Master Mix]
Bisulfite-converted template DNA	~10 ng	~10 ng
Sterile H_2O	Up to 20 µL	Up to 20 µL

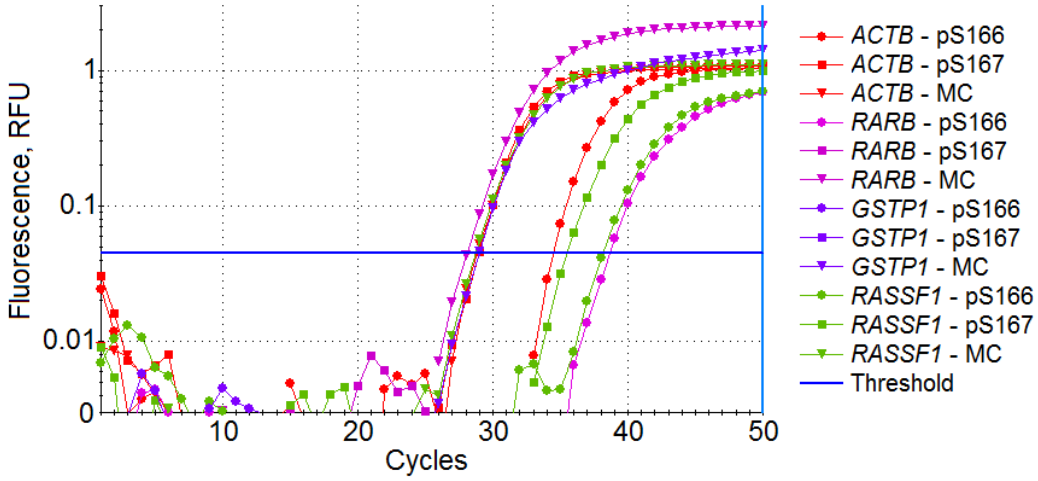


Figure 2.5. An example of quantitative methylation-specific PCR amplification curves. Methylation of the genes *RARB*, *GSTP1*, and *RASSF1* was analyzed using *ACTB* as an endogenous control in two urine samples (pS000). Baseline-subtracted fluorescence signal normalized to the passive reference dye ROX is provided in logarithmic scale in relative fluorescence units (RFU). MC – methylated control.

The results were generated using MxPro v4.0 software (Agilent Technologies) (Fig. 2.5). A run was considered valid when routinely included MCs gave a positive signal and there was no amplification in NTC wells. A sample was classified as valid if the cycle of quantification (Cq)-value of *ACTB* did not exceed 40 and amplification above threshold (automatically determined applying the amplification-based algorithm) was detected in all replicates for a particular gene. Methylation level of a particular gene was

estimated as a percentage of methylated reference DNA (PMR), which was calculated based on ΔΔCq algorithm using MC as a reference (the calibrator) and expressed in percentage [257].

2.8. RNA extraction

Total RNA (including miRNA fraction) was extracted using mirVana™ miRNA Isolation Kit (Ambion®, Thermo Fisher Scientific, Foster City, CA, USA) as described previously [243,244]. Briefly, ~30 mg of homogenized tissue powder were treated with 500 µL Lysis/ Binding Buffer and 50 µL of miRNA Homogenate Additive for 10 min in ice-water bath. Total RNA was extracted with 500 µL of acid-phenol:chloroform and purified using the supplied Filter Cartridges. One hundred µL of preheated (95°C) Elution Solution was used to recover purified RNA.

Concentration of total RNA was evaluated spectrophotometrically with NanoDrop 2000 (Thermo Scientific™). The RNA integrity number (RIN) was ≥7.0 (mean 8.3, range 7.0-9.7), as measured with 2100 Bioanalyzer system using RNA 6000 Nano kit and 2100 Expert software, version B.02.08.SI648 (Agilent Technologies; Fig. 2.6). Samples were stored at -80°C until further use.

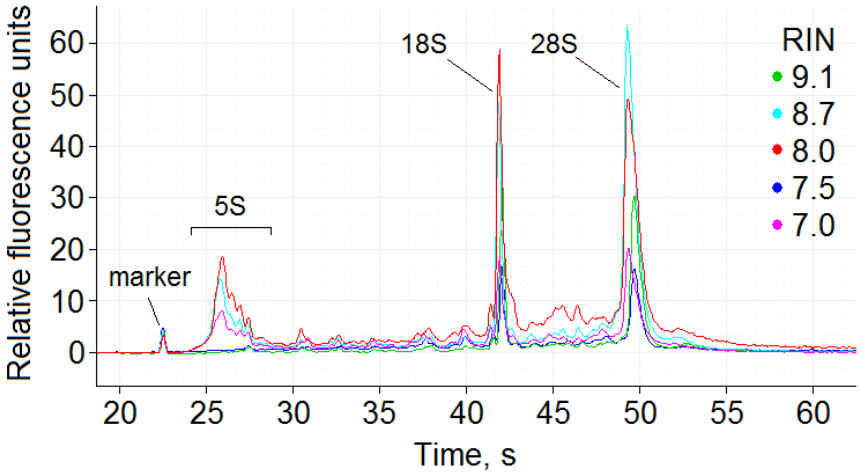


Figure 2.6. Evaluation of extracted total RNA integrity using 2100 Bioanalyzer (Agilent Technologies). Profiles of several samples with different RNA integrity numbers (RIN), automatically calculated by the software, are shown. Peaks of 5S, 18S, and 28S ribosomal RNA are marked.

2.9. Gene expression analysis by RT-qPCR

For cDNA synthesis, 250 ng of RNA was reverse transcribed (RT) using High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor according to the manufacturer‘s instructions (Applied Biosystems™).

Expression levels of 5 protein-coding genes *PRKCB*, *CCDC181*, *ADAMTS12*, *NAALAD2*, *ZMIZ1*, 3 miRNA target genes *DNMT1*, *KDM1A*, *KDM5B*, and endogenous control *HPRT1* were evaluated by means of quantitative PCR (RT-qPCR) using TaqMan® Gene Expression Assays (Hs00176998_m1, Hs00221792_m1, Hs00229594_m1, Hs01119919_m1, Hs00277476_m1, Hs00945900_g1, Hs01002741_m1, Hs00981910_m1, and Hs02800695_m1, respectively; Applied Biosystems™). The reaction mix (20 µL) consisted of 10 µL of TaqMan® Universal Master Mix II no UNG (Applied Biosystems™), 0.6 µL of TaqMan® assay, and 2 µL of RT reaction product (equivalent to 25 µL of total RNA). RT-qPCR was performed on the Mx3005P qPCR System (Agilent Technologies) in triplicates per gene. Thermal cycling conditions consisted of 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Multiple NTCs were included in each RT-qPCR run. Data preprocessing was performed separately for protein-coding and miRNA target genes with GenEx v6.0.1 software (MultiD Analyses AB, Göteburg, Sweden). Relative gene expression values in a linear scale were used for the analysis.

2.10. MiRNA expression analysis

Quantitative miRNA expression data was extracted from our previous study based on the samples from the same PCa cohort and described in detail elsewhere [244]. For the reanalysis, raw Cq-values of miRNA expression from the TaqMan® Low Density Array-based (TLDA) screening using Human MicroRNA A+B Cards Set v3.0 (Applied Biosystems™) were obtained for a set of 41 PCa and 12 NPT samples (Table S2). Data preprocessing was performed using GenEx v6.0.1 software (MultiD Analyses AB). Considering the epigenetic silencing of miRNA expression, missing Cq-values (due to the

amplification below the threshold or none at all) were equated to 40. Expression levels of miR-155-5p, miR-152-3p, miR-137, and miR-31-5p were normalized to RNU44 and RNU48 as endogenous controls, as suggested by the manufacturer, and converted to relative values in a linear scale. Expression of miR-155-3p was undetected in all of the selected samples, therefore, it was omitted from the analysis.

2.11. The Cancer Genome Atlas Dataset of prostate cancer

Publicly available data from TCGA, a collaboration between the National Cancer Institute (NCI, USA) and National Human Genome Research Institute (NHGRI, USA), were used to verify the significant findings. Clinical annotation of the samples was obtained from the marker TCGA PRAD publication [4]. Global DNA methylation profiling data using Illumina Infinium HumanMethylation450K (HM450) platform, RNA and miRNA expression by RNA-seq and miRNA-seq, respectively, were utilized in this study. Level 3 data were obtained from the cBioPortal (http://www.cbioportal.org) [258] and methHC (http://methhc.mbc.nctu.edu.tw) [259] data analyses portals in July 2016. Samples with significant degradation levels, as described in [4], were excluded from the analysis yielding 333 tumors and 19 NPT in total.

2.12. Statistical analysis

Statistical analysis was performed using STATISTICA™ v8.0 (StatSoft, Tulsa, OK, USA) and GraphPad Prism v5.03 software (GraphPad Software, La Jolla, CA, USA). The two-sided Fisher's exact test or Kruskal-Wallis test were used for comparison of categorical variables, while continuous variables were compared by Student's t-test or by Mann-Whitney U test for normally or non-normally distributed variables, respectively. Pearson (R_P) and/ or Spearman's rank (R_S) correlation coefficients were calculated to test the associations between two continuous variables. TCGA dataset of PRAD was used for comparison. For the analysis of RT-qPCR and TCGA data, appropriate parametric tests were applied. For BCR-free survival analysis, Kaplan-Meier

curves were compared with log-rank test and Cox proportional hazards modeling was performed. Hazard ratios (HR) with 95% confidence intervals (CI) were estimated. To avoid overfitting of models, only up to 4 covariates were included in multivariate analysis. Differences and associations were considered statistically significant at $P < 0.0500$.

Microarray data and pathway enrichment analyses were performed as described above, using specialized software.

3. RESULTS

The present study of DNA methylation biomarkers of PCa consisted of two stages of experimental work. During the first stage, several TSGs, previously reported to be aberrantly methylated in various malignancies, were selected for promoter methylation in tissue samples of PCa patients followed by evaluation of the most potential biomarkers for noninvasive screening in urine. Although the obtained results were promising, the relatively low methylation frequencies and specificity for PCa in Lithuanian cohort encouraged to search for new biomarkers with better performance. Therefore, the second stage of the study was oriented towards identification and validation of novel DNA methylation biomarkers of PCa, evaluating both their diagnostic and prognostic potential. All the genes analyzed in the present study are listed in Table S5.

3.1. DNA methylation analysis of selected tumor suppressor genes in prostate cancer

The first stage of the study was focused on the qualitative methylation analysis of 7 promoters of TSGs selected from previous publications (*RARB*, *GSTP1*, *RASSF1*, *MGMT*, *DAPK1*, $p16^{INK4a}$/*CDKN2A*, and $p14^{ARF}$/*CDKN2A*) in the FFPE set of PCa and tumor-adjacent NPT samples. Snap-frozen BPH tissue samples were included in the analysis as an additional control group. The 3 most frequently methylated genes – *RARB*, *RASSF1*, and *GSTP1* – were quantitatively analyzed in urine of PCa and BPH patients in order to assess their potential as noninvasive biomarkers (Fig. 3.1).

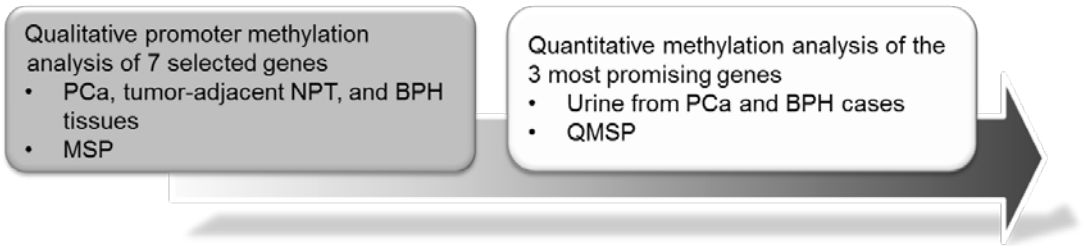


Figure 3.1. The analysis workflow of the selected tumor suppressor genes as DNA methylation biomarkers of prostate cancer (PCa). NPT – noncancerous prostate tissues, BPH – benign prostatic hyperplasia, MSP – methylation-specific PCR, QMSP – quantitative methylation-specific PCR.

3.1.1. DNA methylation of tumor suppressor genes in prostate tissues

First, DNA methylation of the selected TSGs was analyzed in PCa tissue samples (Fig. 3.2). The most frequent aberrant methylation was identified in promoter regions of *RARB* (97 of 149; 65.1%), *GSTP1* (84 of 149; 56.4%), and *RASSF1* (80 of 148; 54.1%), and less frequently occurred in promoters of *MGMT* (38 of 149; 25.5%), *DAPK1* (16 of 149; 10.7%), *p16* (12 of 149; 8.1%), and *p14* (10 of 149; 6.7%). In overall, at least one gene of the gene panel was methylated in 90.6% of PCa cases (135 of 149).

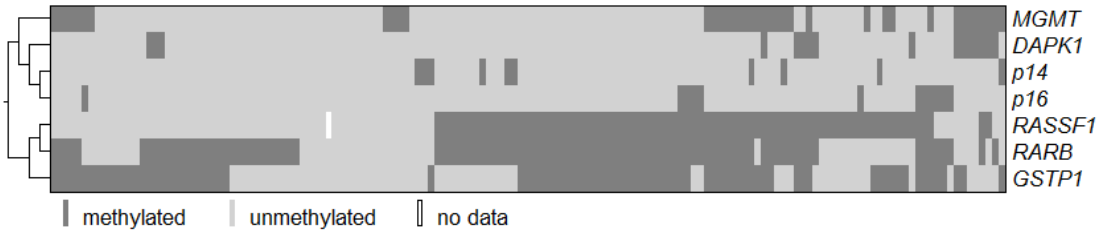


Figure 3.2. Promoter methylation profile of *RARB*, *GSTP1*, *RASSF1*, *MGMT*, *DAPK1*, *p16*, and *p14* genes in prostate cancer samples.

Methylation of the 3 most frequently methylated genes in PCa tissues – *RARB*, *GSTP1*, and *RASSF1* – was further analyzed in NPT (N = 37) and BPH (N = 17) samples. Methylation of *RARB* and *GSTP1* rarely occurred ($P < 0.0500$) in NPT and BPH as compared to PCa. Methylation of *RASSF1* was more prevalent ($P = 0.0283$) in PCa than in NTP, while the difference to BPH was not statistically significant ($P > 0.0500$; Fig. 3.3). Methylation of at least one of the three genes was detected in 85.2% (127 of 149) of PCa tissues, while the specificity was 62.2% (23 of 37) and 47.1% (8 of 17) according to NPT and BPH, respectively.

In addition, quantitative comparison of methylation levels in randomly selected 15 PCa and 15 BPH tissues showed higher average PMR values (all $P < 0.0500$) for all 3 genes in PCa cases (Fig. 3.4A).

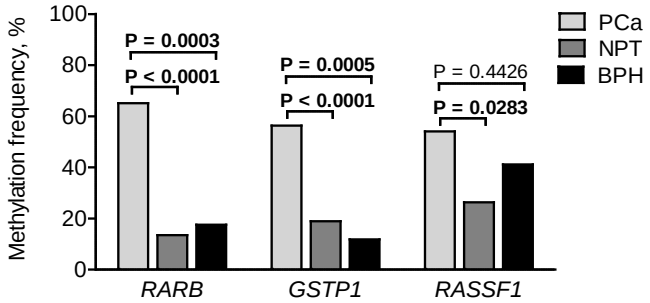
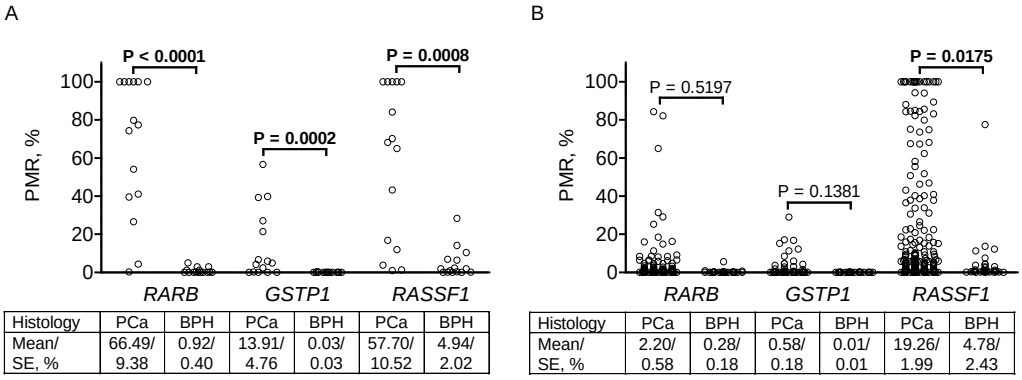


Figure 3.3. Methylation frequencies of *RARB*, *GSTP1*, and *RASSF1* according to prostate tissue histology. PCa – prostate tumors, NPT – noncancerous prostate tissues, BPH – benign prostatic hyperplasia samples. Significant P-values are in bold.



Histology	PCa	BPH	PCa	BPH	PCa	BPH
Mean/ SE, %	66.49/ 9.38	0.92/ 0.40	13.91/ 4.76	0.03/ 0.03	57.70/ 10.52	4.94/ 2.02

Histology	PCa	BPH	PCa	BPH	PCa	BPH
Mean/ SE, %	2.20/ 0.58	0.28/ 0.18	0.58/ 0.18	0.01/ 0.01	19.26/ 1.99	4.78/ 2.43

Figure 3.4. Distribution of methylation intensity values (PMR) of genes *RARB*, *GSTP1*, and *RASSF1* in samples of prostate cancer (PCa) and benign prostatic hyperplasia (BPH) patients. A – PMR values in tissues of PCa and BPH patients; B – in urine samples of PCa and BPH patients. Mean values with standard errors (SE) are given below for each group. Statistically significant P-values are in bold.

3.1.2. *RARB, GSTP1,* and *RASSF1* methylation analysis in urine

QMSP was used for the detection of DNA methylation of *RARB*, *GSTP1*, and *RASSF1* in urine samples of 253 PCa and 32 BPH patients (Table S2). In urine from PCa patients, the observed PMR values were higher than in BPH cases (Fig. 3.4B). However, only the difference of *RASSF1* promoter methylation intensity was statistically significant (P = 0.0175).

For each gene, the mean PMR value in BPH was used as a cut-off for the qualitative assessment of DNA methylation status in urine sediments. Methylation of *RASSF1* was found in 44.7% (113 of 253) of PCa samples, but was less frequent for the other two genes (29.2%; 74 of 253 and 11.1%; 28 of 253 for *RARB* and *GSTP1*, respectively). Aberrant methylation of at least one of the 3 genes was detected in 152 of 253 (60.1%) urine samples of PCa

patients. The combined specificity of this assay was 68.8% (22 of 32), while it reached 81.3% (26 of 32), 96.9% (31 of 32), and 84.4% (27 of 32) for *RARB*, *GSTP1*, and *RASSF1*, respectively.

3.1.3. Tumor suppressor gene methylation according to clinical-pathological and molecular characteristics

Methylation frequencies in PCa tissues were compared according to clinical-pathological characteristics and *TMPRSS2-ERG* fusion status (Table S2). Preoperative PSA level was significantly higher in cases with *RASSF1* methylation (P = 0.0098), while *GSTP1* methylation was more frequent in PCa of pT3 than ≤pT2 (P = 0.0190). Higher methylation frequency of *GSTP1* was also observed in tumors with Gleason score ≥7 as compared to 6 (P = 0.0072; Table S6), whereas the analysis according to ISUP grade groups revealed that the largest difference of methylation frequencies was between groups I and II (43.2% and 70.0%, respectively, P = 0.0056; data not shown).

In urine, methylation of *GSTP1* was more common (P = 0.0073) and more intense (P = 0.0059) in cases with Gleason score ≥7 in comparison to Gleason score 6 (Table S6). Similarly to the analysis in tumors, largest differences of both the frequency and the level of *GSTP1* methylation were observed comparing ISUP grade groups I and II (P = 0.0036 and P = 0.0042, respectively; data not shown). There were also inter-correlations among the PMR values of the genes: PMR of *GSTP1* correlated both with *RARB* and *RASSF1* (R_S = 0.17, P = 0.0083, and R_S = 0.18, P = 0.0034, respectively; data not shown).

No significant associations were observed between gene methylation and *TMPRSS2-ERG* fusion status or patients' age. Some other associations are presented in Table S6.

3.1.4. Tumor suppressor gene methylation and BCR-free survival

More frequent methylation of *RASSF1* was identified in PCa tissues of BCR-positive cases in comparison to BCR-negative cases (Fig. 3.5). In

Kaplan-Meier analysis, the advanced pT and Gleason score, and *RASSF1* methylation in tumor were significant predictors of BCR (Fig. 3.6A-C). Specifically, *RASSF1* methylation correlated with the higher rate of BCR in Gleason score 6 but not in Gleason score ≥7 tumors (Fig. 3.6D). In a univariate Cox proportional hazard models (Table S7), the presence of *RASSF1* and *DAPK1* promoter methylation in PCa significantly correlated with shorter BCR-free survival.

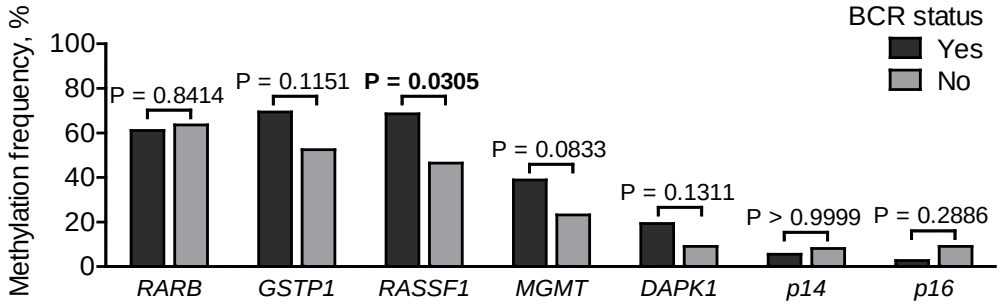


Figure 3.5. Promoter methylation frequencies of tumor suppressor genes in prostate tumors according to biochemical disease recurrence (BCR) status (last updated in July 2013). Significant P-values are in bold.

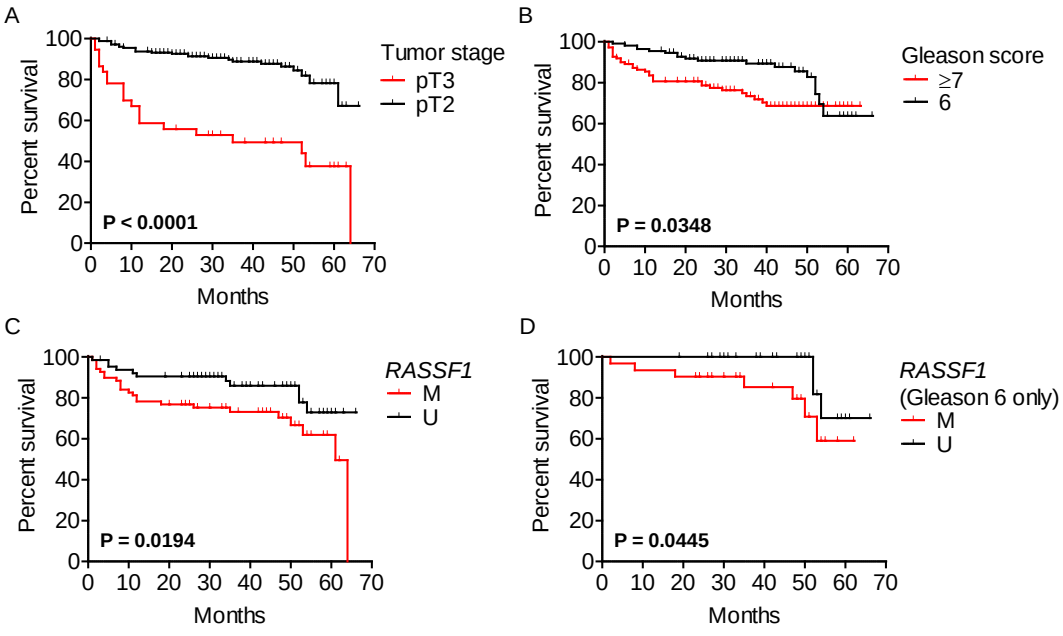


Figure 3.6. Kaplan-Meier curves according to tumor stage pT (A), Gleason score (B), methylation status of *RASSF1* in tumor tissues (C), and methylation status of *RASSF1* in Gleason score 6 tumors only (D). M/ U – methylated/ unmethylated promoter status. Significant P-values are in bold.

In multivariate analysis, pT showed significant prognostic value, while none of the methylation biomarkers was an independent predictor for BCR. However, in patients diagnosed with Gleason score 6 tumor, multivariate analysis using stepwise backward entering of covariates revealed the combination of pT and *RASSF1* methylation in PCa tissue as significant predictors of BCR-free survival (model's P = 0.0007). Similarly, *RASSF1* methylation in urine samples was predictive for time to BCR of patients with Gleason 6 tumor, when included in multivariate model together with pT (Table S7).

3.2. Identification of novel DNA methylation biomarkers of prostate cancer

Although the BCR-free survival analysis of *RASSF1* promoter methylation yielded promising results, the relatively low methylation frequencies of the analyzed genes in Lithuanian cohort prompted the search for new biomarkers with superior performance. A small set of well-characterized PCa cases were selected for the initial screening using DNA methylation microarrays. Based on comparisons by tissue histology and BCR status, 15 genes – 10 protein-coding and 5 miRNA host genes – with differences in methylation levels were selected for validation in an expanded set of samples. Expression analysis of the 5 most promising protein-coding genes and 4 miRNAs was performed by means of RT-qPCR and associated with promoter methylation status of the respective genes. For protein-coding genes, promoter methylation of 3 selected genes was also analyzed quantitatively in a subset of urine samples. For miRNA host genes, *in silico* target analysis of the 3 respective miRNAs was performed and 3 target genes (each gene being regulated by 2 of the selected miRNAs) were chosen for the expression analysis. The PRAD dataset of TCGA [4] was utilized as an independent validation cohort of the findings. The workflow is summarized in Fig. 3.7.

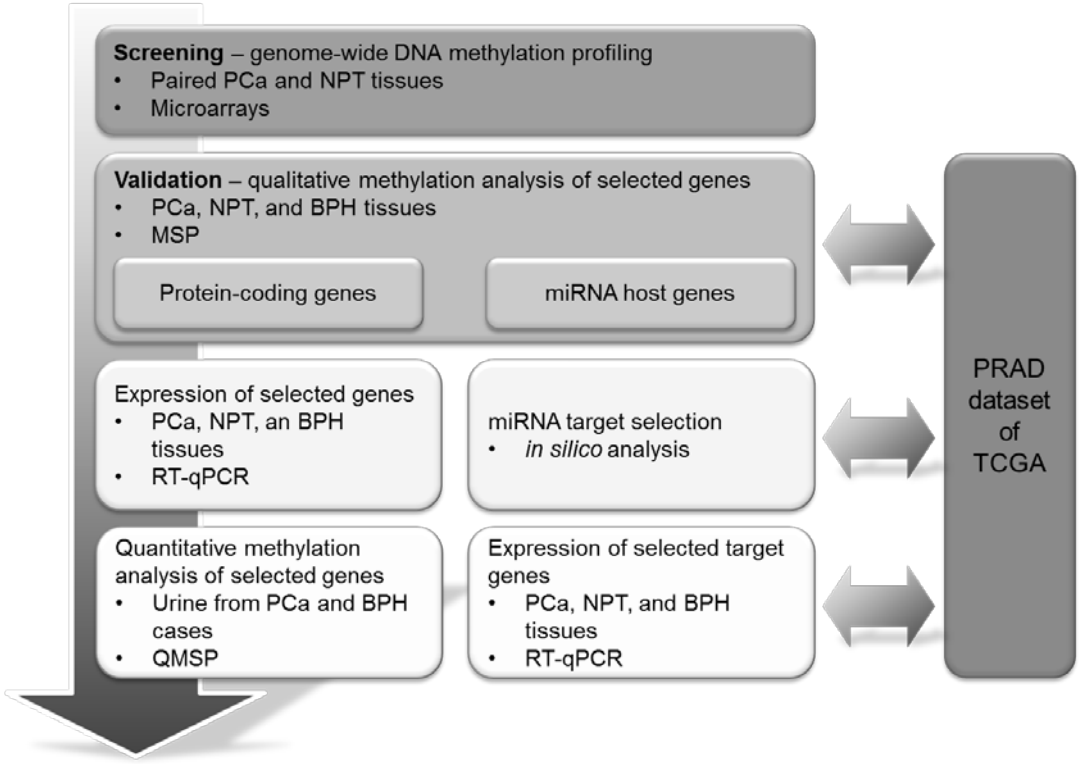


Figure 3.7. Scheme of the workflow for the search of novel DNA methylation biomarkers of prostate cancer. PCa – prostate tumors, NPT – noncancerous prostate tissues, BPH – benign prostatic hyperplasia, miRNA – microRNA, TCGA – The Cancer Genome Atlas project, PRAD – the prostate cancer cohort of TCGA, MSP – methylation-specific PCR, QMSP – quantitative methylation-specific PCR, RT-qPCR – reverse-transcription quantitative PCR.

3.2.1. Genome-wide DNA methylation profiling

To elucidate the extent of epigenetic changes in prostate tumors, global DNA methylation profile was analyzed of 9 pairs of PCa and NPT samples (Fig. 3.8A-C). Significant methylation differences (FC ≥ 1.5, P < 0.0500) were associated with 1143 genes in PCa as compared to NPT samples, of which 199 (17.4%) genes were hypermethylated and 954 (83.5%) genes were hypomethylated, including 10 (0.9%) genes with concurrent changes (Fig. 3.8A). In intragenic regions, the number of hypomethylated genes was much higher than the number of hypermethylated genes (87.3 and 13.4%, respectively). The same difference, only to a lesser extent, was also observed in promoter regions (59.9% and 40.1%, respectively; Fig. 3.9A). The list of the top 50 genes with the most significant differences between PCa and NPT samples is provided in Table S8.

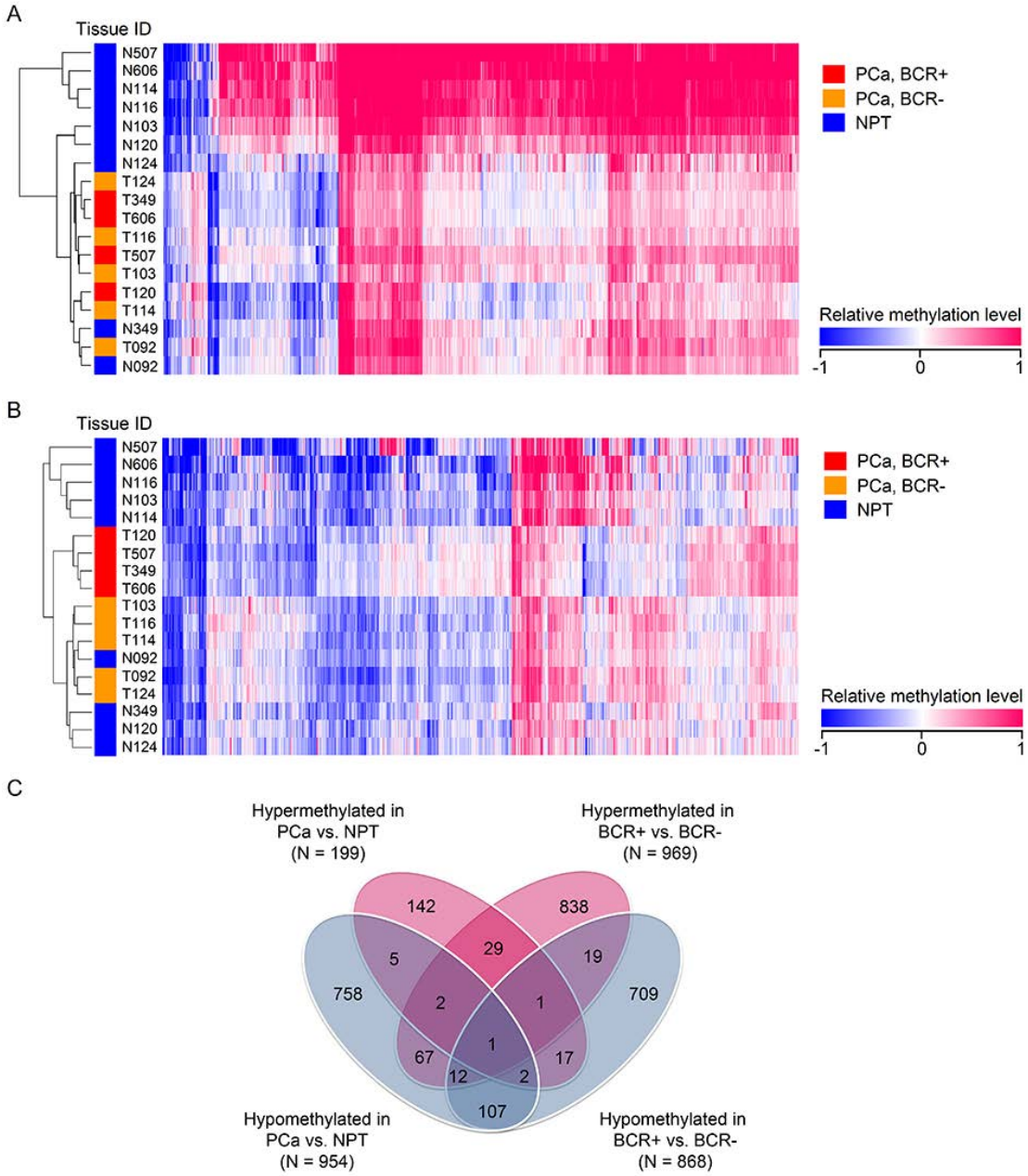


Figure 3.8. DNA methylation profile of genes with significant differences between analyzed groups. A – profile of significant methylation differences comparing prostate tumors (PCa) and noncancerous prostate tissues (NPT); B – profile of significant methylation differences comparing cases according to biochemical disease recurrence positive/ negative (BCR+/-) status; C – Venn diagram of genes with significantly different methylation levels according to tissue histology and disease progression.

Smaller-scale methylation differences were observed comparing BCR-positive and BCR-negative PCa cases and, therefore, the cut-off FC value was set to ≥1.2 for this analysis (Fig. 3.8B). Of 1804 genes with significant methylation differences, 969 (53.7%) genes were hypermethylated and 868

(48.1%) genes were hypomethylated, including 33 (1.8%) overlapping genes. In promoter regions, increase of methylation level was more commonly detected than reduction (53.2% and 47.6%, respectively, with 0.9% overlap) in BCR-positive cases as compared to BCR-negative cases, whereas the reverse tendency was observed in intragenic loci (44.8% and 56.8%, respectively; Fig. 3.9B). The list of the top 50 genes with the most significant differences between BCR-positive and BCR-negative cases is provided in Table S9.

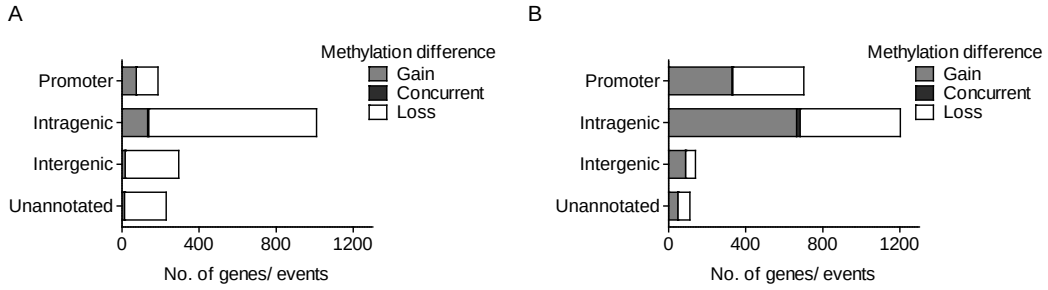


Figure 3.9. DNA methylation differences according to various genomic regions. A – between prostate tumors and noncancerous tissues, B – between biochemical disease recurrence (BCR)-positive and BCR-negative cases.

Hypermethylation of 33 overlapping genes was detected comparing both PCa *vs.* NPT and BCR-positive *vs.* BCR-negative tumor samples, while 122 genes were hypomethylated in both comparisons (Fig. 3.8C). Interestingly, some genes also showed both gain and loss of methylation according to different probes (Fig. 3.8C).

In both group comparisons, differences of methylation levels were also observed in intergenic regions. Hypomethylation events were more frequent than hypermethylation in PCa as compared to NPT samples (94.2% and 5.8%, respectively), while the opposite tendency was observed comparing BCR-positive and BCR-negative cases (37.1% and 62.9%, respectively). Variation of methylation levels was also detected in unannotated loci (Fig. 3.9A and B).

According to the DNA methylation microarray design, 7.1% of biological probes (16853 of 237203) were annotated as associated with miRNA host genes, representing 88.6% (1248 of 1409) of miRNA genes residing in host genes in total (Fig. 3.10A). According to the NCBI Gene database

(www.ncbi.nlm.nih.gov/gene), 93.8% (917 of 978) of the host genes covered by the microarray were annotated as protein-coding genes, while the rest consisted of 4.1% (40 of 978) ncRNA genes and 2.0% (20 of 978) of uncertain type. Regarding the location of miRNA genes, the majority of them were found to be located within introns of the protein-coding host genes (80.3%, 736 of 917).

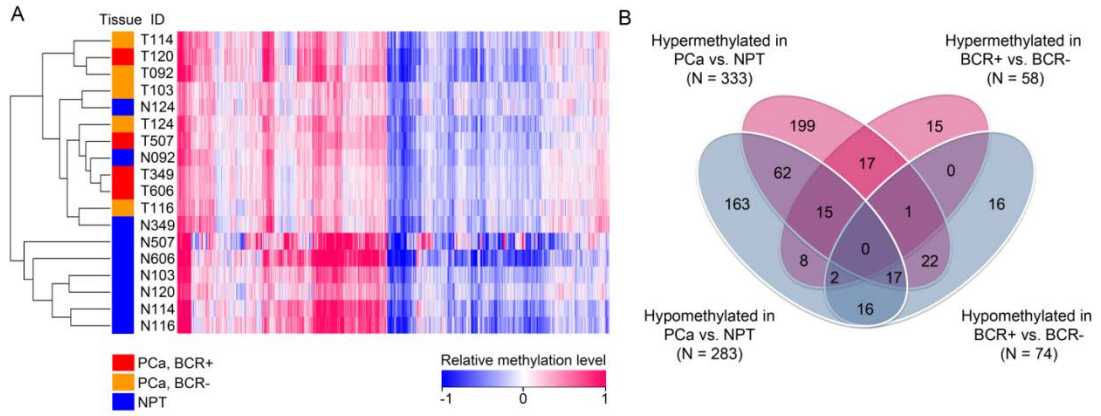


Figure 3.10. DNA methylation differences in miRNA host genes obtained from microarray data. A – DNA methylation profile according to all miRNA host gene-associated probes; B – Venn diagram of genes with significantly different methylation levels according to tissue histology and disease progression. PCa – prostate cancer, NPT – noncancerous prostate tissue, BCR+/- – biochemical disease recurrence status (positive/ negative).

Comparing PCa to NPT samples, significant methylation differences (FC ≥ 1.2; P < 0.0500) were identified in 522 miRNA host genes of which 333 were hypermethylated and 283 were hypomethylated, including 94 overlapping genes. In BCR-positive as compared to BCR-negative PCa cases, methylation differences were less common and significantly differed according to 129 host genes (FC ≥ 1.2, P < 0.0500), consisting of 58 hypermethylated and 74 hypomethylated genes with 3 gene overlap (Fig. 3.10B). Lists of top 50 genes with significant methylation differences in group comparisons are provided in Tables S10 and S11.

3.2.2. Functional term enrichment analysis of differentially methylated genes

In order to gain a better insight of the cellular processes that might be altered due to the aberrant methylation of the genes, the Biological Process

category of GO terms was analyzed for all the differentially methylated genes identified in group comparisons.

Analyzing the genes that showed promoter or intragenic methylation changes in PCa as compared to NPT samples (FC ≥ 1.5), enrichment of the gene groups related to cell differentiation, cell fate commitment, regulation of gene expression and transcription was identified (Fig. S1). Among the genes with methylation differences between BCR-positive and BCR-negative cases (FC ≥ 1.2), GO terms like cell differentiation, cell death, cellular response to hormone stimulus, regulation of gene specific transcription were significantly overrepresented (Fig. S2).

Analysis of the miRNA host genes showed significant over-representation of rather general GO terms (e.g. negative regulation of biological or cellular process, cellular component morphogenesis). However, some more specific GO terms, such as epithelial-mesenchymal transition (EMT), maintenance of cell polarity, and histone H3 or H4 deacetylation, were enriched among the miRNA host genes differentially methylated between BCR-positive and BCR-negative cases (Fig. S3).

To further elucidate the molecular mechanisms that occur during prostate carcinogenesis and disease progression, GSEA was performed. Gene sets involved in cell cycle regulation, estrogen response, and apical junction were among the most significantly enriched hallmark pathways in PCa as compared to NPT samples (Fig. 3.11A). The increase of methylation levels was the most significant among the genes downregulated in response to ultraviolet (UV) exposure, involved in EMT or TNFα signaling via NF-κB, while decreased methylation was commonly observed in genes associated with mitotic spindle or estrogen response. Similar gene sets (e.g. EMT, response to UV) were enriched for hypermethylated genes between BCR-positive and BCR-negative cases, while response to androgens or estrogen and hypoxia-related gene sets were enriched for the decrease of methylation (Fig. 3.11A).

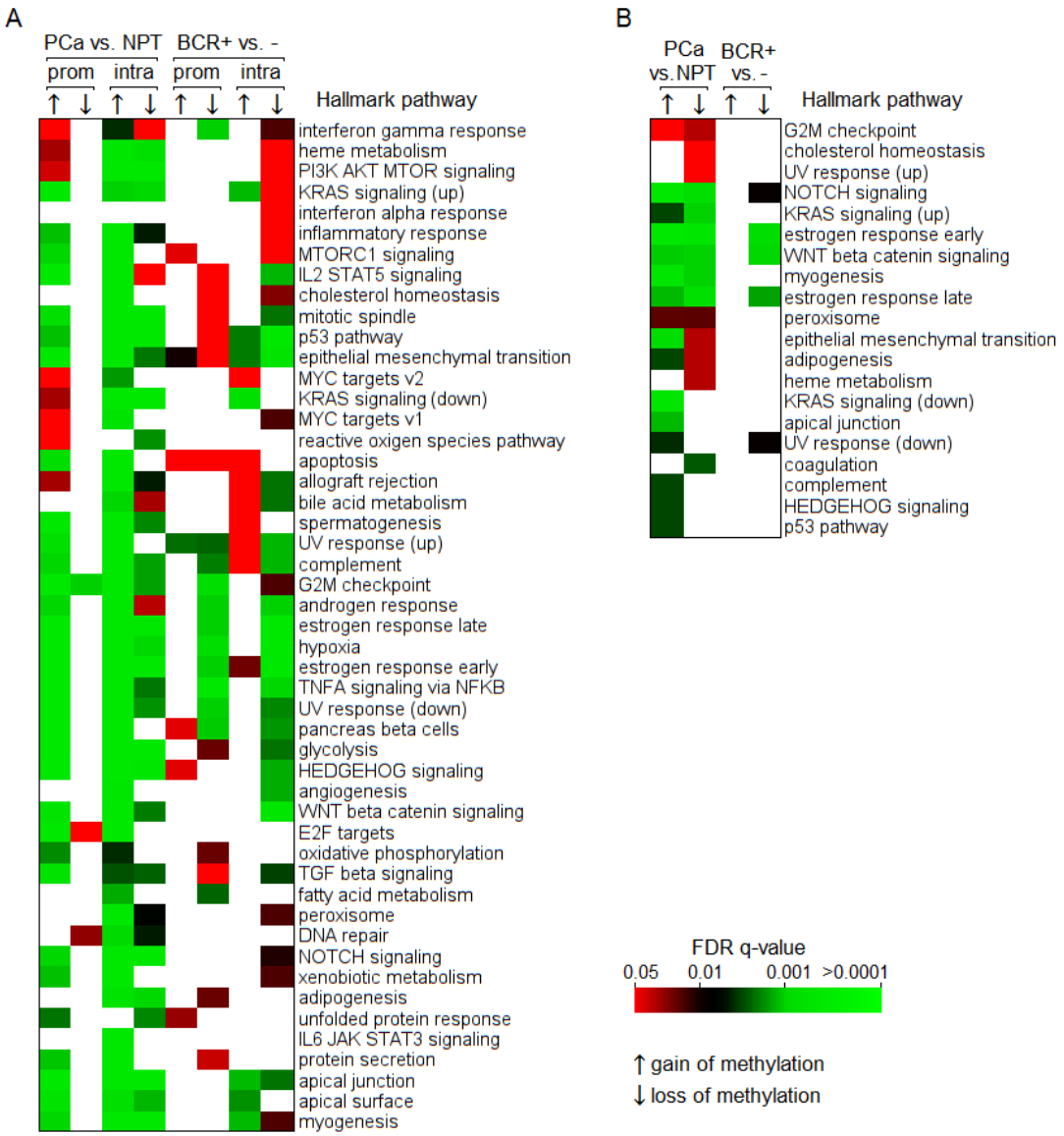


Figure 3.11. Gene set enrichment analysis of differentially methylated genes identified in genome-wide methylation profiling. A – functional gene sets for all identified differentially methylated genes; B – functional gene sets for differentially methylated miRNA host genes. Genes with significant methylation differences with fold change values ≥1.2 were included. The collection of Hallmark gene sets (pathways) as defined in MSigDB (http://software.broadinstitute.org/gsea/) were selected for the enrichment analysis. The color intensities indicate the level of false discovery rate (FDR) adjusted P-values (q-values). PCa – prostate cancer, NPT – noncancerous prostate tissue, BCR+/ - – biochemical disease recurrence status (positive/ negative), prom – promoter regions, intra – intragenic regions.

In the analysis of the differentially methylated miRNA host genes, estrogen response and WNT signaling through accumulation of β-catenin pathways were significantly enriched among genes with decreased methylation levels analyzing both PCa *vs.* NPT and BCR-positive *vs.* BCR-negative

groups. However, no pathways were significantly enriched for increased methylation levels comparing samples according to BCR status (Fig. 3.11B).

In general, GO and GSEA analysis revealed that biological and molecular processes occurring during tumor development and progression were commonly enriched among the sets of genes found to be differentially methylated comparing groups of samples from PCa patients.

3.2.3. Promoter methylation analysis of selected protein-coding genes

For further validation steps, 9 genes – *PRKCB*, *CCDC181*, *ADAMTS12*, *ZMIZ1*, *FILIP1L*, *NAALAD2*, *KCTD8*, *EPAS1*, and *NEK9* – were selected based on differences according to prostate tissue histology and BCR status observed in genome-wide DNA methylation data (Fig. 3.12A and B), while *CD44* was mainly included for comparison with regard to the previous reports [171,184], making a total of 10 genes.

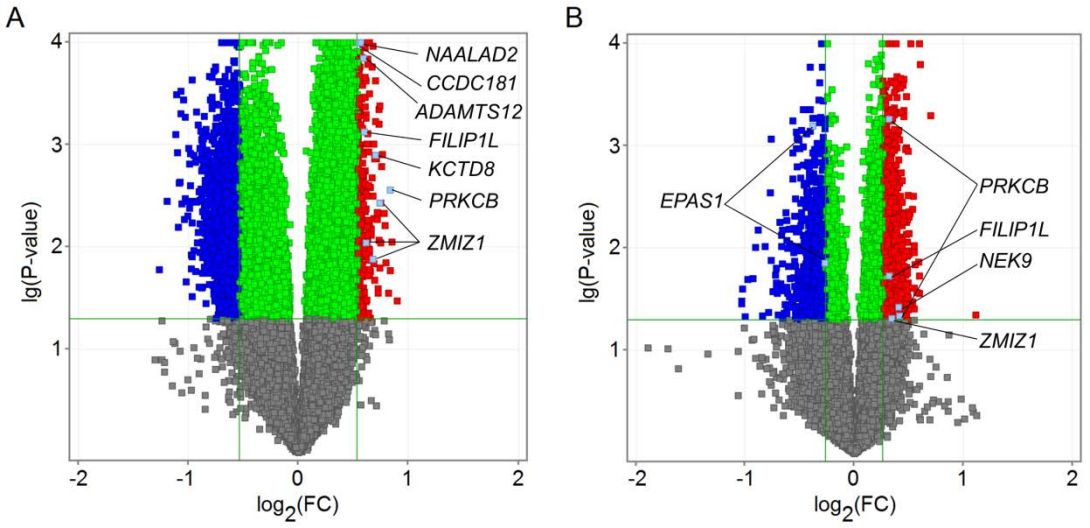


Figure 3.12. Volcano plots of DNA methylation profiling in tissues of prostate cancer (PCa) patients. A – methylation differences between PCa and noncancerous tissues; B – methylation differences in tumors of biochemical disease recurrence (BCR)-positive and BCR-negative cases. All probes are depicted as squares colored according to cut-off fold change (FC) and P-values. Due to different ranges of significant methylation differences, cut-off FC values of ≥1.5 and ≥1.2 were used for the analysis according to tissue histology and BCR status, respectively. Labels indicate genes selected for further validation. All significant probes per gene are highlighted.

Significantly higher methylation frequencies of genes *PRKCB*, *CCDC181*, *ADAMTS12*, *ZMIZ1*, *FILIP1L*, *NAALAD2*, *KCTD8*, and *CD44* were identified in PCa samples, ranging from 22.5% to 91.5%, as compared to NPT (range 0-37.1%) or BPH (0% for all genes; all P < 0.0500; Fig. 3.13A).

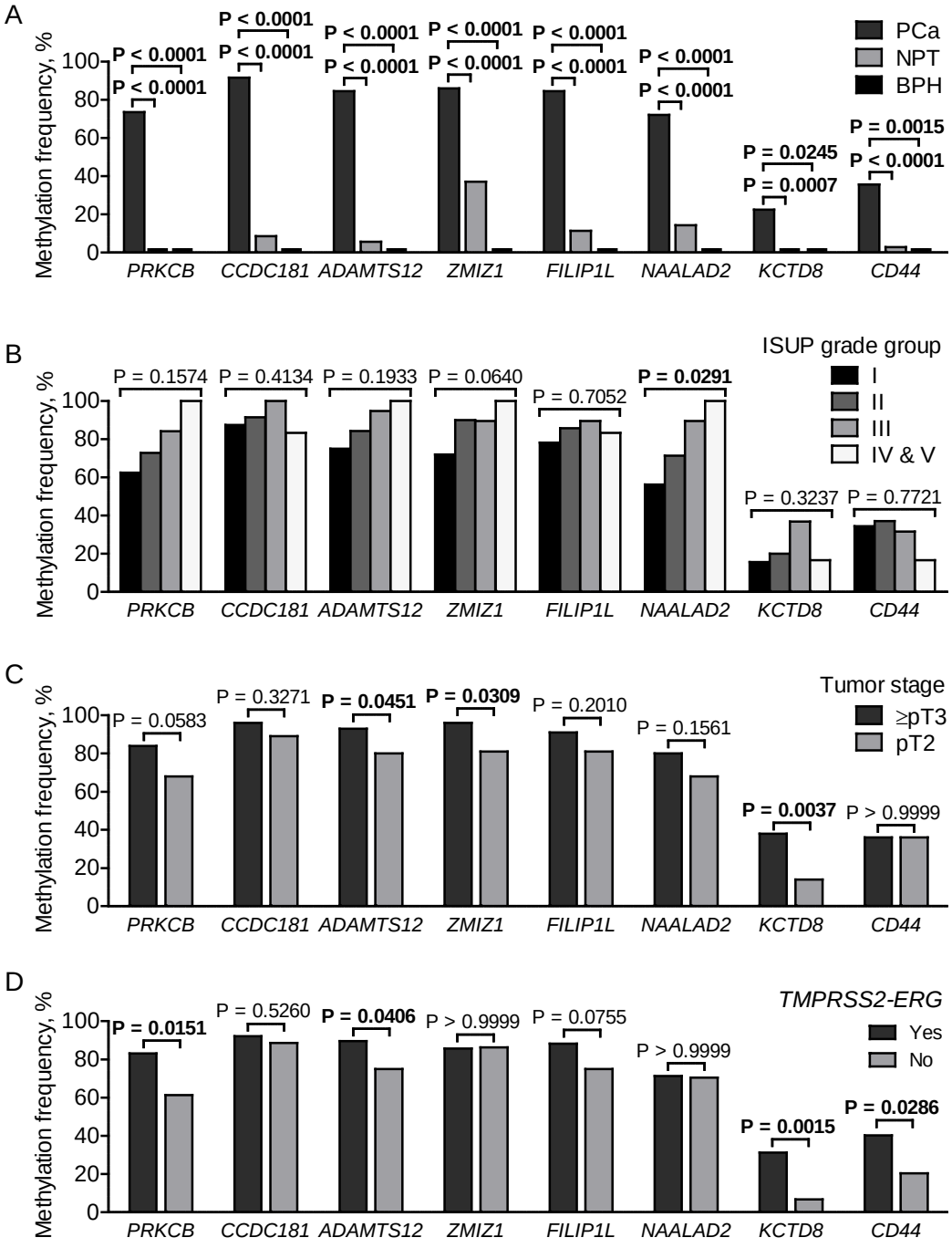
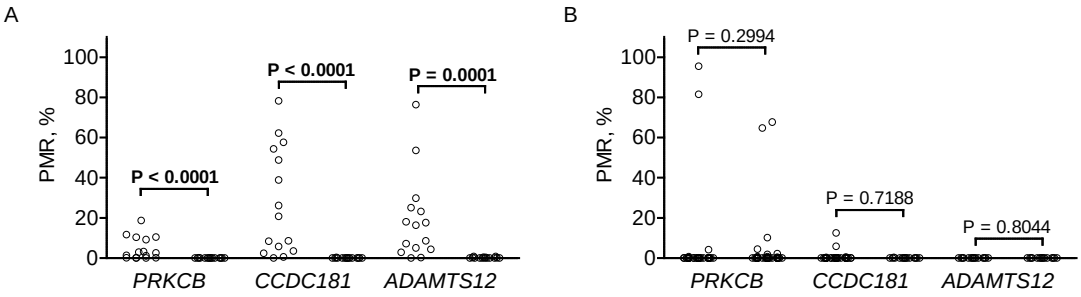


Figure 3.13. Promoter methylation frequencies of protein-coding genes in prostate samples according to tissue histology (A), grade groups by International Society of Urological Pathology (ISUP; B), tumor stage (C), and *TMPRSS2-ERG* fusion status (D). PCa – prostate tumors, NPT – noncancerous prostate tissues, BPH – benign prostatic hyperplasia tissues. Significant P-values are in bold.

The combined sensitivity for PCa of the genes *PRKCB*, *CCDC181*, and *ADAMTS12* was 94.6% (122 of 129), while the specificity reached 88.6% (31 of 35) and 100% (17 of 17) according to NPT and BPH, respectively. Besides, quantitative analysis showed that methylation levels of the three genes were higher in randomly selected 15 PCa than in 15 BPH tissue samples ($P < 0.0001$, $P < 0.0001$, and $P = 0.0001$, respectively; Fig. 3.14A).



Histology	PCa	BPH	PCa	BPH	PCa	BPH
Mean/ SE, %	4.83/ 1.50	0.01/ 0.01	27.76/ 6.87	0.02/ 0.02	19.23/ 5.48	0.19/ 0.07

Histology	PCa	BPH	PCa	BPH	PCa	BPH
Mean/ SE, %	3.37/ 2.30	2.12/ 1.82	2.84/ 1.72	0.54/ 0.53	0.36/ 0.25	1.23/ 0.89

Figure 3.14. Distribution of methylation intensity values (PMR) of genes *PRKCB*, *CCDC181*, and *ADAMTS12* in samples of prostate cancer (PCa) and benign prostatic hyperplasia (BPH) patients. A – PMR value distribution in tissues of PCa (N = 15) and BPH (N = 15) patients; B – PMR value distribution in urine samples of PCa (N = 54) and BPH (N = 15) patients. Mean values with standard errors (SE) are given below for each group. Statistically significant P-values are in bold.

No aberrant methylation events at promoter regions of genes *EPAS1* and *NEK9* were detected in a subset of 51 PCa and 15 NPT samples (data not shown) and, therefore, their further analysis was discontinued.

For the quantitative analysis of the promoter methylation of protein-coding genes, the PRAD dataset of TCGA was used (333 cases in total). In accordance with our data, significantly higher methylation levels were identified in tumors as compared to normal tissues for all 8 genes (all $P < 0.0001$; Fig. S1). Additionally, in tumors, methylation levels of genes *CD44* and *KCTD8* (median β-values 0.26 and 0.15, respectively) were lower than those of the other genes (median β-values ≥0.48), while *FILIP1L* was characterized by relatively high methylation in normal tissues (median β-values 0.89 and 0.84 in tumors and normal tissues, respectively; Fig. S1).

3.2.4. Association of protein-coding gene promoter methylation with clinical-pathological variables and *TMPRSS2-ERG* status

Aberrant promoter methylation of the protein-coding genes was further analyzed according to clinical-pathological patient's characteristics and *TMPRSS2-ERG* fusion status. Methylation frequency of *NAALAD2* significantly increased according to ISUP grade groups (P = 0.0291; Fig. 3.13B). Promoter methylation of *ADAMTS12*, *ZMIZ1*, and *KCTD8* was more frequent in ≥pT3 tumors as compared to pT2 (P = 0.0451, P = 0.0309, and P = 0.0037, respectively) and the same tendency was observed for *PRKCB* (P > 0.0500; Fig. 3.13C). Furthermore, genes *PRKCB*, *ADAMTS12*, *KCTD8*, and *CD44* were more commonly methylated in tumors positive for *TMPRSS2-ERG* fusion (P = 0.0151, P = 0.0406, P = 0.0015, and P = 0.0286, respectively; Fig. 3.13D). Associations with other clinical-pathological characteristics are provided in Table S12.

3.2.5. Protein-coding gene expression and association with promoter methylation and clinical-pathological variables

Based on promoter methylation frequencies and with regard to correlations with clinical-pathological variables and fusion transcript status, genes *PRKCB*, *CCDC181*, *ADAMTS12*, *NAALAD2*, and *ZMIZ1* were selected for expression analysis. RNA of sufficient quality was available of 81 PCa, 25 NPT, and 17 BPH samples (Table S2). Expression levels of *PRKCB*, *CCDC181*, *ADAMTS12* and *NAALAD2* were significantly lower in PCa as compared to NPT and BPH samples (all P < 0.0500; Fig. 3.15A-D). In the case of *ZMIZ1*, lower expression was observed in PCa than in NPT, but higher than in BPH samples (all P < 0.0500; Fig. 3.15E). Furthermore, lower expression levels of *PRKCB*, *CCDC181*, *ADAMTS12*, and *NAALAD2* in tissues of PCa patients correlated with methylated promoter status (all P ≤ 0.0001; Fig. 3.15F-I), while no such association was observed for *ZMIZ1* (P > 0.0500; Fig. 3.15J).

Consistent with our data, lower expression levels of *PRKCB*, *CCDC181*, *ADAMTS12*, and *NAALAD2* were observed in tumors as compared to normal

tissues in the PRAD cohort of TCGA (all P < 0.0500), however, the same association for *ZMIZ1* was not confirmed (P > 0.0500; Fig. S5). Moreover, the *PRKCB*, *CCDC181*, and *ZMIZ1* were expressed at lower levels in PCa samples with higher methylation intensity (all P < 0.0500), while no such association was observed for *ADAMTS12* and *NAALAD2* (P > 0.0500; Fig. S6).

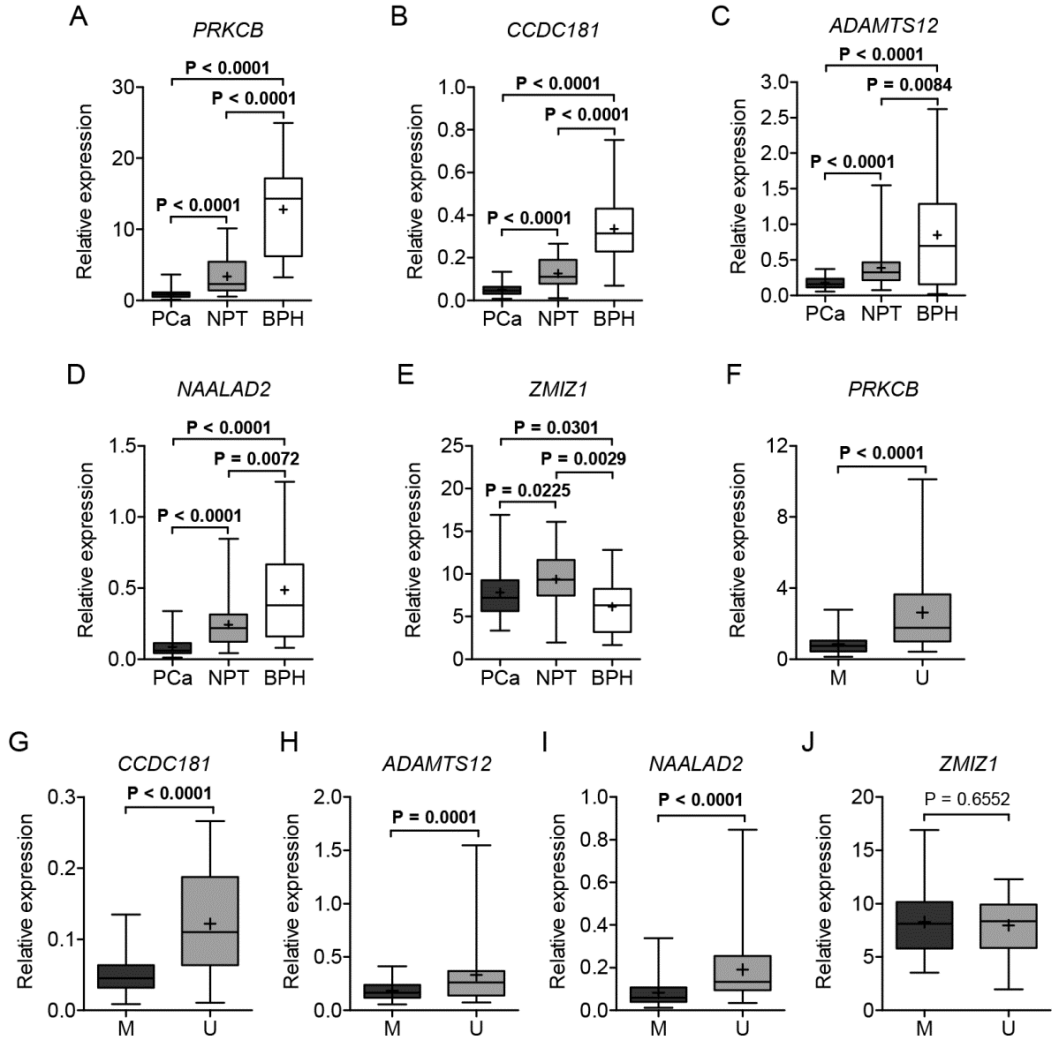


Figure 3.15. Relative expression levels of *PRKCB*, *CCDC181*, *ADAMTS12*, *NAALAD2*, and *ZMIZ1* in prostate tissues. A-E – expression of the genes in prostate tumors (PCa), noncancerous tissues (NPT), and benign prostatic hyperplasia (BPH) samples; F-J – expression of the genes in PCa according to the methylated/ unmethylated promoter status (M/ U). The box extends from the 25th to 75th percentiles; the line in the middle of the box is plotted at the median; the plus sign depicts the mean; the whiskers represent the range. Significant P-values are in bold.

In the Lithuanian cohort, decreasing expression levels of *CCDC181* and *NAALAD2* were significantly associated with higher ISUP grade group

(P = 0.0225 and P = 0.0139, respectively; Table S13). Furthermore, higher expression levels of *CCDC181* were specific for tumors expressing *TMPRSS2-ERG* transcript (P = 0.0136). No associations between gene expression and tumor stage pT, PSA level, or prostate mass were identified, however, expression of *NAALAD2* positively correlated with patients' age (R_S = 0.27, P = 0.0153; Table S13).

3.2.6. Quantitative methylation analysis of protein-coding genes in urine

Genes *PRKCB*, *CCDC181*, and *ADAMTS12* were selected for further quantitative evaluation of promoter methylation levels in urine. Methylation of the three genes was detectable in urine of PCa cases (N = 54), while lower or undetectable methylation levels were observed in BPH cases (N = 17), however, none of the differences were significant (all P > 0.0500; Fig. 3.14B).

As in the previous analysis, the PMR value in BPH of each gene was used as a cut-off for qualitative evaluation. Methylation of *PRKCB*, *CCDC181*, and *ADAMTS12* was detected in 5.6% (3 of 54), 14.8% (8 of 54), and 3.7% (2 of 54) of urine samples, with the combined sensitivity of 14.8% (8 of 54). The combined specificity of this assay reached 88.2% (15 of 17), whereas it was 88.2% (15 of 17), 94.1% (16 of 17), and 88.2% (15 of 17) for *PRKCB*, *CCDC181*, and *ADAMTS12*, respectively. No associations were detected between gene methylation intensity or frequency in urine and clinical-pathological variables (data not shown).

3.2.7. Prognostic value of protein-coding gene methylation

To investigate the prognostic value of the newly identified genes showing tumor-specific promoter methylation, the BCR-free survival analysis was performed. Aberrant methylation of *PRKCB*, *ADAMTS12*, and *NAALAD2* was significantly more frequent in BCR-positive than BCR-negative cases (P = 0.0038, P = 0.0030, and P = 0.0010, respectively; Fig. 3.16).

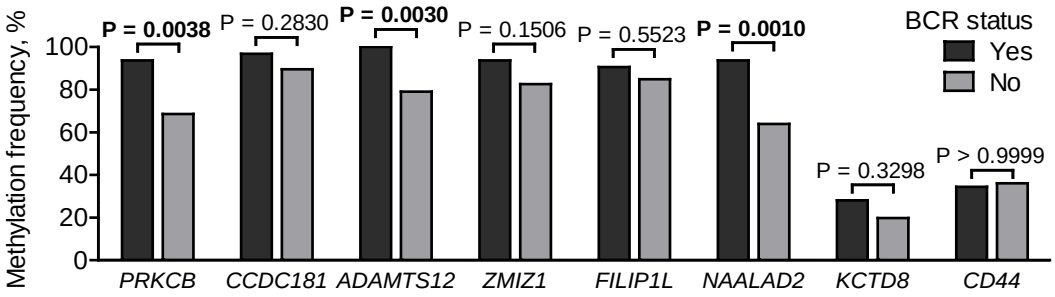


Figure 3.16. Protein-coding gene promoter methylation frequencies in prostate tumors according to biochemical disease recurrence (BCR) status. Significant P-values are in bold.

Kaplan-Meier curve comparison showed significantly earlier BCR in PCa cases with methylated *PRKCB*, *ADAMTS12*, and *NAALAD2* promoter status (P = 0.0066, P = 0.0036, and P = 0.0005, respectively), while no associations were observed for the other analyzed genes (Fig. 3.17). The significance of *PRKCB*, *ADAMTS12*, and *NAALAD2* promoter methylation as independent prognostic factors was supported by univariate and multivariate Cox proportional hazard analyses, however, regarding gene expression levels only *NAALAD2* showed significant prognostic value (all models' P < 0.0500; Tables S14 and S15). Furthermore, methylation status of the three genes was also prognostic for BCR-free survival in combination with *TMPRSS2-ERG* fusion status (all P < 0.0500; Table S15). Forward entering of covariates revealed that promoter methylation of *PRKCB* or *NAALAD2* together with tumor stage pT surpassed the prognostic value of the pathological variable alone (models' P < 0.0001 and P < 0.0001, respectively), and the similar impact was observed of *NAALAD2* methylation status on ISUP grade group (model's P < 0.0001). Some other combinations with clinical-pathological variables are provided in Table S15.

Additionally, methylation status of *PRKCB*, *ADAMTS12*, and *NAALAD2* was a significant independent predictor of BCR-free survival in subgroups of patients <60 years of age (P = 0.0038, P = 0.0088, and P = 0.0124, respectively), diagnosed with pT2 tumors (P = 0.0024, P = 0.0092, and P = 0.0087, respectively), ISUP grade group I or II tumors (P = 0.0069,

P = 0.0024, and P = 0.0005, respectively), or in cases with <10 ng/µL preoperative PSA level (P = 0.0163, P = 0.0036, and P = 0.0011, respectively; data not shown).

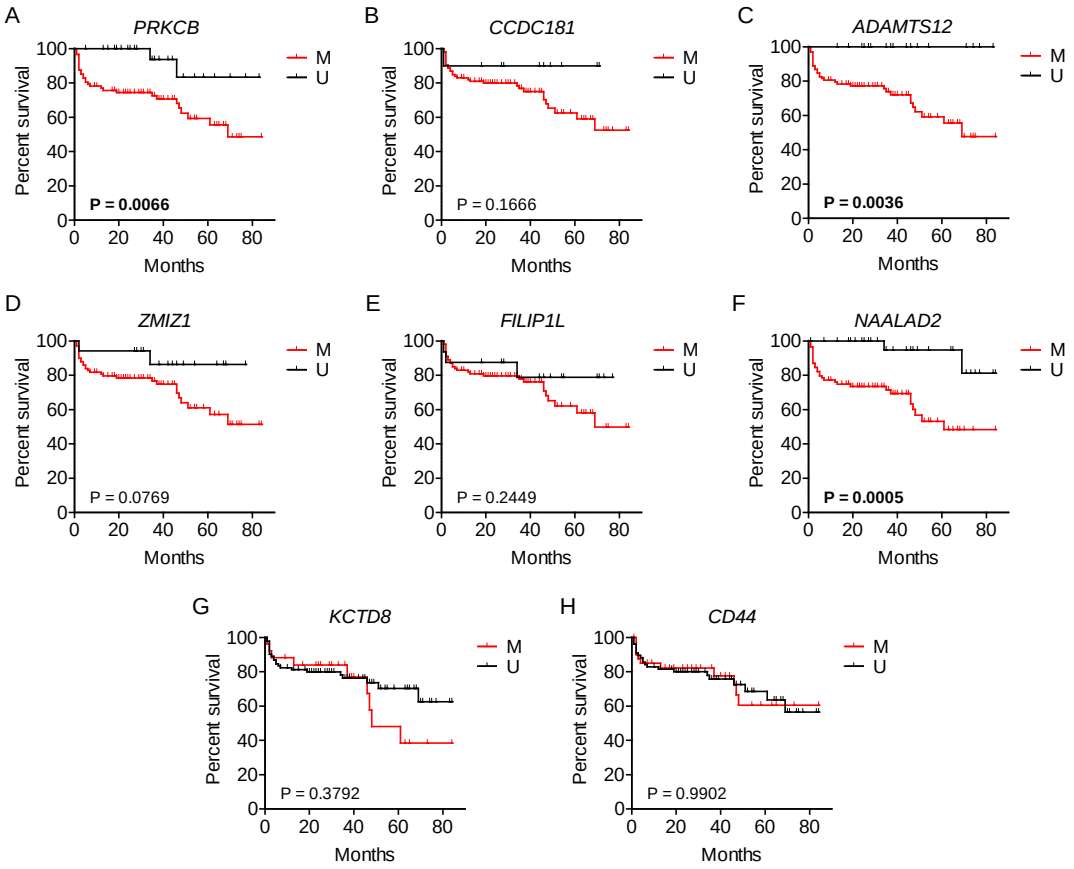


Figure 3.17. Kaplan-Meier curve analysis of *PRKCB* (A), *CCDC181* (B), *ADAMTS12* (C), *ZMIZ1* (D), *FILIP1L* (E), *NAALAD2* (F), *KCTD8* (G), and *CD44* (H) gene promoter methylation status. M/ U – methylated/ unmethylated promoter status. Significant P-values are in bold.

Quantitative DNA methylation analysis of *PRKCB*, *CCDC181*, and *ADAMTS12* in urine of PCa patients did not show any significant associations with BCR-free survival (data not shown).

For TCGA data, disease-free survival, in addition to BCR-free survival, was utilized as an endpoint for the evaluation of prognostic potential of the selected biomarkers. PCa cases with prior cancer diagnosis and/ or prior neoadjuvant therapy, having Gleason score 10, and metastatic cases were filtered out in survival analysis in order to better match the Lithuanian cohort. In the PRAD dataset of TCGA, methylation levels of the analyzed protein-

coding genes were not associated with neither BCR nor disease progression in univariate Cox models, however, decreased expression levels of *PRKCB* and *ADAMTS12* were prognostic for BCR-free survival (both P < 0.0500; Table S14). Among the multivariate models, significant prognostic value was confirmed only for methylation levels of *PRKCB* or *NAALAD2* in combination with pT and methylation of *NAALAD2* together with ISUP grade group (all P < 0.0500; Table S15).

3.2.8. Validation of miRNA host gene promoter methylation

Based on methylation levels according to tissue histology or BCR status and with regard to the number of a particular gene-associated probes showing significant methylation differences, five miRNA coding loci were selected for further DNA methylation analysis (Table 3.1).

Table 3.1. MiRNAs and their host genes selected for promoter methylation analysis.

miRNA		Host gene				DNA methylation differences (probes, N)	
Mature miRNA	Gene symbol	Location (strand)	Gene symbol	Biotype	miRNA location	PCa *vs.* NPT	BCR+ *vs.* BCR-
miR-155-5p, -3p	*MIR155*	21q21.3 (+)	*MIR155HG/ BIC/ MIRHG2*	ncRNA	exon	2	0
miR-152-5p, -3p	*MIR152*	17q21.32 (-)	*COPZ2*	protein-coding	intron	1	2
miR-137	*MIR137*	1p21.3 (-)	*MIR137HG*	ncRNA	exon	9	0
miR-31-5p, -3p	*MIR31*	9p21.3 (-)	*MIR31HG*	ncRNA	intron	5	0
miR-642a-5p, -3p	*MIR642A*	19q13.32 (+)	*GIPR/ PGQTL2*	protein-coding	intron*	2	3
miR-642b-5p, -3p	*MIR642B*	19q13.32 (-)					

Nomenclature of miRNAs and genes is given according to miRBase release 21 (http://www.mirbase.org) and GeneCards® databases (http://www.genecards.org). PCa – prostate cancer, NPT – noncancerous prostate tissue, BCR+/- – biochemical disease recurrence status (positive/ negative).
*5'-UTR/ exon of predicted transcripts.

DNA methylation status of the regulatory regions of the 5 selected miRNAs was examined by means of MSP. Promoter-associated CGIs of mir-155, mir-152, and mir-137 host genes were frequently hypermethylated in

PCa, but not in NPT and/ or BPH samples (all P < 0.0500; Fig. 3.18A). Hypermethylation of mir-31 and mir-642a, -b host gene promoters was rarely detected in PCa, while mir-642a, -b was the only miRNA whose hypermethylation was observed in BPH (Fig. 3.18A).

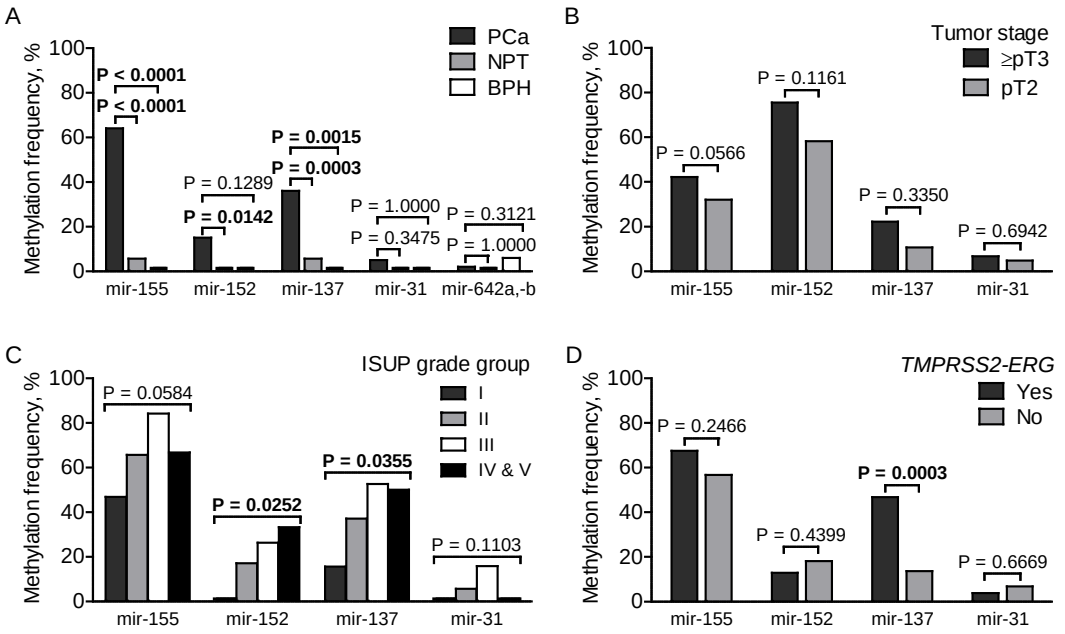


Figure 3.18. MiRNA host gene promoter methylation frequencies in prostate samples according to tissue histology (A), tumor stage (B), Gleason grade group by International Society of Urological Pathology (ISUP; C), and *TMPRSS2-ERG* fusion transcript (D). PCa – prostate tumors, NPT – noncancerous prostate tissues, BPH – benign prostatic hyperplasia tissues. Significant P-values are marked in bold.

In accordance with our data, mir-155, mir-152, mir-137, and mir-31 showed significantly higher methylation levels in tumors than in normal prostate tissues in the PRAD cohort of TCGA (all P < 0.0001; Fig. S7). Besides, in this dataset, the difference of methylation levels was also significant for mir-642a, -b host gene (P = 0.0013).

3.2.9. Association of miRNA host gene promoter methylation with clinical-pathological variables and *TMPRSS2-ERG* status

Methylation frequencies in PCa samples were further analyzed according to clinical-pathological patients' characteristics and *TMPRSS2-ERG* fusion status. No significant associations were observed between miRNA host gene methylation and pT (Fig. 3.18B). Mir-152 and mir-137 methylation frequency

significantly increased according to ISUP grade group, while a trend was observed for mir-155 ($P = 0.0252$, $P = 0.0355$, and $P > 0.0500$, respectively; Fig. 3.18C). Comparing two adjacent ISUP grade groups, promoter methylation frequencies of mir-152 and mir-137 showed the most significant difference between ISUP groups I and II ($P = 0.0163$ and $P = 0.0367$, respectively). Promoter methylation of mir-137 host gene was more frequent in tumors expressing *TMPRSS2-ERG* fusion transcript ($P = 0.0010$; Fig. 3.18D). Associations with other parameters are provided in detail in Table S16.

3.2.10. MiRNA expression and association with host gene promoter methylation

To assess whether the host gene promoter methylation status was associated with the deregulation of the analyzed miRNAs, the expression levels of the predominant mature miRNA sequences – miR-155-5p, miR-152-3p, miR-137, and miR-31-5p – were reanalyzed in a set of 41 PCa and 12 NPT cases from our previous study [244]. Highly significant down-regulation of miR-155-5p, miR-152-3p, and miR-31-5p was observed in PCa as compared to NPT samples, however, there was no difference in expression levels of miR-137 (Fig. 3.19A-D).

Matched miRNA expression and DNA methylation data were available for a subset of 30 PCa and 7 NPT cases. Decreased miR-155-5p expression significantly correlated with the methylated promoter status ($P = 0.0375$), however, no associations between miRNA expression and host gene promoter methylation was observed for miR-152-3p, miR-137, and miR-31 (all $P > 0.0500$; Fig. 3.19E-H).

For the validation in a larger set of samples, expression of the selected miRNAs was also investigated in the PRAD cohort of TCGA. In this dataset, down-regulation of mir-155 and mir-152 significantly correlated with higher promoter methylation levels of their host genes ($R_P = -0.23$, $P < 0.0001$ and $R_P = -0.25$, $P < 0.0001$, respectively; Fig. S8). Besides, promoter methylation

was associated with the decreased expression of mir-155 and mir-152 host gene mRNAs (R_P = -0.32, P < 0.0001 and R_P = -0.57, P < 0.0001, respectively). Expression analysis also indicated that these two miRNAs were co-expressed with their host gene mRNAs (R_P = 0.74, P < 0.0001 and R_P = 0.32, P < 0.0001, respectively). Expression levels of mir-137 and mir-31 did not correlate with promoter methylation (both P > 0.0500; Fig. S8), while expression data on host genes of these miRNAs was not obtainable for more thorough analysis.

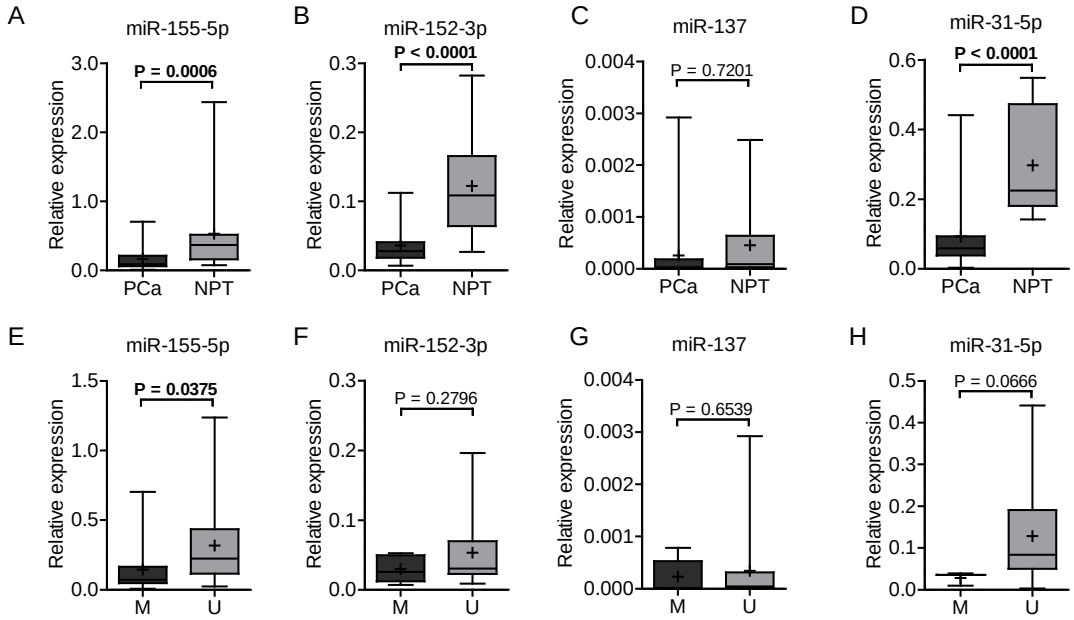


Figure 3.19. Comparison of the relative expression levels of the analyzed miRNAs according to prostate tissue histology and promoter methylation status of the respective host genes. A-D – miRNA expression in prostate tumors (PCa) and noncancerous prostate tissues (NPT). E-H – miRNA expression in prostate tissues according to the methylation status of the respective miRNA host gene promoters. M/ U – methylated/ unmethylated promoter status. The box extends from the 25th to 75th percentiles; the line in the middle of the box is plotted at the median; the plus sign depicts the mean; the whiskers represent the range. Significant P-values are in bold.

3.2.11. Analysis of miRNA target genes

MiRNAs with the most frequent methylation of the host gene promoters in PCa – miR-155-5p, miR-152-3p, and miR-137 – were selected for their target gene analysis. The putative transcriptional targets of these miRNAs were obtained querying six online databases that utilize different methods of target prediction (Table S17). Large numbers of obtained predicted targets were

screened for genes associated with epigenetic modulation as defined in EpiFactors database (http://epifactors.autosome.ru) [260]. Overlapping epigenetic regulator genes were chosen for the analysis. DNA methyltransferase 1 (*DNMT1*) was selected as a target of miR-155-5p and miR-152-3p, while lysine demethylases 1A (*KDM1A*/ *LSD1*) and 5B (*KDM5B*/ *JARID1B*) – as targets of miR-155-5p and miR-137 (Fig. 3.20A and B, and Table S18).

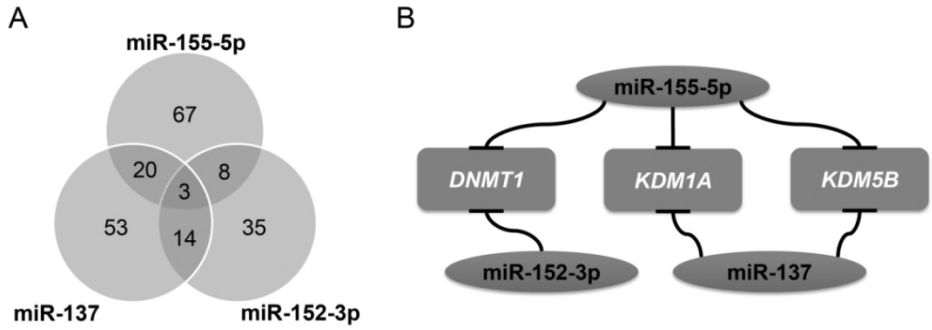


Figure 3.20. MiRNA-targeted genes associated with epigenetic regulation and identified in at least 2 of 6 databases (A) and interaction of miRNAs with transcriptional targets selected for gene expression analysis (B).

In our cohort, expression levels of *DNMT1* and *KDM5B* were higher in tumors and noncancerous tissues of PCa cases as compared to BPH and, thus, distinguishing PCa and BPH patients (Fig. 3.21A-C). Only expression of *KDM5B* was higher in PCa than in NPT cases, while no significant differences of *KDM1A* expression levels were observed among the histological groups.

Expression of the selected target genes was further linked to the promoter methylation status of the corresponding miRNA host genes. *KDM5B* was expressed at higher levels in samples with methylated mir-137 or mir-155 host gene promoters ($P = 0.0001$ and $P < 0.0001$, respectively), whereas increased expression of *DNMT1* and *KDM1A* was significantly associated with mir-152 and mir-155 host gene methylation, respectively ($P = 0.0093$ and $P = 0.0302$; Fig. 3.21D-F).

Expression of *KDM1A* and *KDM5B* correlated inversely with preoperative PSA level ($R_S = -0.29$, $P = 0.0096$ and $R_S = -0.26$, $P = 0.0184$, respectively; Fig. 3.21G and H). Moreover, the two genes were upregulated in PCa cases

expressing *TMPRSS2-ERG* fusion transcript (P = 0.0048 and P = 0.0011, respectively; Fig. 3.21I). No associations between *DNMT1*, *KDM1A*, or *KDM5B* and pT, ISUP grade group, prostate mass, or patients' age were observed (data not shown).

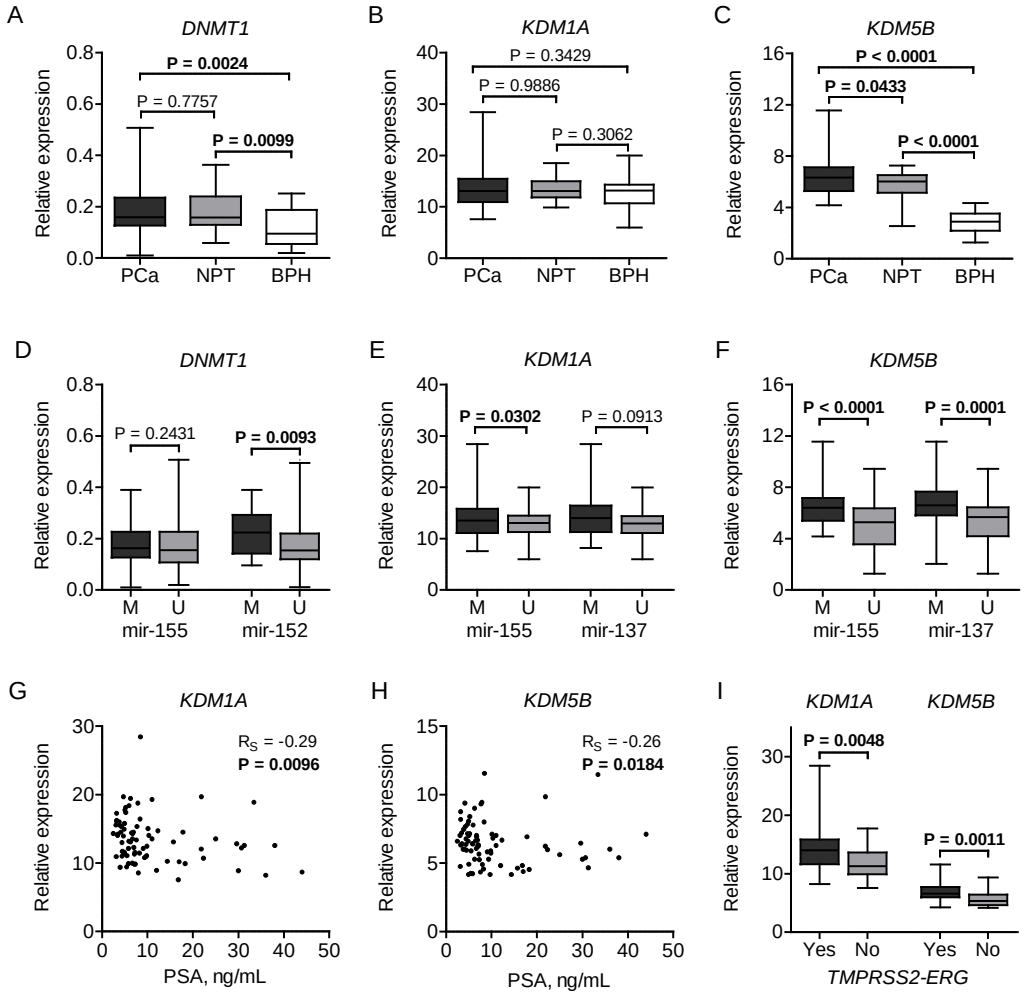


Figure 3.21. Relative expression levels of *DNMT1*, *KDM1A*, and *KDM5B* in prostate tissues. A-C – expression of the genes in prostate tumors (PCa), noncancerous tissues (NPT), and benign prostatic hyperplasia (BPH) samples; D-F – expression of the genes in PCa according to the promoter methylation status of the miRNA host genes; G and H – correlations between gene expression and PSA level; I – expression of the genes according to *TMPRSS2-ERG* fusion transcript status. Only significant associations with PSA and fusion transcript status are shown. The box extends from the 25th to 75th percentiles; the line in the middle of the box is plotted at the median; the whiskers represent the range. Significant P-values are in bold.

In TCGA cohort, significant correlations were validated between *KDM5B* expression and mir-155 or mir-137 promoter methylation (R_P = 0.18, P = 0.0007; and R_P = 0.32, P < 0.0001, respectively) and between *KDM1A*

expression and mir-155 methylation (R_P = 0.28, P < 0.0001; Fig. S9). Conversely to our data, mir-137 methylation level correlated with increased *KDM1A* expression (R_P = 0.40, P < 0.0001), while higher methylation levels of mir-155 were linked to increased *DNMT1* expression (R_P = 0.16, P = 0.0032). The up-regulation of *DNMT1* in association with methylated mir-152 promoter was also supported by the data from TCGA cohort (R_P = 0.12, P = 0.0236; Fig. S9).

3.2.12. Prognostic value of miRNA host gene promoter methylation

Having shown the impact of miRNA host gene methylation on target mRNA expression, we aimed to determine the prognostic value of these observations. Methylation of mir-152 and mir-31 was significantly more common in BCR-positive than BCR-negative cases (P = 0.0213 and P = 0.0191, respectively) and the same tendency was observed for mir-155 (P > 0.0500; Fig. 3.22).

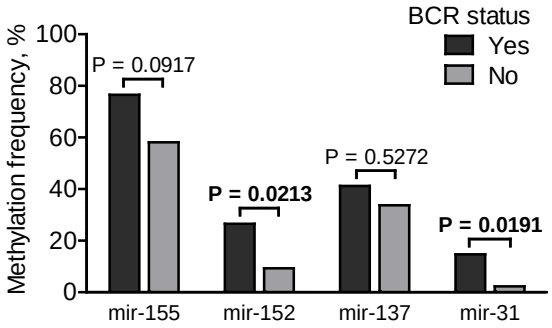


Figure 3.22. MiRNA host gene promoter methylation frequencies in prostate tumors according to biochemical disease recurrence (BCR) status. Significant P-values are in bold.

Kaplan-Meier analysis revealed that methylation of mir-155, mir-152, or mir-31 promoter had significant negative impact on BCR-free survival, whereas no association was observed for mir-137 (P = 0.0371, P = 0.0193, P = 0.0102, and P > 0.0500, respectively; Fig. 3.23A-D). For mir-155, mir-152, and mir-31, the 5-year BCR-free survival was 51.9% *vs.* 75.6% (23.7% difference), 27.3% *vs.* 69.7% (42.4%), and 0% *vs.* 67.4% (67.4%) in cases with methylated and unmethylated promoter status, respectively. Expression levels

of miRNA target genes *DNMT1*, *KDM1A*, and *KDM5B* were not associated with BCR status (all P > 0.0500; data not shown).

Univariate Cox proportional hazard analysis confirmed the association of promoter methylation status of mir-155, mir-152, and mir-31 host genes with BCR-free survival (models‘ P = 0.0447, P = 0.0236 and P = 0.0168, respectively; Table S14). Combinations of methylation status of various miRNA host genes also showed prognostic value in multivariate models, among which a combination of mir-152 and mir-31 was the most significant (model’s P = 0.0130). Moreover, multivariate analysis revealed that promoter methylation of mir-155 host gene alone or together with mir-152 was prognostic for BCR-free survival in combination with *TMPRSS2-ERG* fusion-negative status (models‘ P = 0.0333 and P = 0.0151, respectively; Table S19).

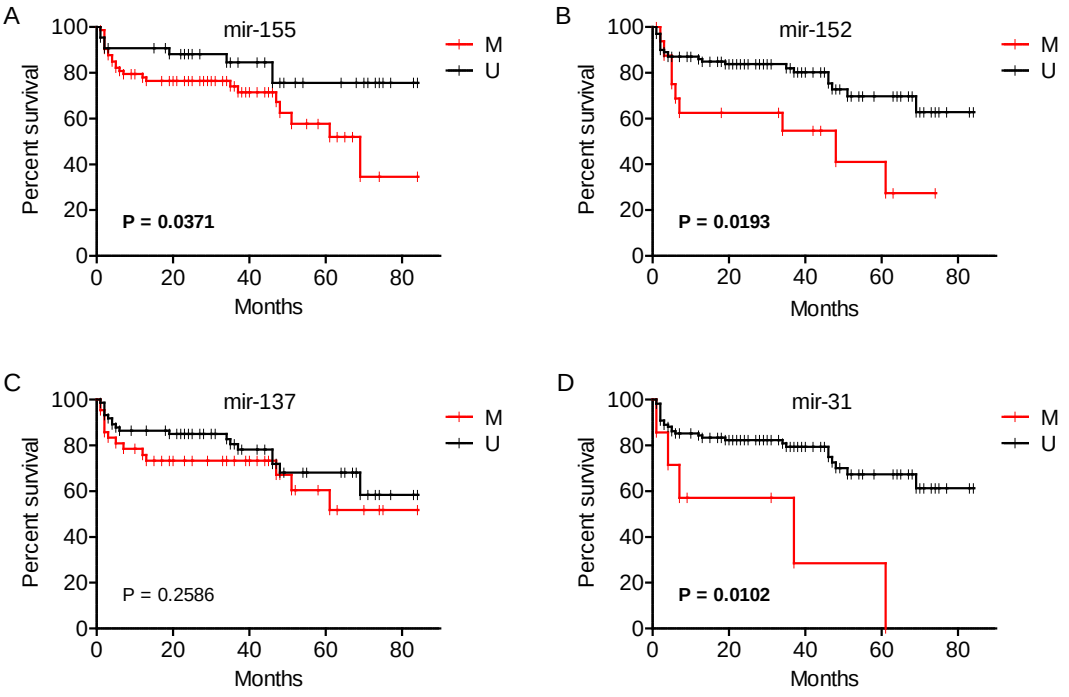


Figure 3.23. Kaplan-Meier curve analysis of mir-155 (A), mir-152 (B), mir-137 (C), and mir-31 (D) host gene promoter methylation status. Numbers of patients at risk and survival proportions are provided below each graph. M/ U – methylated/ unmethylated promoter status. Significant P-values are in bold.

Among clinical-pathological parameters, various combinations were prognostic of BCR-free survival, but only pT and ISUP grade group showed

independent prognostic value (models‘ P = 0.0001 and P < 0.0001, respectively; Table S14). Forward entering of covariates revealed that methylation status of mir-152 or mir-31 host genes together with pT even better predicted BCR-free survival than the pathological variable alone (both models‘ P < 0.0001). Moreover, methylation of mir-31 also had the same effect in combination with ISUP grade group (model‘s P < 0.0001). Other combinations of miRNAs with clinical-pathologic variables were also rather promising (Table S19).

In TCGA cohort, higher methylation levels of mir-137 and mir-31 host genes were associated with disease-free (P = 0.0110 and P = 0.0130, respectively), but not with BCR-free survival (both P > 0.0500; data not shown), and these associations were supported by univariate Cox models (models’ P = 0.0122 and P = 0.0188, respectively; Table S14). In multivariate analysis, forward entering of covariates showed that methylation levels of mir-137 or mir-31 host genes increased the prognostic value of pT (models’ P = 0.0006 and P = 0.0007, respectively; Table S19).

Although expression of the analyzed miRNA targets was not prognostic for BCR status in the Lithuanian cohort, higher expression levels of *KDM5B* were characteristic to shorter BCR-free survival in TCGA cohort (P = 0.0093), while increased *DNMT1* expression significantly predicted time to disease relapse (P = 0.0033; Table S14). Various multivariate models of target genes alone or in combinations with promoter methylation of the respective regulatory miRNAs were also prognostic for BCR-free and/ or disease-free survival. However, none of these multivariate models surpassed the significance of the individual variables (data not shown).

4. DISCUSSION

Despite the emerging improvements in diagnosis and treatment, PCa remains one of the most prevalent malignancies and a leading cause of death in men. Early diagnosis of PCa, as well as the selection of the most proper treatment strategy, is critical for successful management of the disease. Although the introduction of PSA test has significantly decreased the numbers of men diagnosed with already metastatic PCa, it has also resulted in the increase of overdiagnosis and/ or overtreatment of men with non-life threatening disease. This encouraged the search of new molecular tools with better performance in risk stratification. DNA methylation of tumor suppressor genes is a potential source of such biomarkers.

4.1. DNA methylation of known tumor suppressor genes in prostate cancer

In the first stage of the present study, we tested a panel of *RARB*, *GSTP1*, *RASSF1*, *MGMT*, *DAPK1*, $p16^{INK4a}$/*CDKN2A*, and $p14^{ARF}$/*CDKN2A* promoters for the DNA methylation status in 149 PCa tissues and 253 urine samples from PCa patients. Aberrant methylation of at least one of the genes was identified in 90.6% of PCa cases, and the most informative biomarkers were *RASSF1*, *GSTP1*, and *RARB*. Methylation analysis of this 3-gene panel in urine showed 60.1% sensitivity and 68.8% specificity for PCa. Moreover, in patients with Gleason score 6 tumors, *RASSF1* methylation together with pT were significant predictors of BCR when analyzed in tissue or urine samples.

In consistency with previous studies [162,164,165,261], *RARB*, *GSTP1*, and *RASSF1* were the most commonly methylated genes in PCa with the frequencies of over 50%, whereas aberrant methylation of *MGMT*, *DAPK1*, *p16*, and *p14*, was detected in ≤25% of PCa cases in our study. Similarly to other studies [157,262], promoter methylation of *RARB* or *GSTP1* rarely occurred in NPT or BPH tissues, resulting in the specificity of these biomarkers of up to 85%. Although the specificity of *RASSF1* methylation for PCa was relatively low (67%) in the qualitative analysis, the quantification of

methylation levels revealed statistically significant differences from BPH and NPT. Similarly, quantification of methylation levels in *RASSF1* and some other genes revealed significant differences in PCa in comparison to BPH and HGPIN in another QMSP-based study [157], showing a progressive increase of methylation levels in *RASSF1* during prostate carcinogenesis.

Since methylation assessment in panels of several genes shows increased diagnostic and prognostic power, multiple gene combinations have been analyzed in PCa [157,160,162,175]. In the present study, the 3-gene panel (*RARB*, *GSTP1*, and *RASSF1*) showed 85.2% sensitivity for PCa when analyzed in tumors and 60.1% sensitivity when assayed in urine samples. In other studies, more than 85% sensitivity has been reported for panels of 2-4 genes in PCa tissues [160,162,175], with lower sensitivity (≥50%) obtained in similar studies of urine samples [203,204,196]. Although *GSTP1* has been recurrently reported as the most commonly methylated gene in body fluids of PCa patients (reviewed in [205]), *RASSF1* was the most frequently methylated gene in urine in our study. Methylation of *RASSF1* was detected in 45% of urine samples, with the mean methylation level close to 20%. In several previous studies [203,204], even higher frequencies (≥73%) of *RASSF1* methylation has been detected in urine samples. High frequency and intensity of *RASSF1* methylation in urine samples from PCa patients makes it an easily detectable and specific biomarker of PCa. In the present study, *RARB* and *GSTP1* methylation in urine was less frequent, but in agreement with the recent meta-analysis [182,206] the specificity of these biomarkers, especially of *GSTP1*, was very high (>96%), exceeding that of the PSA test. The limited sensitivity of biomarkers in body fluids is a common issue [182,206]. This might be at least partly explained by varying amounts of tumor-derived DNA in the context of nontumor cells, such as leucocytes, which might be present at considerable amounts in urine. Nevertheless, the high specificity of such assays could be particularly useful for appointing or changing the treatment strategy as reported in the recent studies [211,212].

In agreement with previous reports [175,263], advanced pT and higher Gleason score were significant predictors of BCR in our study. Promoter methylation of a single particular gene is rarely found to be an independent predictor for disease outcome (discussed in [152]). In this study, methylation of *RASSF1* in PCa tissue correlated with PSA level and BCR. In combination with pT, methylated promoter status of *RASSF1* was a significant predictor of BCR-free survival following RP in cases with Gleason score 6 tumor, i.e. the group of low- or intermediate-risk PCa, and such association was also observed in urine samples. In fact, this is the first study ever to report the potential prognostic value of a DNA methylation biomarker in urine of PCa. The reason why methylation of *RASSF1* was a significant predictor for BCR-free survival only in Gleason score 6 tumors may be related to higher intratumoral morphological heterogeneity of Gleason score 7 tumors, which, according to the most recent ISUP recommendations [78], are currently subdivided into 3 + 4 and 4 + 3 tumors based on differences in prognosis. Besides, based on the accumulating data from genome-wide PCa studies [8,188], it might be presumed that other prognostic factors might overwhelm the impact of *RASSF1* methylation in higher-grade PCa. No correlations between *RASSF1* methylation and BCR have been reported in previous studies [160,164], which might be also due to the small-size samples analyzed.

GSTP1 is the most frequently studied gene in PCa, but the data regarding its prognostic value are quite contradictory [263]. In the present study, methylation of *GSTP1* in PCa tissues and in urine showed associations with clinical markers of poor prognosis, but was not BCR-predictive. However, the combination of *DAPK1* and *RASSF1* methylation was a significant predictor for BCR-free survival. In agreement with other studies, this suggests that panels of genes, in addition to improved sensitivity, might also provide a superior power to predict the aggressive PCa course after curative treatment.

To date, a DNA methylation biomarker-based PCa test, “ConfirmMDx for Prostate Cancer”, is commercially available, which has been developed at

around the same time as the implementation of this study. Aiming to assist in treatment decision-making process after a negative biopsy, the test evaluates methylation of three TSGs, two of which, namely *GSTP1* and *RASSF1*, have been also analyzed in the present study. As this test has been already included in the EAU-ESTRO-SIOG Guidelines on Prostate Cancer 2016 [59], this is an excellent example of translating DNA methylation research into clinical applicability.

Presented findings support the notion that aberrant promoter methylation of a limited panel of genes might provide additional information on PCa aggressiveness. Aberrant promoter methylation of *RASSF1* in both tissues and urine could be considered as a biomarker of a potentially unfavorable course of the disease in low- or intermediate-risk PCa patients after RP. Further studies are mandated to validate the prognostic significance of this biomarker.

4.2. Identification of novel DNA methylation biomarkers of prostate cancer

The rapid development of novel technologies applicable for high-throughput profiling, such as microarrays and NGS, have made it possible to conduct genome-wide methylation analysis in cancer, primarily searching for novel biomarkers that could outperform the currently existing ones. Indeed, recent studies utilizing these methods have made significant insights into PCa methylome in general and identified a large number of novel potential DNA methylation biomarkers. However, only a few of these biomarkers have been successfully validated in clinical samples, usually in a single cohort [3,53,190], while the vast majority have not been further tested at all.

In the present study, microarray-based approach was used for screening of diagnostic and prognostic DNA methylation biomarkers in a small set of well characterized PCa and matched NPT samples. In agreement with other studies [3,7,53,190,264], differential methylation of numerous loci comparing PCa *vs.* NPT and BCR-positive *vs.* BCR-negative groups was identified in our study,

and the gene sets were significantly enriched for genes involved in various pathways commonly deregulated in cancer. Consistent with previous reports [3,5], methylation differences were observed in both intragenic and intergenic loci. Several previously reported genes, such as *MAGI2* [3,9], *CCDC181* [265], *HOXA7* [266], *HOXD3* [5], *TERT* [267], and *SEPT9* [268], were among those with the most significant methylation differences in PCa. However, the majority of the top-listed genes identified in our study have never been examined in PCa or in other types of cancer before, and similar observations have been made in previous studies [5,190]. Interestingly, none of the most established DNA methylation biomarkers for PCa, namely *GSTP1*, *RASSF1*, *RARB*, and *APC*, were detected as differentially methylated in our analysis, which could have been caused by the limited sample size. Consistently, Kron *et al.* were also unable to capture differential methylation of *GSTP1*, the most characteristic epigenetic aberration in PCa [5], whereas Devaney *et al*. reported evidence of methylation of both *GSTP1* and *RASSF1* promoters, although neither of the two genes was present in the top list of cancer *vs.* reference [190]. Despite the variable sample numbers and different analysis techniques used for PCa methylome profiling in these studies, the inability to detect significant methylation differences of widely studied TSGs might be attributable to PCa heterogeneity and/ or multifocality.

Although the microarray design, utilized in the present study, covered various CGIs regardless of their co-localization with annotated regulatory regions, for further target gene validation we focused on promoter-associated methylation differences as downregulated gene expression by DNA methylation has been extensively described in PCa [269]. However, recent genome-wide methylation studies have provided contradictory data. Although a previous study [3] showed that increasing promoter CGI methylation indeed correlated with gene repression, another study determined this association in only 12.5% of CGIs [53]. As we did not examine the transcriptome in the current study, our selection of targets for validation was primarily based on the

fact that gene silencing is commonly induced by methylation of CpGs that are in proximity to promoter or exon 1 loci.

Among the large number of differentially methylated genes identified in the present study, 10 protein-coding genes (*PRKCB*, *CCDC181*, *ADAMTS12*, *ZMIZ1*, *FILIP1L*, *NAALAD2*, *KCTD8*, *EPAS1*, *NEK9*, and *CD44*) were selected for further validation in 129 PCa, 35 NPT, and 17 BPH tissue samples. Methylation of the three gene panel, *PRKCB*, *CCDC181*, and *ADAMTS12*, showed high sensitivity (94.6%) and specificity for PCa (≥88.6%), and was also detectable in urine.

Protein kinase C beta (*PRKCB*) belongs to the family of genes coding serine/ threonine kinases, which are associated with diverse functions in cancer cells, including transformation, proliferation, and inhibition of apoptosis, and mainly acts through microtubule regulation in mitosis [270]. In the present study, methylation of *PRKCB* was frequent (73.6%) and specific (100%) to PCa. Besides, it was also successfully detected in urine of PCa patients. To date, promoter methylation of *PRKCB* has only been investigated in lung tumors by QMSP and the reported methylation levels ranged up to ~25% [271]. Intriguingly, by analyzing TCGA dataset of lung cancer authors identified a positive correlation between methylation of two CpGs and gene expression [271]. In our analysis of PCa and control tissues both in the Lithuanian and TCGA cohorts, hypermethylated *PRKCB* promoter status was strongly associated with tumor-specific down-regulation of gene expression. In support, treatment with demethylating agent has been previously shown to restore *PRKCB* expression [272]. However, this is contradictory to several gene expression studies that all reported up-regulation of *PRKCB* in PCa [273,274] and its oncogenic role through increased proliferation of PCa cells and activated angiogenesis [275]. Analysis of *PRKCB* 5'-promoter revealed two Sp1 binding sites, an important transcriptional regulator, in a region between -110 bp and the first exon [272]. Furthermore, Metzger *et al.* showed that PRKCB interacts with both KDM1A and AR, and specifically associates

with AREs at promoter or enhancer regions of androgen-regulated genes [274]. In fact, the authors demonstrated that PRKCB determines the dual specificity of KDM1A by phosphorylation of histone H3 at threonine 6 (H3T6) and, therefore, controls AR-dependent gene expression during PCa progression [274]. Owing to the known fact that *PRKCB* encodes several distinct isoforms, the role of promoter methylation on *PRKCB* expression and its function in PCa requires further investigation.

The function of *coiled-coil domain containing 181* gene (*CCDC181*; also known as *C1orf114*) is currently unknown. Highly PCa-specific methylation of *CCDC181* was first reported by Haldrup *et al.* [265] and most recently by Moller *et al.* [52]. In our study, *CCDC181* was the most frequently methylated gene (91.5%) of the 10-gene panel, with 100% specificity for PCa. Besides, our analysis revealed that methylation of *CCDC181* could be also detectable in urine of PCa patients and, thus, shows potential for noninvasive diagnostics. Consistent to our observations, a four-gene panel including *CCDC181* was recently proposed to increase sensitivity for PCa diagnostics through detection of epigenetic field effects in histologically-normal prostate tissues [52]. Although in our study the prognostic potential of *CCDC181* was not determined neither in the Lithuanian nor TCGA cohorts, methylation of this gene contained independent predictive value for BCR-free survival in two RP patient cohorts in the previous analysis [265].

ADAM metallopeptidase with thrombospondin type 1 motif 12 gene (*ADAMTS12*) gene, associated with angiogenesis [276-278], was also selected for validation in the present study. So far, *ADAMTS12* has never been evaluated for promoter methylation in PCa, thus, this is the first study to report the potential diagnostic value of this sensitive (84.5%) and highly PCa-specific (100%) biomarker. Besides, methylation of this gene was also observed in urine. Epigenetic silencing of *ADAMTS12*, as a novel TSG, has been previously reported in colon cancer and, similarly to our data, was associated with disease progression [279].

NAALAD2 gene encodes N-acetylated alpha-linked acidic dipeptidase 2, which has been mostly studied only from the biochemical point of view. Despite the representative member of this NAALADase gene family, prostate-specific membrane antigen gene (*PSM/PSMA/FOLH1*), a marker of PCa, little is known about *NAALAD2* itself. To the best of our knowledge, this is the first study to investigate *NAALAD2* promoter methylation and gene expression in clinical samples. Our analysis revealed potential diagnostic and prognostic value of this putative biomarker. Despite its high sensitivity (72.1%) and specificity (100%) for PCa, promoter methylation of *NAALAD2* had independent predictive value for BCR-free survival in various combinations with *PRKCB* and/ or *ADAMTS12*. Besides, *NAALAD2* methylation in combination with pT or ISUP grade group increased the prognostic power of the pathological variable alone. Further investigation is needed to confirm our preliminary observations, as well as to evaluate the potential of *NAALAD2* methylation as a noninvasive DNA methylation biomarker.

The protein encoded by *zinc finger MIZ-type containing 1* gene (*ZMIZ1*) is a transcription factor, best known as a modulator of AR-dependent transcription [280]. Our study indicated that in the Lithuanian cohort promoter methylation of *ZMIZ1* was frequent in PCa (86.0%) and also moderately common, as compared to the other genes, in NPT samples (37.1%), although the frequencies were still significantly different. Surprisingly, down-regulation of *ZMIZ1* was not associated with the promoter methylation in Lithuanian samples, although this correlation was observed in TCGA dataset. Besides, in TCGA cohort, the differences of both methylation frequencies and levels were the largest among the analyzed genes, while no differential expression was observed between PCa and normal tissues. This suggests that other mechanisms, despite DNA methylation, might be responsible for transcriptional regulation of this gene. Previously, ZMIZ1 recruitment on the promoter of *KLK3* (*PSA*) gene was demonstrated in PCa cells [281], indicating the potential significance of this gene in PCa progression to CRPC under

androgen deprivation therapy. Consistently, PSA level in combination with *ZMIZ1* methylation status was predictive of BCR-free survival in the Lithuanian cohort, however, neither variable had independent prognostic value. Until now, promoter methylation of *ZMIZ1* in cancer has not been reported and, therefore, its role in PCa development and progression remains to be elucidated in the future.

Filamin A interacting protein 1 like gene (*FILIP1L*; also known as *downregulated in ovarian cancer 1*, *DOC1*) encodes an important regulator of cell invasion and metastasis and acts as a TSG. In agreement with the previous study [282], promoter methylation of *FILIP1L* was frequent in PCa tissues (84.5%) and significantly differentiated tumors from benign lesions (≥88.6% specificity). It has been noted that methylation of this gene, particularly coding isoform 2, was associated with downregulated expression in various tumors, including PCa [282-284]. Most of the studies also reported a significant association of *FILIP1L* methylation and/ or down-regulation with an invasive cancer phenotype [283,284]. In our study, no significant differences of methylation frequencies were observed in association with BCR or other clinical-pathological parameters, most likely due to the stringent inclusion of early or intermediate stage, non-metastatic PCa cases in the study group.

Analysis results of the rest four genes were less promising, primarily due to the relatively rare or absent promoter methylation in PCa. Methylation of *CD44* and *potassium channel tetramerization domain containing 8* gene (*KCTD8*) was uncommonly detected in PCa (≤35.7%), although cancer-specific (≥97.1% specificity). *CD44* encodes a transmembrane glycoprotein (a receptor of hyaluronic acid, growth factors, and cytokines), which is important in cell adhesion, migration, and cell-cell interactions [285]. Although *CD44* showed some evidence of differential methylation in microarray-based screening, inclusion of this gene in the validation set was merely based on previously observed promoter methylation and decreased expression in PCa associated with aggressive course of the disease [171,184]. In the present

study, *CD44* methylation was more common in PCa than in benign tissue samples. However, we were not able to confirm the prognostic value of this gene, which is in agreement with the previous study by Alumkal *et al.* [173]. These discrepancies might be at least partly explained by the differences in cohort structures among the studies.

No significant associations with clinical-pathological variables, including BCR status, were observed between *KCTD8* methylation and clinical-pathological parameters, with the exception of higher *KCTD8* methylation frequency in tumors of more advanced stage. Similarly, in the only reported *KCTD8* methylation study, higher median methylation levels of this gene were detected in breast tumors than in normal tissues and differed significantly between invasive and *in situ* carcinomas [286]. Together this indicates that promoter methylation of *KCTD8* might have a putative role in cancer progression, although its diagnostic potential appears to be uncertain.

Due to the complete absence of methylation events in both PCa and control tissues of a pilot set of samples, investigation of *endothelial PAS domain protein 1* (*EPAS1*; also known as *hypoxia-inducible factor 2 alpha*, *HIF2A*) and *NIMA related kinase 9* (*NEK9*) genes was suspended. Both genes had been chosen for validation based on their differential methylation in PCa tissues comparing BCR-positive and BCR-negative cases. *EPAS1* was the only gene showing hypomethylation that was selected for validation in this study. This is contradictory to the studies reporting epigenetic silencing of *EPAS1* in other tumors [287,288]. Presently, no studies exist describing methylation analysis of *NEK9*. Considering its vital function even in cancerous cells [289], it could be speculated that DNA methylation-induced repression of this gene might not be attributable to PCa. Nevertheless, further analysis is required to elucidate the role of epigenetic mechanisms in regulating *NEK9* and *EPAS1* expression.

4.3. Clinical significance of miRNA host gene promoter methylation in prostate cancer

For the last decade, the role of miRNAs in the maintenance of cellular homeostasis and tumor development has been widely accepted. A rapidly growing amount of miRNA expression profiling studies has revealed a wide range of alterations in miRNA expression. Although the mechanisms responsible for deregulated miRNA expression are still not fully understood, miRNA silencing linked to aberrant DNA methylation at promoter regions appears to be of major importance in disrupting miRNA profile in various cancer types [290,291], including prostate cancer [234,292-294].

The present study let us identify PCa-specific promoter methylation of mir-155 host gene accompanied by significant down-regulation of the mature miRNAs. To date, DNA methylation, as a mechanism involved in mir-155 regulation, has been described in few studies analyzing breast cancer [291], multiple myeloma [295], and gastric cancer cell lines [296], therefore, this is the first study to report mir-155 host gene methylation in PCa associated with miRNA silencing. Due to interactions with its various targets, miR-155-5p is considered as a multifunctional miRNA as ambiguous data have been obtained from different types of cancer [297-299]. This discrepancy might be partly explained by differences in cohort composition and/ or model systems used.

MiR-152 is another frequently studied miRNA, however, only a few publications investigating its regulation by CpG island methylation exist [300,301]. In addition to aberrant methylation in PCa samples as compared to NPT or BPH, our results showed that the promoter of mir-152 host gene was also more frequently methylated in higher ISUP grade tumors and was associated with higher PSA level, demonstrating potential clinical significance. Supporting our data, miR-152 inactivation was found associated with promoter methylation in PCa cell lines by other authors [293]. Interestingly, decrease in miR-152 expression was more common in PCa tissues of African American patients as compared to Caucasians [293].

Recently, extensive methylation of mir-137 locus has emerged as a significant phenomenon in different types of tumors [290,291,302] including preclinical PCa cell models of localized and castrate resistant cancer [294]. In our study, more frequent promoter methylation of mir-137 host gene was detected in PCa *vs.* NPT and BPH samples. Similarly to the analysis of a clinical dataset in a recent study [294], this miRNA was more commonly methylated in tumors of higher ISUP grade groups in the present study. However, promoter methylation-associated miR-137 silencing was confirmed only in TCGA data, most likely due to a limited number of cases in our cohort.

Survival analysis revealed that promoter methylation status of mir-155, mir-152, and mir-31 host genes was significant as independent predictors of BCR-free survival in univariate and various multivariate models and even enhanced the prognostic potential of clinical-pathological variables. Methylation status of mir-155 as a covariate contributed to various multivariate models. In particular, mir-155 alone or together with mir-152 significantly predicted time to BCR in PCa cases with negative *TMPRSS2-ERG* fusion status. Despite the low methylation frequency in PCa, mir-31 augmented the prognostic value of pT and ISUP grade group and also significantly predicted disease-free survival in TCGA data analysis. To date, the prognostic potential of these miRNAs has been mostly evaluated at the miRNA expression level [303,304], whereas studies analyzing the effect of epigenetic repression on PCa patient's survival are lacking [227]. Consistent to our data, loss of miR-155, miR-152, miR-137, and miR-31 and other tumor-suppressive miRNAs has been associated with BCR or disease progression in PCa, but the main set of these studies has been performed on treatment-resistant cell lines and rarely in large groups of PCa cases [244,293,303-305].

Recent evidence suggest the idea that deregulated miRNAs may lead to aberrant DNA methylation and histone modification profile observed in cancer (reviewed in [239]). Therefore, after querying 6 miRNA target databases, as well as recent publications, we focused on epigenetic regulatory genes,

previously shown to be essential in cancer epigenome remodeling. DNA methyltransferase 1 (*DNMT1*), mainly responsible for maintaining methylation patterns following DNA replication, has been validated as a target of miR-152-3p in endometrial tumors and PCa cell lines [226,293,306], whereas targeting by miR-155-5p is predicted by bioinformatic tools and is weakly substantiated [307]. Our results indicated higher *DNTM1* expression levels in cancerous and NPT samples as compared to BPH. In support with previous studies, up-regulation of *DNMT1* was determined in PCa cases with methylated mir-152 promoter, but no such correlation was observed for mir-155. However, associations of *DNMT1* expression with promoter methylation levels of the host genes were observed for the both miRNAs analyzing TCGA data, which might be produced due to different samples sizes and other characteristics. Besides the epigenetic regulation of *DNMT1*, a direct binding of miR-155-5p to the catalytic region of DNMT1 enzyme was recently shown [241] suggesting further need for functional studies of miR-155-5p in cancer cells.

KDMs play important roles in PCa progression and particularly in transition from the androgen-dependent to androgen-independent state. KDM1A is the first identified and the most deeply studied lysine demethylase. It specifically demethylates H3K4me1/2 and, therefore, favors gene silencing. However, when recruited by AR, KDM1A mainly acts as a coactivator in PCa by removing H3K9me1/2 repressive marks and leading to the activation of AR signaling [308,309]. In the present study, no expression differences of *KDM1A* were observed between PCa and NPT samples or in association with BCR status in Lithuanian cohort. However, higher *KDM1A* levels alone or in combination with other molecular markers had prognostic significance predicting BCR-free and/ or disease-free survival in TCGA data. High *KDM1A* expression in primary PCa has been also shown to predict higher risk of relapse after prostatectomy in other studies [309].

A recent study has provided mechanistic insight into the formation of the *TMPRSS2-ERG* fusion and showed that *KDM1A* is critically important for this gene rearrangement, as well as for the expression of the fusion transcript [310]. Supporting the fact, we found significantly higher *KDM1A* expression in PCa cases with *TMPRSS2-ERG* fusion. MiR-137 has been previously shown to target *KDM1A* in colorectal cancer cells [311], while miR-155-5p interaction with this lysine demethylase is only predicted. In support to this notion, we determined higher *KDM1A* expression in cases with methylated mir-155 promoter status both in Lithuanian and TCGA datasets, however, similar association with mir-137 methylation was observed only in TCGA cohort. In consideration with the recent study [310] and owing to the above-mentioned associations and the frequent mir-137 host gene promoter methylation in *TMPRSS2-ERG*-positive cases, we hypothesize that this rearrangement might be induced by mir-137 host gene promoter methylation through the increased *KDM1A* expression (Fig. 4.1).

Overexpression of another H3K4me3/2 demethylase, *KDM5B*, has been observed in PCa [312]. Consistently, we found significant differences of *KDM5B* levels in prostate tissues, with the highest expression observed in PCa. Similarly to *KDM1A*, the expression of *KDM5B* was also significantly higher in *TMPRSS2-ERG* positive PCa cases. *KDM5B* has been recently identified as a target of miR-137 in PCa cell lines [294], while the same role of miR-155-5p has been only predicted by computational methods. To the best of our knowledge, our study for the first time associated the upregulated *KDM5B* expression with methylated status of mir-137 and mir-155 host gene promoters. This was validated in TCGA cohort further signifying the importance of these two miRNAs in PCa development. However, further experimental evidence supporting the direct regulation of *KDM5B*, as well as *KDM1A*, expression by miR-155-5p is needed. Previous studies have confirmed that KDM5B, like KDM1A, also interacts with AR and acts as an oncogene [312], although the mechanism is not yet fully understood. Taken

together, this data suggests that KDMs might play a far more complex role in prostate carcinogenesis that transcends our current knowledge.

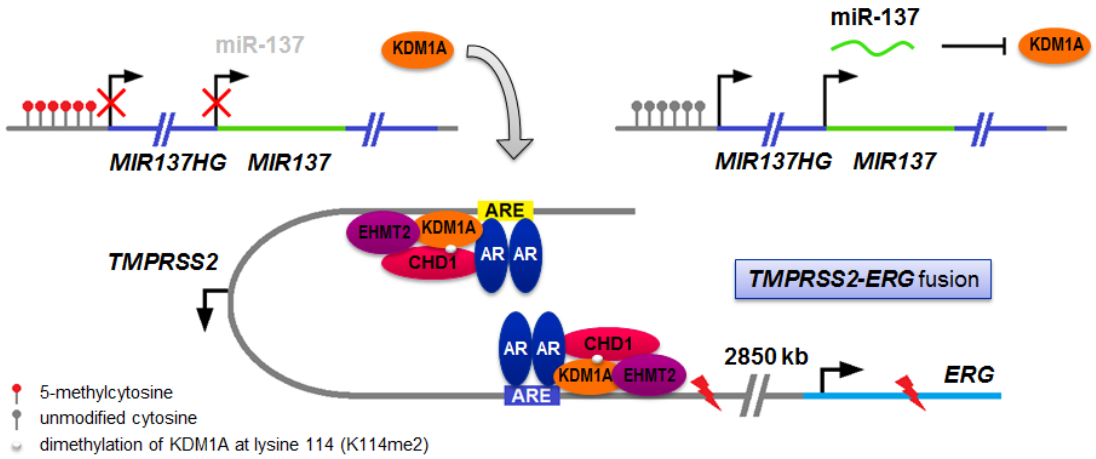


Figure 4.1. The hypothesized role of mir-137 host gene promoter methylation in the formation of *TMPRSS2-ERG* fusion in PCa. Methylation of mir-137 host gene (*MIR137HG*) promoter represses mir-137, as well as *MIR137HG*, expression (upper left). In the absence of the mature miR-137, translation of *KDM1A*, a key player in *TMPRSS2-ERG* formation, is not blocked and, therefore, this lysine demethylase can form complexes with other proteins involved in fusion formation. In the absence of methylation at *MIR137HG* promoter, both *MIR137HG* and mir-137 genes are transcribed (upper right). The mature miR-137 can suppress *KDM1A* expression at post-transcriptional level and, therefore, decrease the amount of KDM1A. As recently demonstrated by Metzger *et al.* (lower part) [310], histone methyltransferase EHMT2 methylates KDM1A at lysine 114 (K114me2), which is a key event controlling androgen-dependent gene transcription. The assembly of methylated KDM1A and chromodomain helicase DNA binding protein 1 (CHD1), the reader of K114me2, at the enhancer (ARE in yellow) and breakpoint (ARE in blue) regions drives AR-dependent loop formation and *TMPRSS2-ERG* fusion. Red lightning bolts show double-strand breaks, transcription start sites are depicted as arrows. Lower part of the figure adapted from [310].

In conclusion, the present study has shown aberrant promoter methylation of mir-155, mir-152, mir-137, and mir-31 host genes as promising diagnostic and/ or prognostic biomarkers of PCa. Our observations suggest that methylation status of particular miRNA genes as independent variables or in combinations can predict BCR-free survival and even increase the prognostic value of clinical-pathological variables. Aberrantly methylated miRNA host genes, as well as epigenetic factors regulated by these miRNAs, might be attractive targets for epigenetic therapeutics.

As a result of the recent technological advancement in molecular sciences, significant insights have been made in the genomic and epigenomic landscape of PCa. This led to identification of numerous putative biomarkers aiming to overcome the limitations of the ones that are currently used, however, few of them have been thoroughly validated in clinical datasets. DNA methylation at gene promoter regions is probably the best source of such biomarkers as detection of such alterations could be easily applicable for noninvasive testing. Within the next decade, new DNA methylation-based biomarker tests are expected to become available that would assist clinicians not only in early PCa diagnostics, but, more importantly, in timely differentiation of aggressive from indolent malignancy and, thus, the most proper treatment selection.

CONCLUSIONS

1. In PCa tissues, promoter methylation of genes *RARB, GSTP1*, and *RASSF1* was ≥54.1%, while methylation of *MGMT, DAPK1*, *p16*, and *p14* was ≤25.5%. The combined sensitivity for PCa of *RARB, GSTP1*, and *RASSF1* was 85.2%, while the specificity was 62.2% or 47.1% analyzing NPT or BPH samples, respectively.

2. In urine of PCa cases, promoter methylation frequencies of *RASSF1*, *RARB*, and *GSTP1* were 44.7%, 29.2%, and 11.1%, respectively, with the combined sensitivity of 60.1%. The specificity according to BPH was 68.8%.

3. Promoter methylation of *RASSF1* and *DAPK1* in PCa tissues had independent prognostic value for BCR-free survival (both $P < 0.0500$). In patients diagnosed with Gleason score 6 tumor, *RASSF1* methylation analyzed both in PCa tissues and urine was a significant predictor for BCR-free survival when combined with pT (both models' $P < 0.0010$).

4. Significant ($P < 0.0500$) methylation differences in 1143 protein-coding ($FC \geq 1.5$) and 522 miRNA host genes ($FC \geq 1.2$) were identified comparing PCa and NPT samples, and in 1804 protein-coding ($FC \geq 1.2$) and 129 miRNA host genes ($FC \geq 1.2$) comparing BCR-positive and BCR-negative PCa cases by microarray-based DNA methylation profiling. Biological and molecular processes occurring during tumor development and progression were commonly enriched among the sets of differentially methylated genes.

5. Methylation frequencies were ≥72.1% for *PRKCB, CCDC181*, *ADAMTS12*, *ZMIZ1*, *FILIP1L*, and *NAALAD2*, but did not exceed 35.7% for *KCTD8* and *CD44*, while only unmethylated promoter status (0%) was observed for *EPAS1* and *NEK9*. The combined sensitivity for PCa of *PRKCB*, *CCDC181*, and *ADAMTS12* was 94.6%, while the specificity reached 88.6% and 100% analyzing NPT and BPH,

respectively. Methylated promoter status was associated with lower expression levels of *PRKCB*, *CCDC181*, *ADAMTS12*, and *NAALAD2* in PCa (all $P \leq 0.0001$). Methylation of *PRKCB*, *CCDC181*, and *ADAMTS12* was also detectable in urine of PCa patients.

6. Promoter methylation of *PRKCB*, *ADAMTS12*, and *NAALAD2* in PCa tissues had independent prognostic value for BCR-free survival (all models' $P < 0.0500$). Methylation of *PRKCB* and *NAALAD2* augmented the prognostic value of pT and/ or ISUP grade group (all $P < 0.0001$).

7. Promoter methylation frequencies of mir-155, mir-152, and mir-137 host genes ranged from 14.7% to 64.3% in PCa and significantly differed from NPT and/ or BPH samples (all $P < 0.0500$), while methylation of mir-31 and mir-642a, -b host genes was rare (≤5.4%). Methylated mir-155 host gene promoter status correlated with decreased miR-155-5p expression ($P = 0.0375$). Mir-137 host gene was more frequently methylated in tumors expressing *TMPRSS2-ERG* fusion transcript ($P = 0.0010$).

8. Promoter methylation status of mir-152 and mir-31 host genes was independent prognostic factors for BCR-free survival and augmented the prognostic value of pT and/ or ISUP grade group (all models' $P < 0.0500$). Promoter methylation of mir-155 host gene alone or together with mir-152 was prognostic for BCR-free survival in combination with *TMPRSS2-ERG* fusion-negative status (both models' $P < 0.0500$).

9. Higher *KDM5B* expression was detected in samples with methylated mir-155 or mir-137 host gene promoters, whereas upregulation of *KDM1A* and *DNMT1* was associated with mir-155 and mir-152 methylation status, respectively (all $P < 0.0500$). *KDM1A* and *KDM5B* were upregulated in PCa cases expressing *TMPRSS2-ERG* fusion transcript (both $P < 0.0050$).

TRANSLATIONAL RELEVANCE

In the present study, selected TSGs and novel proposed targets, identified from PCa methylome screening, were evaluated as potential diagnostic and/ or prognostic DNA methylation biomarkers. With moderate sensitivity and specificity, PCa was successfully detected in urine using a set of 3 TSGs (*RARB*, *RASSF1*, and *GSTP1*). Furthermore, in patients diagnosed with Gleason 6 tumor, methylation of *RASSF1* in combination with pT was predictive of BCR-free survival in tissues or urine. This indicates the potential value of this epigenetic alteration for PCa prognosis and, thus, treatment decision making at early stages of the malignancy.

PCa methylome profiling data revealed numerous differential methylation events between PCa and NPT, as well as BCR-positive and BCR-negative cases. This further led to identification of a novel putative PCa biomarkers (*PRKCB*, *CCDC181*, *ADAMTS12*, and others) with excelling diagnostic and/ or prognostic potential, which were validated in two independent cohorts. Associations between promoter methylation and downregulated gene expression showed the significance of these genes in PCa development and/ or progression, while the detectability of these putative biomarkers in urine indicated their potential utility for noninvasive routine testing.

Common methylation of miRNA host genes proved to be cancer-specific and predictive of PCa progression. Moreover, together with pT or ISUP grade group, methylation of mir-152 and mir-31 host genes augmented the prognostic value of the pathological variables. Furthermore, analysis of epigenetic factors regulated by the miRNAs indicated potential targets for epigenetic therapeutics.

In summary, the present study proposed novel putative DNA methylation biomarkers for PCa diagnosis and/ or prognosis, especially for low- or intermediate-risk PCa, and also provided some new insights into the dysregulated epigenetic mechanisms in PCa development.

PUBLICATIONS AND PRESENTATIONS

Publications

Research articles, directly related to the topic of doctoral dissertation, published in journals with a citation index (IF) in the Clarivate Analytics Web of Science platform:

1. **Daniunaite K**, Dubikaityte M, Gibas P, Lazutka JR, Bakavicius A, Ulys A, Jankevicius F, Jarmalaite S. Clinical significance of microRNA host gene promoter methylation in prostate cancer. *Accepted for print in Hum Mol Genet (Oxford Univ Press)*. doi: 10.1093/hmg/ddx138. [IF_{2015} 5.985; IF_{5yr} 6.353]
2. **Daniunaite, K**, Jarmalaite S, Kalinauskaite N, Petroska D, Laurinavicius A, Lazutka JR, Jankevicius F. Prognostic value of *RASSF1* promoter methylation in prostate cancer. J Urol. 2014; 192(6): 1849-1855. doi: 10.1016/j.juro.2014.06.075. [IF_{2014} 4.360; IF_{5yr} 4.064]

Other publications, not directly related to the topic of doctoral dissertation, in journals with a citation index in the Clarivate Analytics Web of Science platform:

1. Kubiliute R, Sulskyte I, **Daniunaite K**, Daugelavicius R, Jarmalaite S. Molecular features of doxorubicin-resistance development in colorectal cancer CX-1 cell line. Medicina (Kaunas). 2016; 52(5): 298-306. doi: 10.1016/j.medici.2016.09.003. [IF_{2015} 0.609; IF_{5yr} 0.700]
2. Stuopelyte K, **Daniunaite K**, Bakavicius A, Lazutka JR, Jankevicius F, Jarmalaite S. The utility of urine-circulating miRNAs for detection of prostate cancer. Br J Cancer. 2016; 115(6): 707-715. doi: 10.1038/bjc.2016.233. [IF_{2015} 5.569; IF_{5yr} 5.617]
3. Stuopelyte K, **Daniunaite K**, Jankevicius F, Jarmalaite S. Detection of miRNAs in urine of prostate cancer patients. Medicina (Kaunas). 2016; 52(2): 116-124. doi: 10.1016/j.medici.2016.02.007. [IF_{2015} 0.609; IF_{5yr} 0.700]
4. Demidenko R, Razanauskas D, **Daniunaite K**, Lazutka JR, Jankevicius F, Jarmalaite S. Frequent down-regulation of ABC transporter genes in prostate cancer. BMC Cancer. 2015; 15: 683. doi: 10.1186/s12885-015-1689-8. [IF_{2015} 3.265; IF_{5yr} 3.642]
5. **Daniunaite K**, Serenaite I, Misgirdaite R, Gordevicius J, Unguryte A, Fleury-Cappellesso S, Bernotiene E, Jarmalaite S. Epigenetic regulation of human adipose-derived stem cells differentiation. Mol Cell Biochem. 2015; 410(1-2): 111-120. doi: 10.1007/s11010-015-2543-7. [IF_{2015} 2.613; IF_{5yr} 2.438]

6. Serenaite I, **Daniunaite** K, Jankevicius F, Laurinavicius A, Petroska D, Lazutka JR, Jarmalaite S. Heterogeneity of DNA methylation in multifocal prostate cancer. Virchows Arch. 2015; 466(1): 53-59. doi: 10.1007/s00428-014-1678-3. [IF_{2015} 2.627; IF_{5yr} 2.644]

7. Stuopelyte K, **Daniunaite K**, Laurinaviciene A, Ostapenko V, Jarmalaite S. High-resolution melting-based quantitative analysis of *RASSF1* methylation in breast cancer. Medicina (Kaunas). 2013; 49(2): 78-83. [IF_{2013} 0.508; IF_{5yr} 0.632]

Presentations at scientific conferences

Presentations directly related to the topic of doctoral dissertation:

- Oral presentations:

1. **Daniunaite K**, Jarmalaite S. DNA methylation biomarkers for prostate cancer diagnosis and prognosis. Biofuture: The Perspectives of Natural and Life Sciences. Vilnius, Lithuania. December 7, 2016. [award for best presentation]

2. Mackoit A, **Daniunaite K**, Jarmalaite S. Epigenetic profile of hypoxia and angiogenesis-related genes and its prognostic relevance in prostate cancer patients. Riga Stradins University International Student Conference. Riga, Latvia. March 16, 2016.

3. **Daniunaite K**, Jarmalaite S. Epigenetic alterations in prostate cancer. The 1st LUHS Medical Molecular Biology Conference: Biomarkers of Chronic Diseases. Kaunas, Lithuania. April 17, 2015.

4. **Daniunaite K**, Jankevicius F, Laurinavicius A, Lazutka JR, Jarmalaite S. Global DNA methylation analysis in prostate tumors. Biofuture: The Perspectives of Natural and Life Sciences. Vilnius, Lithuania. December 9, 2014. [award for best presentation]

5. **Daniunaite K**, Jarmalaite S, Kalinauskaite N, Jankevicius F, Laurinavicius A, Lazutka JR. Frequent DNA methylation of tumor suppressor genes in urine sediments from patients with early-stage prostate cancer. Urinomics 2013 – 1st International Conference on Urine Omics. Caparica, Portugal. September 9-11, 2013.

- Poster presentations:

1. **Daniunaite K**, Dubikaityte M, Rauluseviciute I, Mackoit A, Razanauskas D, Lazutka JR, Jankevicius F, Jarmalaite S. Novel DNA methylation biomarkers of prostate cancer. The COST Meeting "Creating New Collaborative Research Proposals". Groningen, Netherlands. September 15-16, 2016.

2. **Daniunaite K**, Rauluseviciute I, Dubikaityte M, Mackoit A, Razanauskas D, Lazutka JR, Jankevicius F, Jarmalaite S. Identification of novel DNA methylation biomarkers of prostate cancer. XIV International Conference of the Lithuanian Biochemical Society. Druskininkai, Lithuania. June 28-30, 2016.

3. Mackoit A, **Daniunaite K**, Jarmalaite S. DNA methylation analysis of angiogenesis-related genes *ADAMTS12* and *FILIP1L* in prostate cancer. The COINS International Conference of Life Sciences. Vilnius, Lithuania. February 29 – March 3, 2016.

4. Rauluseviciute I, **Daniunaite K**, Jarmalaite S. DNA methylation of genes, selected by microarray analysis, in prostate cancer. The COINS International Conference of Life Sciences. Vilnius, Lithuania. February 29 – March 3, 2016.

5. **Daniunaite K**, Razanauskas D, Dubikaityte M, Rauluseviciute I, Jankevicius F, Lazutka JR, Jarmalaite S. DNA methylation profile of prostate cancer. Cell Symposia: Human Genomics. Singapore, Republic of Singapore. November 8-10, 2015.

6. Mackoit A, **Daniunaite K**, Jarmalaite S. Analysis of prostate cancer epigenome: identifying new DNA methylation markers. 58th Scientific Conference for Students of Physics and Natural Sciences: Open Readings 2015. Vilnius, Lithuania. March 24-27, 2015.

7. **Daniunaite K**, Unguryte A, Bernotiene E, Lazutka JR, Jarmalaite S. Genome-wide DNA methylation profile of mesenchymal stem cells and cancer. XIIIth International Conference of Lithuanian Biochemical Society: 50th Anniversary of FEBS. Birstonas, Lithuania. June 18-20, 2014.

Other presentations not directly related to the topic of doctoral dissertation:

- Oral presentations:

1. Kubiliute R, **Daniunaite K**, Jarmalaite S. Analysis of chemoresistance development in cancer cells. Biofuture: The Perspectives of Natural and Life Sciences. Vilnius, Lithuania. December 7, 2016.

2. Stuopelyte K, Stankevicius V, **Daniunaite K**, Jankevicius F, Jarmalaite S. MiRNA analysis in urine sediments for prostate cancer detection. Biofuture: The Perspectives of Natural and Life Sciences. Vilnius, Lithuania. December 9, 2014.

3. Jarmalaite S, **Daniunaite K**, Misgirdaite R, Serenaite I, Unguryte A, Bernotiene E. Epigenetic and genetic profile of stemness and cancer. XIIIth International Conference of Lithuanian Biochemical Society: 50th Anniversary of FEBS. Birstonas, Lithuania. June 18-20, 2014.

4. **Daniunaite K**, Matjosaitis K, Misgirdaite R, Unguryte A, Bernotiene E, Jarmalaite S. MiRNA expression profile in human mesenchymal stem cells of osteogenic and adipogenic differentiation. Biofuture: The Perspectives of Life and Natural Sciences. Vilnius, Lithuania. December 5, 2012. [award for best presentation]

- Poster presentations:

1. Kubiliute R, **Daniunaite K**, Daugelavicius R, Jarmalaite S. Development of chemoresistance: *ABCB1* methylation and epithelial-to-mesenchymal

transition as the key players. XIV International Conference of the Lithuanian Biochemical Society. Druskininkai, Lithuania. June 28-30, 2016.

2. Misgirdaite R, **Daniunaite K**, Bernotiene E, Jarmalaite S. MiRNA expression profile of osteogenically differentiated human mesenchymal stem cells. XIV International Conference of the Lithuanian Biochemical Society. Druskininkai, Lithuania. June 28-30, 2016.

3. Kubiliute R, **Daniunaite K**, Jarmalaite S. Epithelial-to-mesenchymal transition contributes to doxorubicin resistance in MX-1 breast cancer cell lines. The COINS International Conference of Natural and Life Sciences. Vilnius, Lithuania. February 29 – March 3, 2016.

4. Misgirdaite R, **Daniunaite K**, Jarmalaite S. MiRNA expression profile of osteogenically differentiated human mesenchymal stem cells. The COINS International Conference of Natural and Life Sciences. Vilnius, Lithuania. February 29 – March 3, 2016.

5. Stuopelyte K, **Daniunaite K**, Jankevicius F, Jarmalaite S. Detecting prostate cancer: urine-circulating miRNAs. The COINS International Conference of Natural and Life Sciences. Vilnius, Lithuania. February 29 – March 3, 2016.

6. Demidenko R, **Daniunaite K**, Laurinavicius A, Lazutka JR, Jankevicius F, Jarmalaite S. Gene expression profile of recurrent prostate cancer. Prostate Cancer World Congress. Cairns, Australia. August 17-21, 2015.

7. Stuopelyte K, **Daniunaite K**, Jankevicius F, Lazutka JR, Jarmalaite S. Urinary miRNAs for early detection of prostate cancer. CLEPSO Clinical Epigenetics Society 5th International Meeting. Dusseldorf, Germany. March 5-6, 2015.

8. Kubiliute R, **Daniunaite K**, Sulskyte I, Daugelavicius R, Jarmalaite S. The alteration of *ABCB1* and *ABCC1* expression in lipophilic agents induced resistant MX-1 and CX-1 cancer cell lines. The COINS International Conference of Natural and Life Sciences. Vilnius, Lithuania. March 3-7, 2015.

9. Misgirdaite R, **Daniunaite K**, Jarmalaite S. Osteogenic differentiation-related changes in miRNA expression profile of human mesenchymal stem cells. The COINS International Conference of Natural and Life Sciences. Vilnius, Lithuania. March 3-7, 2015.

10. Razanauskas D, **Daniunaite K**, Demidenko R, Jarmalaite S. *TMPRSS2-ERG* fusion transcript expression and *ABCA1* gene promoter methylation in prostate cancer. The COINS International Conference of Natural and Life Sciences. Vilnius, Lithuania. March 3-7, 2015.

11. **Daniunaite K**, Stuopelyte K, Stankevicius V, Laurinavicius A, Jankevicius F, Jarmalaite S. Analysis of miRNA expression in prostate tumors and urine samples. FEBS-EMBO 2014 Conference. Paris, France. August 30 – September 4, 2014.

12. Demidenko R, **Daniunaite K**, Laurinavicius A, Jankevicius F, Jarmalaite S, Lazutka JR. Investigation of potential biomarkers of biochemically recurrent prostate cancer. FEBS-EMBO 2014 Conference. Paris, France. August 30 – September 4, 2014.

13. Jarmalaite S, **Daniunaite K**, Misgirdaite R, Serenaite I, Unguryte A, Bernotiene E. Epigenetic and genetic profile of human mesenchymal stem cells from adipose and synovial tissue. FEBS-EMBO 2014 Conference. Paris, France. August 30 – September 4, 2014.

14. Kapustina Z, Valaitiene G, **Daniunaite K**, Jarmalaite S, Jankevicius F, Laurinavicius F, Lazutka JR. Analysis of influence of single nucleotide polymorphisms in *RNASEL*, *LEPR*, *CRY1*, *IL4*, *CHI3L2*, and *AKAP14* genes on clinical course of prostate cancer. XIIIth International Conference of Lithuanian Biochemical Society: 50th Anniversary of FEBS. Birstonas, Lithuania. June 18-20, 2014.

15. Kubiliute R, **Daniunaite K**, Demidenko R, Jarmalaite S, Daugelavicius R. Influence of lipophilic agents on the formation of multidrug resistance in cell cultures. XIIIth International Conference of Lithuanian Biochemical Society: 50th Anniversary of FEBS. Birstonas, Lithuania. June 18-20, 2014.

16. Stuopelyte K, **Daniunaite K**, Stankevicius V, Laurinavicius V, Jankevicius F, Lazutka JR, Jarmalaite S. MiRNA expression analysis in prostate tumors and urine sediments. XIIIth International Conference of Lithuanian Biochemical Society: 50th Anniversary of FEBS. Birstonas, Lithuania. June 18-20, 2014.

17. Stuopelyte K, **Daniunaite K**, Stankevicius V, Jankevicius F, Lazutks JR, Jarmalaite S. MiRNA expression analysis in urine sediments of prostate cancer patients. Urinomics – 1st International Conference on Urine Omics. Caparica, Portugal. September 9-11, 2013.

18. **Daniunaite K**, Misgirdaite R, Matjosaitis K, Stuopelyte K, Bernotiene E, Unguryte A, Jarmalaite S. Differences of miRNA expression and DNA modification profile in undifferentiated and differentiated mesenchymal stem cells. qPCR & NGS 2013 Event: Next Generation Thinking in Molecular Diagnostics. Freising, Germany. March 18-22, 2013.

19. **Daniunaite K**, Girdziusaite A, Matjosaitis K, Valiuniene S, Bernotiene E, Unguryte A, Jarmalaite S. DNA methylation stability of cultured adipose tissue and synovial fluid-derived stem cells. 5th Baltic Congress of Genetics. Kaunas, Lithuania. October 19-22, 2012.

20. Stuopelyte K, **Daniunaite K**, Laurinaviciene A, Ostapenko V, Jarmalaite S. High-resolution melting-based quantitative analysis of *RASSF1* methylation in breast cancer. 5th Baltic Congress of Genetics. Kaunas, Lithuania. October 19-22, 2012.

FINANCIAL SUPPORT

The research was supported by several grants from the Research Council of Lithuania (RCL). Initial sample collection and protocol development was partially supported by grant No. C-03/2007-09 from the former Lithuanian State Science and Studies Foundation. Sample collection was also partially supported by RCL grant No. MIP-075/2010. Tumor suppressor gene analysis in prostate tissues was supported by European Social Fund under National Integrated Programme Biotechnology and Biopharmacy grant No. VP1-3.1-ŠMM-08-K-01-005/KS-560000-1757. Biomarker analysis in body fluids was funded by grant MIP-109 from RCL. Analysis of novel PCa biomarkers was funded by grants LIG-14/2012 and SEN-9/2016, and partially supported by scholarship for Promotion of Student Scientific Activities (grant No. VP1-3.1-SMM-01-V-02-003) from RCL. The author also received multiple Student Research Fellowship awards and doctoral scholarships (2013-2016) from RCL. Expenses, related to participation at international conferences, were partially covered by COST Action CM1406 (the conference in Groningen, Netherlands) and RCL grant for Scientific Mission of Doctoral Students (the conference in Singapore, Republic of Singapure).

CURRICULUM VITAE

Name, Surname	Kristina Daniūnaitė
General information	
Date of birth	24th December, 1987
Main workplace	Human Genome Research Group, Institute of Biosciences, Life Sciences Center, Vilnius University; Sauletekio Ave. 7, LT-10257, Vilnius, Lithuania; Tel. +370 5 223 4441 E-mail: Kristina.Daniunaite@gf.vu.lt
Second workplace	Genetic Diagnostics Center, National Cancer Institute; Baublio St. 3b, LT-08406, Vilnius, Lithuania; Tel. + 370 5 219 0184 E-mail: Kristina.Daniunaite@nvi.lt
Personal contacts	Tel. +370 614 20818 E-mail: vailaomalyn@yahoo.com
Education	
2012-2016	Doctoral studies (Biology, 01 B); Institute of Biosciences (former Faculty of Natural Sciences), Life Sciences Center, Vilnius University.
2010-2012	Master of Genetics (*Magna Cum Laude*), Genetics study program (621C40001); former Faculty of Natural Sciences, Vilnius University. [Lithuanian Academy of Sciences award for best master thesis]
2006-2010	Bachelor of Biology (*Cum Laude*), Molecular Biology study program (61201B105); former Faculty of Natural Sciences, Vilnius University.
1994-2006	Ukmerge Jonas Basanavicius Gymnasium (former Ukmerge Jonas Basanavicius High School)
Training courses & schools	
2014	"Proton System. Hands-on Training and Data Analysis" by Life Technologies; Vilnius, Lithuania
2013	"NGS Data Analysis Workshop" by Genomatix; Freising, Germany
2012	"Applications of Microarray Technology" by Agilent Technologies; Waldbronn, Germany
2011	Course of Laboratory Animal Science at Vilnius University (FELASA C certificate); Vilnius, Lithuania
2011	"qPCR data analysis: Biostatistics & Expression Profiling" by TATAA Biocenter; Freising, Germany
2009	"Real-Time PCR School" by Applied Biosystems; Vilnius, Lithuania
Work experience	
Since 2016	Junior researcher at Genetic Diagnostics Center, National Cancer Institute
2015-2016	Junior researcher at the Faculty of Natural Sciences, Vytautas Magnus University
Since 2012	Junior researcher at Institute of Biosciences (former Faculty of Natural Sciences), Life Sciences Center, Vilnius University
2011-2012	Senior specialist at the former Faculty of Natural Sciences, Vilnius University
2009-2011	Laboratory technician at the former Faculty of Natural Sciences, Vilnius University
Research projects	
Since 2009	Participation in 7 research projects, 1 ongoing
Membership	
Since 2013	Lithuanian Biochemical Society and FEBS

ACKNOWLEDGEMENTS

I would like to thank my supervisor prof. Sonata Jarmalaite for the provided opportunity to perform doctoral studies. Thank you for your accurate insights, long and fruitful discussions, valuable advices, critical comments, all the possibilities to present research results at international events, and support during the studies.

I am grateful to prof. Juozas Rimantas Lazutka, prof. Feliksas Jankevicius, and all the people involved in the implementation of the projects, whose united effort led to the accomplishment of ambitious objectives.

I would like to thank Neringa Kalinauskaite for the guidance at the very beginning of the experimental work at the labs. I am also grateful to Monika Dubikaityte, Ieva Rauluseviciute, Deividas Razanauskas, Agnieska Mackoit, and all other students, who participated in DNA methylation analysis of prostate cancer.

I thank the constantly changing Human Genome Research group and all the co-workers of the former Department of Botany and Genetics for sustaining that pleasant working environment, all the technical and moral support, whenever one or another was needed, and, of course, unforgettable cheerful moments of respite.

My special thanks go to my Biology teacher Giedrius Svitra for all those inspirational lessons full of imaginative teaching methods, which altogether made great influence in chosing the path of a scientist. I am also very grateful to dr. Vida Bendikiene, who was probably the first person at the university to notice the potential in me and, both figuratively and literally speaking, showed the way to the research lab.

I also thank all the patients and their family members for participation in this study.

I extend acknowledgement to all the members of The Cancer Genome Atlas Project, Gene Ontology Consortium, and GSEA/ MSigDB workgroup, who have worked together to create these resources.

REFERENCES

Publications:

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015; 136(5): E359-E386.
2. Shen MM, Abate-Shen C. Molecular genetics of prostate cancer: new prospects for old challenges. Genes Dev. 2010; 24(18): 1967-2000.
3. Kim JH, Dhanasekaran SM, Prensner JR, Cao X, Robinson D, Kalyana-Sundaram S, et al. Deep sequencing reveals distinct patterns of DNA methylation in prostate cancer. Genome Res. 2011; 21(7): 1028-1041.
4. The Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. Cell. 2015; 163(4): 1011-1025.
5. Kron K, Pethe V, Briollais L, Sadikovic B, Ozcelik H, Sunderji A, et al. Discovery of novel hypermethylated genes in prostate cancer using genomic CpG island microarrays. PLos One. 2009; 4(3): e4830.
6. Kristensen H, Thomsen AR, Haldrup C, Dyrskjot L, Hoyer S, Borre M, et al. Novel diagnostic and prognostic classifiers for prostate cancer identified by genome-wide microRNA profiling. Oncotarget. 2016; 7(21): 30760-30771.
7. Kobayashi Y, Absher DM, Gulzar ZG, Young SR, McKenney JK, Peehl DM, et al. DNA methylation profiling reveals novel biomarkers and important roles for DNA methyltransferases in prostate cancer. Genome Res. 2011; 21(7): 1017-1027.
8. Mahapatra S, Klee EW, Young CY, Sun Z, Jimenez RE, Klee GG, et al. Global methylation profiling for risk prediction of prostate cancer. Clin Cancer Res. 2012; 18(10): 2882-2895.
9. Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, Sivachenko AY, et al. The genomic complexity of primary human prostate cancer. Nature. 2011; 470(7333): 214-220.
10. Ross AE, D'Amico AV, Freedland SJ. Which, when and why? Rational use of tissue-based molecular testing in localized prostate cancer. Prostate Cancer Prostatic Dis. 2016; 19(1): 1-6.
11. Park JY. Promoter hypermethylation as a biomarker in prostate adenocarcinoma. Methods Mol Biol. 2015; 1238: 607-625.
12. American Cancer Society. Cancer Facts & Figures 2016. Report. Atlanta: American Cancer Society; 2016.
13. Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends--an update. Cancer Epidemiol Biomarkers Prev. 2016; 25(1): 16-27.
14. Lin DW, Porter M, Montgomery B. Treatment and survival outcomes in young men diagnosed with prostate cancer: a population based cohort study. Cancer. 2009; 115(13): 2863-2871.
15. Markozannes G, Tzoulaki I, Karli D, Evangelou E, Ntzani E, Gunter MJ, et al. Diet, body size, physical activity and risk of prostate cancer: an umbrella review of the evidence. Eur J Cancer. 2016; 69: 61-69.
16. Fine SW, Reuter VE. Anatomy of the prostate revisited: implications for prostate. Histopathology. 2012; 60(1): 142-152.
17. Andreoiu MD, Cheng L. Multifocal prostate cancer: biologic, prognostic, and therapeutic implications. Hum Pathol. 2010; 41(6): 781-793.
18. Liu AY, True LD. Characterization of prostate cell types by CD cell surface molecules. Am J Pathol. 2002; 160(1): 37-43.

19. Bok RA, Small EJ. Bloodborne biomolecular markers in prostate cancer development and progression. Nat Rev Cancer. 2002; 2(12): 918-926.
20. Goldstein AS, Huang J, Guo C, Garraway IP, Witte ON. Identification of a cell of origin for human prostate cancer. Science. 2010; 329(5991): 568-571.
21. Park JW, Lee JK, Phillips JW, Huang P, Cheng D, Huang J, et al. Prostate epithelial cell of origin determines cancer differentiation state in an organoid transformation assay. Proc Natl Acad Sci U S A. 2016; 113(16): 4482-4487.
22. Stoyanova T, Cooper AR, Drake JM, Liu X, Armstrong AJ, Pienta KJ, et al. Prostate cancer originating in basal cells progresses to adenocarcinoma propagated by luminal-like cells. Proc Natl Acad Sci U S A. 2013; 110(50): 20111-20116.
23. Attard G, Parker C, Eeles RA, Schroder F, Tomlins SA, Tannock I, et al. Prostate cancer. Lancet. 2016; 387(10013): 70-82.
24. Serenaite I, Daniunaite K, Jankevicius F, Laurinavicius A, Petroska D, Lazutka JR, et al. Heterogeneity of DNA methylation in multifocal prostate cancer. Virchows Arch. 2015; 466(1): 53-59.
25. Cooper CS, Eeles R, Wedge DC, Van Loo P, Gundem G, Alexandrov LB, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. Nat Genet. 2015; 47(4): 367-372.
26. Lindberg J, Klevebring D, Liu W, Neiman M, Xu J, Wiklund P, et al. Exome sequencing of prostate cancer supports the hypothesis of independent tumour. Eur Urol. 2013; 63(2): 347-353.
27. Abate-Shen C, Shen MM. Molecular genetics of prostate cancer. Genes Dev. 2000; 14(19): 2410-2432.
28. Bostwick DG, Cheng L. Precursors of prostate cancer. Histopathology. 2012; 60(1): 4-27.
29. De Marco AM, Haffner MC, Lotan TL, Yegnasubramanian S, Nelson WG. Premalignancy in prostate cancer: rethinking what we know. Cancer Prev Res (Phila). 2016; 9(8): 648-656.
30. Packer JR, Maitland NJ. The molecular and cellular origin of human prostate cancer. Biochim Biophys Acta. 2016; 1863(6): 1238-1260.
31. Mundbjerg K, Chopra S, Alemozaffar M, Duymich C, Lakshminarasimhan R, Nichols PW, et al. Identifying aggressive prostate cancer foci using a DNA methylation classifier. Genome Biol. 2017; 18(1): 3.
32. Schmidt H, DeAngelis G, Eltze E, Gockel I, Semjonow A, Brandt B. Asynchronous growth of prostate cancer is reflected by circulating tumor cells delivered from distinct, even small foci, harboring loss of heterozygosity of the *PTEN* gene. Cancer Res. 2006; 66(18): 8959-8965.
33. Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JM, Papaemmanuil E, et al. The evolutionary history of lethal metastatic prostate cancer. Nature. 2015; 520(7547): 353-357.
34. Merrimen JL, Evans AJ, Srigley JR. Preneoplasia in the prostate gland with emphasis on high grade prostatic intraepithelial neoplasia. Pathology. 2013; 45(3): 251-263.
35. Tomlins SA, Laxman B, Varanbally S, Cao X, Yu J, Helgeson BE, et al. Role of *TMPRSS2-ERG* gene fusion in prostate cancer. 2008; 10(2): 177-188.
36. Nakayama M, Bennett CJ, Hicks JL, Epstein JI, Platz EA, Nelson WG, et al. Hypermethylation of the human glutathione S-transferase-pi gene (*GSTP1*) CpG island in present in a subset of proliferative inflammatory atrophy lesions but not in normal or hyperplastic epithelium of the prostate. Am J Pathol. 2003; 163(3): 923-933.
37. Montironi R, Diamanti L, Thomson D, Bartels HG, Bartels PH. Analysis of the capillary architecture in the precursors of prostate cancer: recent findings and new concepts. Eur Urol. 1996; 30(2): 191-200.

38. Lane BR, Zippe CD, Abouassaly R, Schoenfield L, Magi-Galluzzi C, Jones JS. Saturation technique does not decrease cancer detection during followup after initial prostate biopsy. J Urol. 2008; 179(5): 1746-1750.
39. Humphrey PA. Atypical adenomatous hyperplasia (adenosis) of the prostate. J Urol. 2012; 188(6): 2371-2372.
40. Choi YJ, Kim JK, Kim N, Kim KW, Choi EK, Cho KS. Functional MR imaging of prostate cancer. Radiographics. 2007; 27(1): 63-75.
41. Birnie R, Bryce SD, Roome C, Dussupt V, Droop A, Lang SH, et al. Gene expression profiling of human prostate cancer stem cells reveals a pro-inflammatory phenotype and the importance of extracellular matrix interactions. Genome Biol. 2008; 9(5): R83.
42. Nonn L, Ananthanarayanan V, Gann PH. Evidence for field cancerization of the prostate. Prostate. 2009; 69(13): 1470-1479.
43. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer. 1953; 6(5): 963-968.
44. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter‘s concept of field cancerization: evidence and clinical implications. Cancer Res. 2003; 63(8): 1727-1730.
45. Braakhuis BJ, Leemans CR, Brakenhoff RH. Using tissue adjacent to carcinoma as a normal control: an obvious but questionable practice. J Pathol. 2004; 203(2): 620-621.
46. Chandran UR, Dhir R, Ma C, Michalopoulos G, Becich M. Differences in gene expression in prostate cancer, normal appearing prostate tissue adjacent to cancer and prostate tissue from cancer free organ donors. BMC Cancer. 2005; 5: 45.
47. Haaland CM, Heaphy CM, Butler KS, Fischer EG, Griffith JK, Bisoffi M. Differential gene expression in tumor adjacent histologically normal prostatic tissue indicates field cancerization. Int J Oncol. 2009; 35(3): 537-546.
48. Jones AC, Trujillo KA, Phillips GK, Fleet TM, Murton JK, Severns V, et al. Early growth response 1 and fatty acid synthase expression is altered in tumor adjacent prostate tissue and indicates field cancerization. Prostate. 2012; 72(11): 1159-1170.
49. Troyer DA, Lucia MS, de Bruine AP, Mendez-Meza R, Baldewijns MM, Dunscomb N, et al. Prostate cancer detected by methylated gene markers in histopathologically cancer-negative tissues from men with subsequent positive biopsies. Cancer Epidemiol Biomarkers Prev. 2009; 18(10): 2717-2722.
50. Mehrotra J, Varde S, Wang H, Chiu H, Vargo J, Gray K, et al. Quantitative, spatial resolution of the epigenetic field effect in prostate cancer. Prostate. 2008; 68(2): 152-160.
51. Trock BJ, Brotzman MJ, Mangold LA, Bigley JW, Epstein JI, McLeod D, et al. Evaluation of *GSTP1* and *APC* methylation as indicators for repeat biopsy in a high-risk cohort of men with negative initial prostate biopsies. BJU Int. 2012; 110(1): 56-62.
52. Moller M, Strand SH, Mundbjerg K, Liang G, Gill I, Haldrup C, et al. Heterogeneous patterns of DNA methylation-based field effects in histologically normal prostate tissue from cancer patients. Sci Rep. 2017; 7: 40636.
53. Luo JH, Ding Y, Chen R, Michalopoulos G, Nelson J, Tseng G, et al. Genome-wide methylation analysis of prostate tissues reveals global methylation patterns of prostate cancer. Am J Pathol. 2013; 182(6): 2028-2036.
54. Yang B, Bhusari S, Kueck J, Weeratunga P, Wagner J, Leverson G, et al. Methylation profiling defines an extensive field defect in histologically normal prostate tissues associated with prostate cancer. Neoplasia. 2013; 15(4): 399-408.

55. Stewart GD, Van Neste L, Delvenne P, Delree P, Delga A, McNeill SA, et al. Clinical utility of an epigenetic assay to detect occult prostate cancer in histopathologically negative biopsies: results of the MATLOC study. J Urol. 2013; 189(3): 1110-1116.
56. Partin AW, Van Neste L, Klein EA, Marks LS, Gee JR, Troyer DA, et al. Clinical validation of an epigenetic assay to predict negative histopathological results in repeat prostate biopsies. J Urol. 2014; 192(4): 1081-1087.
57. Gittelman MC, Hertzman B, Bailen J, Williams T, Koziol I, Henderson RJ, et al. *PCA3* molecular urine test as a predictor of repeat prostate biopsy outcome in men with previous negative biopsies: a prospective multicenter clinical study. J Urol. 2013; 190(1): 64-69.
58. Richiardi L, Fiano V, Grasso C, Zugna D, Delsedime L, Gillio-Tos A, et al. Methylation of *APC* and *GSTP1* in non-neoplastic tissue adjacent to prostate tumour and mortality from prostate cancer. PLoS One. 2013; 8(7): e68162.
59. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol. 2017; 71(4): 618-629.
60. Hayes JH, Barry MJ. Screening for prostate cancer with the prostate-specific antigen test: a review of current evidence. JAMA. 2014; 311(11): 1143-1149.
61. Sartor AO, Hricak H, Wheeler TM, Coleman J, Penson DF, Carroll PR, et al. Evaluating localized prostate cancer and identifying candidates for focal therapy. Urology. 2008; 72(6 Suppl): S12-S24.
62. Rosser CJ, Parker A. Re: The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? J Urol. 2005; 174(3): 1154-1155.
63. Gulati R, Inoue LY, Gore JL, Katcher J, Etzioni R. Individualized estimates of overdiagnosis in screen-detected prostate cancer. J Natl Cancer Inst. 2014; 106(2): djt367.
64. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. N Engl J Med. 2004; 350(22): 2239-2246.
65. Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate cancer. J Clin Oncol. 2010; 28(7): 1117-1123.
66. Guazzoni G, Nava L, Lazzeri M, Scattoni V, Lughezzani G, Maccagnano C, et al. Prostate-specific antigen (PSA) isoform p2PSA significantly improves the prediction of prostate cancer at initial extended prostate biopsies in patients with total PSA between 2.0 and 10 ng/ml: results of a prospective study in a clinical setting. Eur Urol. 2011; 60(2): 214-222.
67. Fossati N, Buffi NM, Haese A, Stephan C, Larcher A, McNicholas T, et al. Preoperative prostate-specific antigen isoform p2PSA and its derivatives, %p2PSA and prostate health index, predict pathologic outcomes in patients undergoing radical prostatectomy for prostate cancer: results from a multicentric European prospective study. Eur Urol. 2015; 68(1): 132-138.
68. Parekh DJ, Punnen S, Sjoberg DD, Asroff SW, Bailen JL, Cochran JS, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. Eur Urol. 2015; 68(3): 464-470.
69. Robinson K, Creed J, Reguly B, Powell C, Wittock R, Klein D, et al. Accurate prediction of repeat prostate biopsy outcomes by a mitochondrial DNA deletion assay. Prostate Cancer Prostatic Dis. 2010; 13(2): 126-131.
70. Blume-Jensen P, Berman DM, Rimm DL, Shipitsin M, Putzi M, Nifong TP, et al. Development and clinical validation of an *in situ* biopsy-based multimarker assay for risk stratification in prostate cancer. Clin Cancer Res. 2015; 21(11): 2591-2600.
71. Shore N, Concepcion R, Saltzstein D, Lucia MS, van Breda A, Welbourn W, et al. Clinical utility of a biopsy-based cell cycle gene expression assay in localized prostate cancer. Curr Med Res Opin. 2014; 30(4): 547-553.

72. Klein EA, Cooperberg MR, Magi-Galluzzi C, Simko JP, Falzarano SM, Maddala T, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. Eur Urol. 2014; 66(3): 550-560.
73. Erho N, Crisan A, Vergara IA, Ghadessi M, Buerki C, Bergstralh EJ, et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. PLoS One. 2013; 8(6): e66855.
74. de Bono JS, Scher HI, Montgomery RB, Parker C, Miller MC, Tissing H, et al. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res. 2008; 14(19): 6302-6309.
75. McKiernan J, Donovan MJ, O'Neill V, Bentink S, Noerholm M, Belzer S, et al. A novel urine exosome gene expression assay to predict high-grade prostate cancer at initial biopsy. JAMA Oncol. 2016; 2(7): 882-889.
76. Beahrs OH, Henson DE, Hutter RVP, Kennedy BJ. AJCC manual for staging of cancer. 4th ed. Philadelphia: Lippincott; 1992.
77. Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol. 1974; 111(1): 58-64.
78. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason score. Eur Urol. 2016; 69(3): 428-435.
79. Epstein JI, Allsbrook WCJ, Amin MB, Egevad LL, ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason rrading of prostatic carcinoma. Am J Surg Pathol. 2005; 29(9): 1228-1242.
80. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, et al. The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol. 2016; 40(2): 244-252.
81. Epstein JI, Amin M, Boccon-Gibod L, Egevad L, Humphrey PA, Mikuz G, et al. Prognostic factors and reporting of prostate carcinoma in radical prostatectomy and pelvic lymphadenectomy specimens. Scand J Urol Nephrol Suppl. 2005; 216: 34-63.
82. Marks RA, Kock MO, Lopez-Beltran A, Montironi R, Juliar BE, Cheng L. The relationship between the extent of surgical margin positivity and prostate specific antigen recurrence in radical prostatectomy specimens. Hum Pathol. 2007; 38(8): 1207-1211.
83. Priemer DS, Montironi R, Wang L, Williamson SR, Lopez-Beltran A, Cheng L. Neuroendocrine tumors of the prostate: emerging insights from molecular data and updates to the 2016 World Health Organization classification. Endocr Pathol. 2016; 27(2): 123-135.
84. Cindolo L, Cantile M, Vacherot F, Terry S, de la Taille A. Neuroendocrine differentiation in prostate cancer: from lab to bedside. Urol Int. 2007; 79(4): 287-296.
85. Hu CD, Choo R, Huang J. Neuroendocrine differentiation in prostate cancer: a mechanism of radioresistance and treatment failure. Front Oncol. 2015; 5: 90.
86. Davis KM, Kelly SP, Luta G, Tomko C, Miller AB, Taylor KL. The association of long-term treatment-related side effects with cancer-specific and general quality of life among prostate cancer survivors. Urology. 2014; 84(2): 300-306.
87. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001; 98(19): 10869-10874.
88. Kgatle MM, Kalla AA, Islam MM, Satheke M, Moorad R. Prostate cancer: epigenetic alterations, risk factors, and therapy. Prostate Cancer. 2016; 2016: 5653862.

89. Antonarakis ES, Feng Z, Trock BJ, Humphreys EB, Carducci MA, Partin AW, et al. The natural history of metastatic progression in men with prostate-specific antigen recurrence after radical prostatectomy: long-term follow-up. BJU Int. 2012; 109(1): 32-39.
90. Partin AW, Pearson JD, Landis PK, Carter HB, Pound CR, Clemens JQ, et al. Evaluation of serum prostate-specific antigen velocity after radical prostatectomy to distinguish local recurrence from distant metastases. Urology. 1994; 43(5): 649-659.
91. Graham L, Schweizer MT. Targeting persistent androgen receptor signaling in castration-resistant prostate cancer. Med Oncol. 2016; 33(5): 44.
92. Katzenwadel A, Wolf Androgen deprivation of prostate cancer: Leading to a therapeutic dead end. Cancer Lett. 2015; 367(1): 12-17.
93. Coutinho I, Day TK, Tilley WD, Selth LA. Androgen receptor signaling in castration-resistant prostate cancer: a lesson in persistence. Endocr Relat Cancer. 2016; 23(12): T179-T197.
94. Shafi AA, Yen AE, Weigel NL. Androgen receptors in hormone-dependent and castration-resistant prostate cancer. Pharmacol Ther. 2013; 140(3): 223-238.
95. de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and increased survival in metastatic prostate cancer. N Engl J Med. 2011; 364(21): 1995-2005.
96. Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, Penson DF, et al. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. N Engl J Med. 2010; 363(5): 411-422.
97. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. JAMA. 1999; 281(17): 1591-1597.
98. Rouviere O, Vitry T, Lyonnet D. Imaging of prostate cancer local recurrences: why and how? Eur Radiol. 2010; 20(5): 1254-1266.
99. Bubendorf L, Schopfer A, Wagner U, Sauter G, Moch H, Willi N, et al. Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. Hum Pathol. 2000; 31(5): 578-583.
100. Manca P, Pantano F, Iuliani M, Ribelli G, De Lisi D, Danesi R, et al. Determinants of bone specific metastasis in prostate cancer. Crit Rev Oncol Hematol. 2017; 112: 59-66.
101. Bell KJL, Del Mar C, Wright G, Dickinson J, Glasziou P. Prevalence of incidental prostate cancer: a systematic review of autopsy studies. Int J Cancer. 2015; 137(7): 1749-1757.
102. Freedland SJ, Presti JCJ, Kane CJ, Aronson WJ, Terris MK, Dorey F, et al. Do younger men have better biochemical outcomes after radical prostatectomy? Urology. 2004; 63(3): 518-522.
103. Kimura T. East meets West: ethnic differences in prostate cancer epidemiology between East Asians and Caucasians. Chin J Cancer. 2012; 31(9): 421-429.
104. McGinley KF, Tay KJ, Moul WJ. Prostate cancer in men of African origin. Nat Rev Urol. 2016; 13(2): 99-107.
105. Resnick MJ, Canter DJ, Guzzo TJ, Brucker BM, Bergey M, Sonnad SS, et al. Does race affect postoperative outcomes in patients with low-risk prostate cancer who undergo radical prostatectomy? Urology. 2009; 73(3): 620-623.
106. Bensen JT, Xu Z, Smith GJ, Mohler JL, Fontham ET, Taylor JA. Genetic polymorphism and prostate cancer aggressiveness: a case-only study of 1,536 GWAS and candidate SNPs in African-Americans and European-Americans. Prostate. 2013; 73(1): 11-22.
107. Giri VN, Beebe-Dimmer JL. Familial prostate cancer. Semin Oncol. 2016; 43(5): 560-565.
108. Gianferrari L, Arrigoni G, Cresseri A, Lovati G, Morganti G. [Genetic and clinico-statistical research on neoplasms of the prostate]. Acta Genet Med Gemellol (Roma). 1956; 5(2): 224-233.
109. Zeegers MP, Jellema A, Ostrer H. Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. Cancer. 2003; 97(8): 1894-1903.

110. Johns LE, Houlston RS. A systematic review and meta-analysis of familial prostate cancer risk. BJU International. 2003; 91(9): 789-794.

111. Negri E, Pelucchi C, Talamini R, Montella M, Gallus S, Bosetti C, et al. Family history of cancer and the risk of prostate cancer and benign prostatic hyperplasia. Int J Cancer. 2005; 114(4): 648-652.

112. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer – analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000; 343(2): 78-85.

113. Lynch HT, Kosoko-Lasaki O, Leslie SW, Rendell M, Shaw T, Snyder C, et al. Screening for familial and hereditary prostate cancer. Int J Cancer. 2016; 138(11): 2579-2591.

114. Smith JR, Freije D, Carpten JD, Gronberg H, Xu J, Isaacs SD, et al. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. Science. 1996; 274(5291): 1371-1374.

115. Barbieri CE, Bangma CH, Bjartell A, Catto JWF, Culig Z, Gronberg H, et al. The mutational landscape of prostate cancer. Eur Urol. 2013; 64(4): 567-576.

116. Cavanagh H, Rogers KMA. The role of *BRCA1* and *BRCA2* mutations in prostate, pancreatic and stomach cancers. Hered Cancer Clin Pract. 2015; 13(1): 1-7.

117. Sun J, Kader AK, Hsu FC, Kim ST, Zhu Y, Turner AR, et al. Inherited genetic markers discovered to date are able to identify a significant number of men at considerably elevated risk for prostate cancer. Prostate. 2011; 71(4): 421-430.

118. Ryan S, Jenkins MA, Win AK. Risk of prostate cancer in lynch syndrome: a systematic review and meta-analysis. Cancer Epidemiol Biomarkers Prev. 2014; 23(3): 437-449.

119. LeMasters GK, Genaidy AM, Succop P, Deddens J, Sobeih T, Barriera-Viruet H, et al. Cancer risk among firefighters: a review and meta-analysis of 32 studies. J Occup Environ Med. 2006; 48(11): 1189-1202.

120. Caini S, Gandini S, Dudas M, Bremer V, Severi E, Gherasim A. Sexually transmitted infections and prostate cancer risk: a systematic review and meta-analysis. Cancer Epidemiol. 2014; 38(4): 329-338.

121. Liu LH, Kang R, He J, Zhao SK, Li SK, Wan SP, et al. Vasectomy and risk of prostate cancer: a systematic review and meta-analysis of cohort studies. Andrology. 2015; 3(4): 643-649.

122. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell. 2010; 18(1): 11-22.

123. Matsumoto T, Sakari M, Okada M, Yokoyama A, Takahashi S, Kouzmenko A, et al. The androgen receptor in health and disease. Annu Rev Physiol. 2013; 75: 201-224.

124. Lorente D, De Bono JS. Molecular alterations and emerging targets in castration resistant prostate cancer. Eur J Cancer. 2014; 50(4): 753-764.

125. Jin HJ, Kim J, Yu J. Androgen receptor genomic regulation. Transl Androl Urol. 2013; 2(3): 157-177.

126. Vaarala MH, Hirvikoski P, Kauppila S, Paavonen TK. Identification of androgen-regulated genes in human prostate. Mol Med Rep. 2012; 6(3): 466-472.

127. Coutinho-Camillo CM, Salaorni S, Sarkis AS, Nagai MA. Differentially expressed genes in the prostate cancer cell line LNCaP after exposure to androgen and anti-androgen. Cancer Genet Cytogenet. 2006; 166(2): 130-138.

128. Pihlajamaa P, Sahu B, Janne OA. Determinants of receptor- and tissue-specific actions in androgen signaling. Endocr Rev. 2015; 36(4): 357-384.

129. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of *TMPRSS2* and *ETS* transcription factor genes in prostate cancer. Science. 2005; 310(5748): 644-648.

130. Mani RS, Tomlins SA, Callahan K, Ghosh A, Nyati MK, Varambally S, et al. Induced chromosomal proximity and gene fusions in prostate cancer. Science. 2009; 326(5957): 1230.

131. Crea F, Sun L, Mai A, Chiang Y, Farrar W, Danesi R, et al. The emerging role of histone lysine demethylases in prostate cancer. Mol Cancer. 2012; 11: 52.

132. Gao L, Alumkal J. Epigenetic regulation of androgen receptor signaling in prostate cancer. Epigenetics. 2010; 5(2): 100-104.

133. Cai C, Yuan X, Balk SP. Androgen receptor epigenetics. Transl Androl Urol. 2013; 2(3): 148-157.

134. Ong CT, Corcer VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat Rev Genet. 2011; 12(4): 283-293.

135. Boyd LK, Mao X, Lu YJ. The complexity of prostate cancer: genomic alterations and heterogeneity. Nat Rev Urol. 2012; 9(11): 652-664.

136. Sabaliauskaite R, Jarmalaite S, Petroska D, Dasevicius D, Laurinavicius A, Jankevicius F, et al. Combined analysis of *TMPRSS2-ERG* and *TERT* for improved prognosis of biochemical recurrence in prostate cancer. Genes Chromosomes Cancer. 2012; 51(8): 781-791.

137. Tan MH, Li J, Xu HE, Melcher K, Yong EL. Androgen receptor: structure, role in prostate cancer and drug discovery. Acta Pharmacol Sin. 2015; 36(1): 3-23.

138. Rubin MA, Maher CA, Chinnaiyan AM. Common gene rearrangements in prostate cancer. J Clin Oncol. 2011; 29(27): 3659-3668.

139. Borno ST, Fischer A, Kerick M, Falth M, Laible M, Brase JC, et al. Genome-wide DNA methylation events in *TMPRSS2-ERG* fusion-negative prostate cancers implicate an *EZH2*-dependent mechanism with miR-26a hypermethylation. Cancer Discov. 2012; 2(11): 1024-1035.

140. Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A, et al. Aberrant *ERG* expression cooperates with loss of *PTEN* to promote cancer progression in the prostate. Nat Genet. 2009; 41(5): 619-624.

141. Chen M, Pratt CP, Zeeman ME, Schultz N, Taylor BS, O'Neill A, et al. Identification of *PHLPP1* as a tumor suppressor reveals the role of feedback activation in *PTEN*-mutant prostate cancer progression. Cancer Cell. 2011; 20(2): 173-186.

142. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, et al. Exome sequencing identifies recurrent *SPOP*, *FOXA1* and *MED12* mutations in prostate cancer. Nat Genet. 2012; 44(6): 685-689.

143. Fraser M, Sabelnykova VY, Yamaguchi TN, Heisler LE, Livingstone J, Huang V, et al. Genomic hallmarks of localized, non-indolent prostate cancer. Nature. 2017; 541(7637): 359-364.

144. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. Nature. 2012; 487(7406): 239-243.

145. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, et al. Punctuated evolution of prostate cancer genomes. Cell. 2013; 153(3): 666-677.

146. Xu K, Wu ZJ, Groner AC, He HH, Cai C, Lis RT, et al. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science. 2012; 338(6113): 1465-1469.

147. Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D, Stern HM, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. Nature. 2010; 466(7308): 869-873.

148. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. Cell. 2015; 161(5): 1215-1228.

149. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012; 13(7): 484-492.

150. Shen H, Laird PW. Interplay between the cancer genome and epigenome. Cell. 2013; 153(1): 38-55.

151. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: prospects and challenges. Trends Genet. 2014; 30(2): 75-84.

152. Strand SH, Orntoft TF, Sorensen KD. Prognostic DNA methylation markers for prostate cancer. Int J Mol Sci. 2014; 15(9): 16544-16576.

153. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. Cancer Res. 2001; 61(8): 3225-3229.

154. Yegnasubramanian S, Haffner MC, Zhang Y, Gurel B, Cornish TC, Wu Z, et al. DNA hypomethylation arises later in prostate cancer progression than CpG island hypermethylation and contributes to metastatic tumor heterogeneity. Cancer Res. 2008; 68(21): 8954-8967.

155. Van Neste L, Bigley J, Toll A, Otto G, Clark J, Delree P, et al. A tissue biopsy-based epigenetic multiplex PCR assay for prostate cancer detection. BMC Urol. 2012; 12: 16.

156. Henrique R, Jeronimo C, Teixeira MR, Hoque MO, Carvalho AL, Pais I, et al. Epigenetic heterogeneity of high-grade prostatic intraepithelial neoplasia: clues for clonal progression in prostate carcinogenesis. Mol Cancer Res. 2006; 4(1): 1-8.

157. Jeronimo C, Henrique R, Hoque MO, Mambo E, Ribeiro FR, Varzim G, et al. A quantitative promoter methylation profile of prostate cancer. Clin Cancer Res. 2004; 10(24): 8472-8478.

158. Kang GH, Lee S, Lee HJ, Hwang KS. Aberrant CpG island hypermethylation of multiple genes in prostate cancer and prostatic intraepithelial neoplasia. J Pathol. 2004; 202(2): 233-240.

159. Henrique R, Ribeiro FR, Fonseca D, Hoque MO, Carvalho AL, Costa VL, et al. High promoter methylation levels of *APC* predict poor prognosis in sextant biopsies from prostate cancer patients. Clin Cancer Res. 2007; 13(20): 6122-6129.

160. Rosenbaum E, Hoque MO, Cohen Y, Zahurak M, Eisenberger MA, Epstein JI, et al. Promoter hypermethylation as an independent prognostic factor for relapse in patients with prostate cancer following radical prostatectomy. Clin Cancer Res. 2005; 11(23): 8321-8325.

161. Richiardi L, Fiano V, Vizzini L, De Marco L, Delsedime L, Akre O, et al. Promoter methylation in *APC*, *RUNX3*, and *GSTP1* and mortality in prostate cancer patients. J Clin Oncol. 2009; 27(19): 3161-3168.

162. Ellinger J, Bastian PJ, Jurgan T, Biermann K, Kahl P, Heukamp LC, et al. CpG island hypermethylation at multiple gene sites in diagnosis and prognosis of prostate cancer. Urology. 2008; 71(1): 161-167.

163. Moritz R, Ellinger J, Nuhn P, Haese A, Muller SC, Graefen M, et al. DNA hypermethylation as a predictor of PSA recurrence in patients with low- and intermediate-grade prostate cancer. Anticancer Res. 2013; 33(12): 5249-5254.

164. Yegnasubramanian S, Kowalski J, Gonzalgo ML, Zahurak M, Piantadosi S, Walsh PC, et al. Hypermethylation of CpG islands in primary and metastatic human prostate cancer. Cancer Res. 2004; 64(6): 1975-1986.

165. Maruyama R, Toyooka S, Toyooka KO, Virmani AK, Zochbauer-Muller S, Farinas AJ, et al. Aberrant promoter methylation profile of prostate cancers and its relationship to clinicopathological features. Clin Cancer Res. 2002; 8(2): 514-519.

166. Florl AR, Steinhoff C, Muller M, Seifert HH, Hader C, Engers R, et al. Coordinate hypermethylation at specific genes in prostate carcinoma precedes *LINE-1* hypomethylation. Br J Cancer. 2004; 91(5): 985-994.

167. Yamanaka M, Watanabe M, Yamada Y, Takagi A, Murata T, Takahashi H, et al. Altered methylation of multiple genes in carcinogenesis of the prostate. Int J Cancer. 2003; 106(3): 382-387.

168. Enokida H, Shiina H, Urakami S, Igawa M, Ogishima T, Li LC, et al. Multigene methylation analysis for detection and staging of prostate cancer. Clin Cancer Res. 2005; 11(18): 6582-6588.

169. Bastian PJ, Elinger J, Wellmann A, Wernert N, Heukamp LC, Muller SC, et al. Diagnostic and prognostic information in prostate cancer with the help of a small set of hypermethylated gene loci. Clin Cancer Res. 2005; 11(11): 4097-4106.

170. Singal R, Ferdinand L, Reis IM, Schlesselman JJ. Methylation of multiple genes in prostate cancer and the relationship with clinicopathological features of disease. Oncol Rep. 2004; 12(3): 631-637.

171. Woodson K, Hayes R, Wideroff L, Villaruz L, Tangrea J. Hypermethylation of *GSTP1*, *CD44*, and *E-cadherin* genes in prostate cancer among US Blacks and Whites. Prostate. 2003; 55(3): 199-205.

172. Cho NY, Kim BH, Choi M, Yoo EJ, Moon KC, Cho YM, et al. Hypermethylation of CpG island loci and hypomethylation of *LINE-1* and *Alu* repeats in prostate adenocarcinoma and their relationship to clinicopathological features. J Pathol. 2007; 211(3): 269-277.

173. Alumkal JJ, Zhang Z, Humphreys EB, Bennett C, Mangold LA, Carducci MA, et al. Effect of DNA methylation on identification of aggressive prostate cancer. Urology. 2008; 72(6): 1234-1239.

174. Kawamoto K, Okino ST, Place RF, Urakami S, Hirata H, Kikuno N, et al. Epigenetic modifications of *RASSF1A* gene through chromatin remodeling in prostate cancer. Clin Cancer Res. 2007; 13(9): 2541-2548.

175. Liu L, Kron KJ, Pethe VV, Demetrashvili N, Nesbitt ME, Trachtenberg J, et al. Association of tissue promoter methylation levels of *APC*, *TGFβ2*, *HOXD3* and *RASSF1A* with prostate cancer progression. Int J Cancer. 2011; 129(10): 2454-2462.

176. Aitchison A, Warren A, Neal D, Rabbitts P. *RASSF1A* promoter methylation is frequently detected in both pre-malignant and non-malignant microdissected prostatic epithelial tissues. Prostate. 2007; 67(6): 638-644.

177. Zhang J, Liu L, Pfeifer GP. Methylation of the retinoid response gene *TIG1* in prostate cancer correlates with methylation of the retinoic acid receptor beta gene. Oncogene. 2004; 23(12): 2241-2249.

178. Verdoodt B, Sommerer F, Palisaar RJ, Noldus J, Vogt M, Nambiar S, et al. Inverse association of *p16 INK4a* and *p14 ARF* methylation of the *CDKN2a* locus in different Gleason scores of prostate cancer. Prostate Cancer Prostatic Dis. 2011; 14(4): 295-301.

179. Kito H, Suzuki H, Ichikawa T, Sekita N, Kamiya N, Akakura K, et al. Hypermethylation of the *CD44* gene is associated with progression and metastasis of human prostate cancer. Prostate. 2001; 49(2): 110-115.

180. Hesson LB, Cooper WN, Latif F. The role of *RASSF1A* methylation in cancer. Dis Markers. 2007; 23(1-2): 73-87.

181. Liu L, Yoon JH, Dammann R, Pfeifer GP. Frequent hypermethylation of the *RASSF1A* gene in prostate cancer. Oncogene. 2002; 21(44): 6835-6840.

182. Pan J, Chen J, Zhang B, Chen X, Huang B, Zhuang J, et al. Association between *RASSF1A* promoter methylation and prostate cancer: a systematic review and meta-analysis. PLoS One. 2013; 8(9): e75283.

183. Nakayama T, Watanabe M, Yamanaka M, Hirokawa Y, Suzuki H, Ito H, et al. The role of epigenetic modifications in retinoic acid receptor beta2 gene expression in human prostate cancers. Lab Invest. 2001; 81(7): 1049-1057.

184. Woodson K, O'Reilly KJ, Ward DE, Walter J, Hanson J, Walk EL, et al. *CD44* and *PTGS2* methylation are independent prognostic markers for biochemical recurrence among prostate cancer patients with clinically localized disease. Epigenetics. 2006; 1(4): 183-186.

185. Nguyen TT, Nguyen CT, Gonzales FA, Nichols PW, Yu MC, Jones PA. Analysis of cyclin-dependent kinase inhibitor expression and methylation patterns in human prostate cancers. Prostate. 2000; 43(3): 233-242.

186. Jiang D, Shen Y, Dai D, Xu Y, Xu C, Zhu H, et al. Meta-analyses of methylation markers for prostate cancer. Tumour Biol. 2014; 35(10): 10449-10455.

187. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007; 128(4): 683-692.

188. Lin PC, Giannopoulou EG, Park K, Mosquera JM, Sboner A, Tewari AK, et al. Epigenomic alterations in localized and advanced prostate cancer. Neoplasia. 2013; 15(4): 373-383.

189. Bhasin JM, Lee BH, Matkin L, Taylor MG, Hu B, Xu Y, et al. Methylome-wide sequencing detects DNA hypermethylation distinguishing indolent from aggressive prostate cancer. Cell Rep. 2015; 13(10): 2135-2146.

190. Devaney JM, Wang S, Funda S, Long J, Taghipour DJ, Tbaishat R, et al. Identification of novel DNA-methylated genes that correlate with human prostate cancer and high-grade prostatic intraepithelial neoplasia. Prostate Cancer Prostatic Dis. 2013; 16(4): 292-300.

191. Kim YJ, Yoon HY, Kim SK, Kim YW, Kim EJ, Kim IY, et al. *EFEMP1* as a novel DNA methylation marker for prostate cancer: array-based DNA methylation and expression profiling. Clin Cancer Res. 2011; 17(13): 4523-4530.

192. Vener T, Derecho C, Baden J, Wang H, Rajpurohit Y, Skelton J, et al. Development of a multiplexed urine assay for prostate cancer diagnosis. Clin Chem. 2008; 54(5): 874-882.

193. Cairns P, Esteller M, Herman JG, Schoenberg M, Jeronimo C, Sanchez-Cespedes M, et al. Molecular detection of prostate cancer in urine by *GSTP1* hypermethylation. Clin Cancer Res. 2001; 7(9): 2727-2730.

194. Goessl C, Muller M, Heicappell R, Krause H, Miller K. DNA-based detection of prostate cancer in blood, urine, and ejaculates. Ann N Y Acad Sci. 2001; 945: 51-58.

195. Jeronimo C, Usadel H, Henrique R, Silva C, Oliveira J, Lopes C, et al. Quantitative *GSTP1* hypermethylation in bodily fluids of patients with prostate cancer. Urology. 2002; 60(6): 1131-1135.

196. Payne SR, Serth J, Schostak M, Kamradt J, Strauss A, Thelen P, et al. DNA methylation biomarkers of prostate cancer: confirmation of candidates and evidence urine is the most sensitive body fluid for non-invasive detection. Prostate. 2009; 69(12): 1257-1269.

197. Roupret M, Hupertan V, Catto JW, Yates DR, Rehman I, Proctor LM, et al. Promoter hypermethylation in circulating blood cells identifies prostate cancer progression. Int J Cancer. 2008; 122(4): 952-956.

198. Sunami E, Shinozaki M, Higano CS, Wollman R, Dorff TB, Tucker SJ, et al. A multimarker circulating DNA assay for assessing prostate cancer patients' blood. Clin Chem. 2009; 55(3): 559-567.

199. Ellinger J, Haan K, Heukamp LC, Kahl P, Buttner R, Muller SC, et al. CpG island hypermethylation in cell-free serum DNA identifies patients with localized prostate cancer. Prostate. 2008; 68(1): 42-49.

200. Payne SR, Serth J, Schostak M, Kamradt J, Strauss A, Thelen P, et al. DNA methylation biomarkers of prostate cancer: confirmation of candidates and evidence urine is the most sensitive body fluid for non-invasive detection. Prostate. 2009; 69(12): 1257-1269.

201. Crocitto LE, Korns D, Kretzner L, Shevchuk T, Blair SL, Wilson TG, et al. Prostate cancer molecular markers *GSTP1* and *hTERT* in expressed prostatic secretions as predictors of biopsy results. Urology. 2004; 64(4): 821-825.

202. Di Meo A, Bartlett J, Cheng Y, Pasic MD, Yousef GM. Liquid biopsy: a step forward towards precision medicine in urologic malignancies. Mol Cancer. 2017; 16(1): 80.

203. Roupret M, Hupertan V, Yates DR, Catto JW, Rehman I, Meuth M, et al. Molecular detection of localized prostate cancer using quantitative methylation-specific PCR on urinary cells obtained following prostate massage. Clin Cancer Res. 2007; 13(6): 1720-1725.

204. Hoque MO, Topaloglu O, Begum S, Henrique R, Rosenbaum E, Van Criekinge W, et al. Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects. J Clin Oncol. 2005; 23(27): 6569-6575.

205. Truong M, Yang B, Jarrad DF. Toward the detection of prostate cancer in urine: a critical analysis. J Urol. 2013; 189(2): 422-429.

206. Wu T, Giovannucci E, Welge J, Mallick P, Tang WY, Ho SM. Measurement of *GSTP1* promoter methylation in body fluids may complement PSA screening: a meta-analysis. Br J Cancer. 2011; 105(1): 65-73.

207. Baden J, Green G, Painter J, Curtin K, Markiewicz J, Jones J, et al. Multicenter evaluation of an investigational prostate cancer methylation assay. J Urol. 2009; 182(3): 1186-1193.

208. Lin YL, Deng QK, Wang YH, Fu XL, Ma JG, Li WP. Aberrant protocadherin17 (*PCDH17*) methylation in serum is a potential predictor for recurrence of early-stage prostate cancer patients after radical prostatectomy. Med Sci Monit. 2015; 21: 3955-3690.

209. Deng QK, Lei YG, Lin YL, Ma JG, Li WP. Prognostic Value of protocadherin10 (*PCDH10*) methylation in serum of prostate cancer patients. Med Sci Monit. 2016; 22: 516-521.

210. Bastian PJ, Palapattu GS, Lin X, Yegnasubramanian S, Mangold LA, Trock B, et al. Preoperative serum DNA *GSTP1* CpG island hypermethylation and the risk of early prostate-specific antigen recurrence following radical prostatectomy. Clin Cancer Res. 2005; 11(11): 4037-4043.

211. Jatkoe TA, Karnes RJ, Freedland SJ, Wang Y, Le A, Baden J. A urine-based methylation signature for risk stratification within low-risk prostate cancer. Br J Cancer. 2015; 112(5): 802-808.

212. Zhao F, Olkhov-Mitsel E, van der Kwast T, Sykes J, Zdravic D, Venkateswaran V, et al. Urinary DNA methylation biomarkers for noninvasive prediction of aggressive disease in patients with prostate cancer on active surveillance. J Urol. 2017; 197(2): 335-341.

213. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004; 116(2): 281-297.

214. Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell. 2009; 136(4): 642-655.

215. Kunej T, Godnic I, Horvat S, Zorc M, Calin GA. Cross talk between microRNA and coding cancer genes. Cancer J. 2012; 18(3): 223-231.

216. Vanacore D, Boccellino M, Rossetti S, Cavaliere C, D'Aniello C, Di Franco R, et al. Micrornas in prostate cancer: an overview. Oncotarget. 2017; [in press].

217. Lee Y, Kim M, Han J, Yeom KH, Le S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004; 23(20): 4051-4060.

218. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. Genome Res. 2004; 14(10A): 1902-1910.

219. Godnic I, Zorc M, Jevsinek Skok D, Calin GA, Horvat S, Dovc P, et al. Genome-wide and species-wide *in silico* screening for intragenic microRNAs in human, mouse and chicken. PLoS One. 2013; 8(6): e65165.

220. Kim YK, Kim VN. Processing of intronic microRNAs. EMBO J. 2007; 26(3): 775-783.

221. Ramalingam P, Palanichamy JK, Singh A, Das P, Bhagat M, Kassab MA, et al. Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. RNA. 2014; 20(1): 76-87.

222. Zhou H, Lin K. Excess of microRNAs in large and very 5' biased introns. Biochem Biophys Res Commun. 2008; 368(3): 709-715.

223. Golan D, Levy C, Friedman B, Shomron N. Biased hosting of intronic microRNA genes. Bioinformatics. 2010; 26(8): 992-995.

224. Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, et al. Clustering and conservation patterns of human microRNAs. Nucleic Acids Res. 2005; 33(8): 2697-2706.

225. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA. 2005; 11(3): 241-247.

226. Tsuruta T, Kozaki K, Uesugi A, Furuta M, Hirasawa A, Imoto I, et al. miR-152 is a tumor suppressor microRNA that is silenced by DNA hypermethylation in endometrial cancer. Cancer Res. 2011; 71(20): 6450-6462.

227. Lin PC, Chiu YL, Banerjee S, Park K, Mosquera JM, Giannopoulou E, et al. Epigenetic repression of miR-31 disrupts androgen receptor homeostasis and contributes to prostate cancer progression. Cancer Res. 2013; 73(3): 1232-1244.

228. He C, Li Z, Chen P, Huang H, Hurst LD, Chen J. Young intragenic miRNAs are less coexpressed with host genes than old ones: implications of miRNA-host gene coevolution. Nucleic Acids Res. 2012; 40(9): 4002-4012.

229. Saito Y, Friedman JM, Chihara Y, Egger G, Chuang JC, Liang G. Epigenetic therapy upregulates the tumor suppressor microRNA-126 and its host gene EGFL7 in human cancer cells. Biochem Biophys Res Commun. 2009; 379(3): 726-731.

230. Tsang J, Zhu J, van Oudenaarden A. MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. Mol Cell. 2007; 26(5): 753-767.

231. Wang G, Wang Y, Shen C, Huang YW, Huang K, Huang TH, et al. RNA polymerase II binding patterns reveal genomic regions involved in microRNA gene regulation. PLoS One. 2010; 5(11): e13798.

232. Suzuki H, Maruyama R, Yamamoto E, Kai M. Epigenetic alteration and microRNA dysregulation in cancer. Front Genet. 2013; 4: 258.

233. Kozaki K, Inazawa J. Tumor-suppressive microRNA silenced by tumor-specific DNA hypermethylation in cancer cells. Cancer Sci. 2012; 103(5): 837-845.

234. Lynch SM, O'Neill KM, McKenna MM, Walsh CP, McKenna DJ. Regulation of miR-200c and miR-141 by methylation in prostate cancer. Prostate. 2016; 76(13): 1146-1159.

235. Torres-Ferreira J, Ramalho-Carvalho J, Gomez A, Menezes FD, Freitas R, Oliveira J, et al. MiR-193b promoter methylation accurately detects prostate cancer in urine sediments and miR-34b/c or miR-129-2 promoter methylation define subsets of clinically aggressive tumors. Mol Cancer. 2017; 16(1): 26.

236. Ramalho-Carvalho J, Graca I, Gomez A, Oliveira J, Henrique R, Esteller M, et al. Downregulation of miR-130b~301b cluster is mediated by aberrant promoter methylation and impairs cellular senescence in prostate cancer. J Hematol Oncol. 2017; 10(1): 43.

237. Padrao NA, Monteiro-Reis S, Torres-Ferreira J, Antunes L, Leca L, Montezuma D, et al. MicroRNA promoter methylation: a new tool for accurate detection of urothelial carcinoma. Br J Cancer. 2017; 116(5): 634-639.

238. He Y, Cui Y, Wang W, Gu J, Guo S, Ma K, et al. Hypomethylation of the hsa-miR-191 locus causes high expression of hsa-mir-191 and promotes the epithelial-to-mesenchymal transition in hepatocellular carcinoma. Neoplasia. 2011; 13(9): 841-853.

239. Suzuki H, Maruyama R, Yamamoto E, Kai M. DNA methylation and microRNA dysregulation in cancer. Mol Oncol. 2012; 6(6): 567-578.

240. Chi P, Allis CD, Wang GG. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nat Rev Cancer. 2010; 10(7): 457-469.

241. Zhang G, Esteve PO, Chin HG, Terragni J, Dai N, Correa IRJ, et al. Small RNA-mediated DNA (cytosine-5) methyltransferase 1 inhibition leads to aberrant DNA methylation. Nucleic Acids Res. 2015; 43(12): 6112-6124.

242. Lopez-Serra P, Esteller M. DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. Oncogene. 2012; 31(13): 1609-1622.

243. Demidenko R, Razanauskas D, Daniunaite K, Lazutka JR, Jankevicius F, Jarmalaite S. Frequent down-regulation of ABC transporter genes in prostate cancer. BMC Cancer. 2015; 15: 683.

244. Stuopelyte K, Daniunaite K, Bakavicius A, Lazutka JR, Jankevicius F, Jarmalaite S. The utility of urine-circulating miRNAs for detection of prostate cancer. Br J Cancer. 2016; 115(6): 707-715.

245. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002; 12(4): 996-1006.

246. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2015 update. Nucleic Acids Res. 2015; 43(Database issue): D670-D681.

247. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2006; 34(Database issue): D140-D144.

248. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2014; 42(Database issue): D68-D73.

249. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25(1): 25-29.

250. Gene Ontology Consorcium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015; 43(Database issue): D1049-D1056.

251. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. 2005; 21(16): 3448-3449.

252. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13(11): 2498-2504.

253. Gray KA, Yates B, Seal RL, Wright MW, Bruford EA. Genenames.org: the HGNC resources in 2015. Nucleic Acids Res. 2015; 43(Database issue): D1079-D1085.

254. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102(43): 15545-15550.

255. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015; 1(6): 417-425.

256. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002; 18(11): 1427-1431.

257. Trinh BN, Long TI, Laird PW. DNA methylation analysis by MethyLight technology. Methods. 2001; 25(4): 456-462.

258. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2(5): 401-404.

259. Huang WY, Hsu SD, Huang HY, Sun YM, Chou CH, Weng SL, et al. MethHC: a database of DNA methylation and gene expression in human cancer. Nucleic Acids Res. 2015; 43(Database issue): D856-D861.

260. Medvedeva YA, Lennartsson A, Ehsani R, Kulakovskiy IV, Vorontsov IE, Panahandeh P, et al. EpiFactors: a comprehensive database of human epigenetic factors and complexes. Database (Oxford). 2015; 2015: bav067.

261. Konishi N, Nakamura M, Kishi M, Nishimine M, Ishida E, Shimada K. DNA hypermethylation status of multiple genes in prostate adenocarcinomas. Jpn J Cancer Res. 2002; 93(7): 767-773.

262. Delgado-Cruzata L, Hruby GW, Gonzalez K, McKiernan J, Benson MC, Santella RM, et al. DNA methylation changes correlate with Gleason score and tumor stage in prostate cancer. DNA Cell Biol. 2012; 31(2): 187-192.

263. Chao C, Chi M, Preciado M, Black MH. Methylation markers for prostate cancer prognosis: a systematic review. Cancer Causes Control. 2013; 24(9): 1615-1641.

264. Kirby MK, Ramaker RC, Roberts BS, Lasseigne BN, Gunther DS, Burwell TC, et al. Genome-wide DNA methylation measurements in prostate tissues uncovers novel prostate cancer diagnostic biomarkers and transcription factor binding patterns. BMC Cancer. 2017; 17(1): 273.

265. Haldrup C, Mundbjerg K, Vestergaard EM, Lamy P, Wild P, Schulz WA, et al. DNA methylation signatures for prediction of biochemical recurrence after radical prostatectomy of clinically localized prostate cancer. J Clin Oncol. 2013; 31(26): 3250-3258.

266. Shui IM, Wong CJ, Zhao S, Kolb S, Ebot EM, Geybels MS, et al. Prostate tumor DNA methylation is associated with cigarette smoking and adverse prostate cancer outcomes. Cancer. 2016; 122(14): 2168-2177.

267. Castelo-Branco P, Leao R, Lipman T, Campbell B, Lee D, Price A, et al. A cancer specific hypermethylation signature of the *TERT* promoter predicts biochemical relapse in prostate cancer: a retrospective cohort study. Oncotarget. 2016; 7(36): 57726-57736.

268. Angulo JC, Andres G, Ashour N, Sanchez-Chapado M, Lopez JI, Ropero S. Development of castration resistant prostate cancer can be predicted by a DNA hypermethylation profile. J Urol. 2016; 195(3): 619-626.

269. Nelson WG, De Marzo AM, Yegnasubramanian S. Epigenetic alterations in human prostate cancers. Endocrinology. 2009; 150(9): 3991-4002.

270. Chen D, Purohit A, Halilovic E, Doxsey SJ, Newton AC. Centrosomal anchoring of protein kinase C betaII by pericentrin controls microtubule organization, spindle function, and cytokinesis. J Biol Chem. 2004; 279(6): 4829-4839.

271. Liu S, Chen X, Chen R, Wang J, Zhu G, Jiang J, et al. Diagnostic role of Wnt pathway gene promoter methylation in non small cell lung cancer. Oncotarget. 2017; [in press].

272. Hagiwara K, Ito H, Murate T, Miyata Y, Ohashi H, Nagai H. *PROX1* overexpression inhibits protein kinase C beta II transcription through promoter DNA methylation. Genes Chromosomes Cancer. 2012; 51(11): 1024-1036.

273. Koren R, Ben Meir D, Langzam L, Dekel Y, Konichezky M, Baniel J, et al. Expression of protein kinase C isoenzymes in benign hyperplasia and carcinoma of prostate. Oncol Rep. 2004; 11(2): 321-326.

274. Metzger E, Imhof A, Patel D, Kahl P, Hoffmeyer K, Friedrichs N, et al. Phosphorylation of histone H3T6 by PKCbeta(I) controls demethylation at histone H3K4. Nature. 2010; 464(7289): 792-796.

275. Kim J, Choi YL, Vallentin A, Hunrichs BS, Hellerstein MK, Peekl DM, et al. Centrosomal PKCβII and pericentrin are critical for human prostate cancer growth and angiogenesis. Cancer Res. 2008; 68(16): 6831-6839.

276. El Hour M, Moncada-Pazos A, Blacher S, Masset A, Cal S, Berndt S, et al. Higher sensitivity of Adamts12-deficient mice to tumor growth and angiogenesis. Oncogene. 2010; 29(20): 3025-3032.

277. Lachance G, Uniacke J, Audas TE, Holterman CE, Franovic A, Payette J, et al. DNMT3a epigenetic program regulates the HIF-2α oxygen-sensing pathway and the cellular response to hypoxia. Proc Natl Acad Sci U S A. 2014; 111(21): 7783-7788.

278. Cui J, Duan B, Zhao X, Chen Y, Sun S, Deng W, et al. MBD3 mediates epigenetic regulation on *EPAS1* promoter in cancer. Tumour Biol. 2016; 37(10): 13455-13467.

279. Moncada-Pazos A, Obaya AJ, Fraga MF, Viloria CG, Capella G, Gausachs M, et al. The *ADAMTS12* metalloprotease gene is epigenetically silenced in tumor cells and transcriptionally activated in the stroma during progression of colon cancer. J Cell Sci. 2009; 122(Pt 16): 2906-2913.

280. Sharma M, Li X, Wang Y, Zarnegar M, Huang CY, Palvimo JJ, et al. hZimp10 is an androgen receptor co-activator and forms a complex with SUMO-1 at replication foci. EMBO J. 2003; 22(22): 6101-6114.

281. Li X, Zhu C, Yang N, Qin H, Sun Z. ZMIZ1 preferably enhances the transcriptional activity of androgen receptor with short polyglutamine tract. PLoS One. 2011; 6(9): e25040.

282. Desotelle J, Truong M, Ewald J, Weeratunga P, Yang B, Huang W. CpG island hypermethylation frequently silences *FILIP1L* isoform 2 expression in prostate cancer. J Urol. 2013; 189(1): 329-335.

283. Burton ER, Gaffar A, Lee SJ, Adeshuko F, Whitney KD, Chung JY, et al. Downregulation of Filamin A interacting protein 1-like is associated with promoter methylation and induces an invasive phenotype in ovarian cancer. Mol Cancer Res. 2011; 9(8): 1126-1138.

284. Kwon M, Lee SJ, Reddy S, Rybak Y, Adem A, Libutti SK. Down-regulation of Filamin A interacting protein 1-like is associated with promoter methylation and an invasive phenotype in breast, colon, lung and pancreatic cancers. PLoS One. 2013; 8(12): e82620.

285. Orian-Rousseau V, Sleeman J. CD44 is a multidomain signaling platform that integrates extracellular matrix cues with growth factor and cytokine signals. Adv Cancer Res. 2014; 123: 231-254.

286. Faryna M, Konermann C, Aulmann S, Bermejo JL, Brugger M, Diederichs S, et al. Genome-wide methylation screen in low-grade breast cancer identifies novel epigenetically altered genes as potential biomarkers for tumor diagnosis. FASEB J. 2012; 26(12): 4937-4950.

287. Rawluszko-Wieczorek AA, Horbacka K, Krokowicz P, Misztal M, Jagodzinski PPrognostic potential of DNA methylation and transcript levels of *HIF1A* and *EPAS1* in colorectal cancer. Mol Cancer Res. 2014; 12(8): 1112-1127.

288. Cruzeiro GA, Dos Reis MB, Silveira VS, Lira RC, Carlotti CG, Neder L, et al. *HIF1A* is overexpressed in medulloblastoma and its inhibition reduces proliferation and increases *EPAS1* and *ATG16L1* methylation. Curr Cancer Drug Targets. 2017; [in press].

289. Kaneta Y, Ullrich A. *NEK9* depletion induces catastrophic mitosis by impairment of mitotic checkpoint control and spindle dynamics. Biohem Biophys Res Commun. 2013; 442(3-4): 139-146.

290. Shimizu T, Suzuki H, Nojima M, Kitamura H, Yamamoto E, Maruyama R, et al. Methylation of a panel of microRNA genes is a novel biomarker for detection of bladder cancer. Eur Urol. 2013; 63(6): 1091-1100.

291. Vrba L, Munoz-Rodriguez JL, Stampfer MR, Futscher BW. miRNA gene promoters are frequent targets of aberrant DNA methylation in human breast cancer. PLoS One. 2013; 8(1): e54398.

292. Lodygin D, Tarasov V, Epanchintsev A, Berking C, Knyazeva T, Korner H, et al. Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. Cell Cycle. 2008; 7(16): 2591-2600.

293. Theodore SC, Davis M, Zhao F, Wang H, Chen D, Rhim J, et al. MicroRNA profiling of novel African American and Caucasian prostate cancer cell lines reveals a reciprocal regulatory relationship of miR-152 and DNA methyltranferase 1. Oncotarget. 2014; 5(11): 3512-3525.

294. Nilsson EM, Laursen KB, Whitchurch J, McWilliam A, Odum N, Persson JL, et al. MiR137 is an androgen regulated repressor of an extended network of transcriptional coregulators. Oncotarget. 2015; 6(34): 35710-35725.

295. Krzeminski P, Sarasquete ME, Misiewicz-Krzeminska I, Corral R, Corchete LA, Martin AA, et al. Insights into epigenetic regulation of microRNA-155 expression in multiple myeloma. Biochim Biophys Acta. 2015; 1849(3): 353-366.

296. Li CL, Nie H, Wang M, Su LP, Li JF, Yu YY, et al. microRNA-155 is downregulated in gastric cancer cells and involved in cell metastasis. Oncol Rep. 2012; 27(6): 1960-1996.

297. Cai ZK, Chen Q, Chen YB, Gu M, Zheng DC, Zhou J, et al. microRNA-155 promotes the proliferation of prostate cancer cells by targeting annexin 7. Mol Med Rep. 2015; 11(1): 533-538.

298. Ji H, Li Y, Jiang F, Wang X, Zhang J, Shen J, et al. Inhibition of transforming growth factor beta/SMAD signal by MiR-155 is involved in arsenic trioxide-induced anti-angiogenesis in prostate cancer. Cancer Sci. 2014; 105(12): 1541-1549.

299. Cha YJ, Lee jH, Han HH, Kim BG, Kang S, Choi YD, et al. MicroRNA alteration and putative target genes in high-grade prostatic intraepithelial neoplasia and prostate cancer: *STAT3* and *ZEB1* are upregulated during prostate carcinogenesis. Prostate. 2016; 76(10): 937-947.

300. Kohler CU, Bryk O, Meier S, Lang K, Rozynek P, Bruning T, et al. Analyses in human urothelial cells identify methylation of miR-152, miR-200b and miR-10a genes as candidate bladder cancer biomarkers. Biochem Biophys Res Commun. 2013; 438(1): 48-53.

301. Sengupta D, Deb M, Rath SK, Kar S, Parbin S, Pradhan N, et al. DNA methylation and not H3K4 trimethylation dictates the expression status of miR-152 gene which inhibits migration of breast cancer cells via DNMT1/CDH1 loop. Exp Cell Res. 2016; 346(2): 176-187.

302. Steponaitiene R, Kupcinskas J, Langner C, Balaguer F, Venclauskas L, Pauzas H, et al. Epigenetic silencing of miR-137 is a frequent event in gastric carcinogenesis. Mol Carcinog. 2016; 55(4): 376-386.

303. Lichner Z, Fendler A, Saleh C, Nasser AN, Boles D, Al-Haddad S, et al. MicroRNA signature helps distinguish early from late biochemical failure in prostate cancer. Clin Chem. 2013; 59(11): 1595-1603.

304. Zhu C, Li J, Ding Q, Cheng G, Zhou H, Tao L, et al. miR-152 controls migration and invasive potential by targeting TGFα in prostate cancer cell lines. Prostate. 2013; 73(10): 1082-1089.

305. Long Q, Johnson BA, Osunkoya AO, Lai YH, Zhou W, Abramovitz M, et al. Protein-coding and microRNA biomarkers of recurrence of prostate cancer following radical prostatectomy. Am J Pathol. 2011; 179(1): 46-54.

306. Ji W, Yang L, Yuan J, Yang L, Zhang M, Qi D, et al. MicroRNA-152 targets DNA methyltransferase 1 in NiS-transformed cells via a feedback mechanism. Carcinogenesis. 2013; 34(2): 446-453.

307. Wang Q, Li C, Zhu Z, Teng Y, Che X, Wang Y, et al. miR-155-5p antagonizes the apoptotic effect of bufalin in triple-negative breast cancer cells. Anticancer Drugs. 2016; 27(1): 9-16.

308. Metzger E, Wissmann M, Yin N, Muller JM, Schneider R, Peters AH, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature. 2005; 437(7057): 436-439.

309. Kahl P, Gullotti L, Heukamp LC, Wolf S, Friedrichs N, Vorreuther R, et al. Androgen receptor coactivators lysine-specific histone demethylase 1 and four and a half LIM domain protein 2 predict risk of prostate cancer recurrence. Cancer Res. 2006; 66(23): 11341-11347.

310. Metzger E, Willmann D, McMillan J, Forne I, Metzger P, Gerhardt S, et al. Assembly of methylated KDM1A and CHD1 drives androgen receptor-dependent transcription and translocation. Nat Struct Mol Biol. 2016; 23(2): 132-139.

311. Balaguer F, Link A, Lozano JJ, Cuatrecasas M, Nagasaka T, Boland CR, et al. Epigenetic silencing of miR-137 is an early event in colorectal carcinogenesis. Cancer Res. 2010; 70(16): 6609-6618.

312. Xiang Y, Zhu Z, Han G, Ye X, Xu B, Peng Z, et al. JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. Proc Natl Acad Sci U S A. 2007; 104(49): 19226-19231.

313. Burbee DG, Forgacs E, Zochbauer-Muller S, Shivakumar L, Fong K, Gao B, et al. Epigenetic inactivation of *RASSF1A* in lung and breast cancers and malignant phenotype suppression. J Natl Cancer Inst. 2001; 93(9): 691-699.

314. Costa VL, Henrique R, Ribeiro FR, Pinto M, Oliveira J, Lobo F, et al. Quantitative promoter methylation analysis of multiple cancer-related genes in renal cell tumors. BMC Cancer. 2007; 23(7): 133.

315. Esteller M, Sanchez-Cespedes M, Rosell R, Sidransky D, Baylin SB, Herman JG. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res. 1999; 59(1): 67-70.

316. Paz MF, Yaya-Tur R, Rojas-Marcos I, Reynes G, Pollan M, Aquirre-Cruz L, et al. CpG island hypermethylation of the DNA repair enzyme methyltransferase predicts response to temozolomide in primary gliomas. Clin Cancer Res. 2004; 10(15): 4933-4938.

317. Esteller M, Tortola S, Toyota M, Capella G, Peinado MA, Baylin SB, et al. Hypermethylation-associated inactivation of *p14*(*ARF*) is independent of *p16*(*INK4a*) methylation and p53 mutational status. Cancer Res. 2000; 60(1): 129-133.

318. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A. 1996; 93(18): 9821-9826.

319. Lehmann U, Langer F, Feist H, Gockner S, Hasemeier B, Kreipe H. Quantitative assessment of promoter hypermethylation during breast cancer development. Am J Pathol. 2002; 160(2): 605-612.

320. Eads CA, Lord RV, Wickramasinghe K, Long TI, Kurumboor SK, Bernstein L, et al. Epigenetic patterns in the progression of esophageal adenocarcinoma. Cancer Res. 2001; 61(8): 3410-3418.

321. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. Nucleic Acids Res. 2015; 43(Database issue): D146-D152.

322. Agarwal V, Bell GW, Nam JW, Bartel DPredicting effective microRNA target sites in mammalian mRNAs. Elife. 2015; 4: e05005.

323. Chou CH, Chang NW, Shrestha S, Hsu SD, Lin YL, Lee WH, et al. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. Nucleic Acids Res. 2016; 44(D1): D239-D247.

324. Dweep H, Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods. 2015; 12(8): 697.

325. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Res. 2008; 36(Database issue): D149-D153.

326. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. Nat Genet. 2005; 37(5): 495-500.

Online resources (by the order of citation):

1. European Cancer Observatory, International Agency for Research on Cancer, World Health Organization [website]; last accessed 11-09-2016. https://www.iarc.fr
2. Cancer Registry, National Cancer Institute (Lithuania) [website]; last accessed 05-09-2016. http://www.nvi.lt
3. Cancer Research UK [website]; last accessed 19-12-2016. https://www.cancerresearchuk.org
4. American Cancer Society [website]; last accessed 19-12-2016. http://www.cancer.org
5. The Cancer Genome Atlas, National Cancer Institute, National Human Genome Research Institute [website]; last accessed 05-08-2016. https://cancergenome.nih.gov
6. miRBase: the microRNA database, release 21, University of Manchester [database]; last accessed 15-12-2016. http://www.mirbase.org
7. Gene Expression Omnibus, National Center for Biotechnology Information [database]; last accessed 05-09-2016. https://www.ncbi. nlm.nih. gov/geo/
8. SureDesign, Agilent Technologies [online tool]; last accessed 25-09-2016. https://earray.chem.agilent.com/suredesign/
9. UCSC Genome Browser, University of California Santa Cruz [online tool]; last accessed 04-10-2016. https://genome.ucsc.edu
10. Genenames.org, HUGO Gene Nomenclature Committee [website], last accessed 15-12-2016. http://www.genenames.org
11. Gene Set Enrichment Analysis (GSEA and Molecular Signatures Database, MSigDB v5.2), Broad Institute [online tool & database]; last accessed 11-11-2016. http://software.broadinstitute.org/gsea/
12. MethPrimer, The Li Lab, Peking Union Medical College Hospital (PUMCH), Chinese Academy of Medical Sciences [online tool]; last accessed 20-10-2015. http://www.urogene.org/methprimer/index.html
13. cBioPortal for Cancer Genomics, Memorial Sloan Kettering Cancer Center (The Cancer Genome Atlas) [online tool & database]; last accessed 25-09-2016. http://www.cbio portal.org
14. MethHC, a database of DNA methylation and gene expression in human cancer, v1.0.3, Institute of Bioinformatics and Systems Biology, National Chiao Tung University [database]; last accessed 30-07-2016. http://methhc.mbc.nctu.edu.tw
15. Gene, National Center for Biotechnology Information [database]; last accessed 16-12-2016. https://www.ncbi. nlm.nih.gov/geo/
16. GeneCards®: The Human Gene Database, Weizmann Institute of Science [database]; last accessed 15-12-2016. http://www.gene cards.org
17. EpiFactors, Supported by RFBR grants 14-04-00180 and 15-34-20423 [database]; last accessed 27-09-2016. http://epifactors.autosome.ru
18. The target prediction database, miRDB, Washington University [database]; last accessed 20-04-2016. http://mirdb.org
19. TargetScanHuman, Prediction of microRNA targets, Whitehead Institute for Biomedical Research [database]; last accessed 20-04-2016. http://www.targetscan.org
20. miRTarBase, Institute of Bioinformatics and Systems Biology, National Chiao Tung University [database]; last accessed 20-04-2016. http://mirtarbase.mbc.nctu.edu.tw
21. miRWalk 2.0, Ruprecht-Karls-Universität Heidelberg [database]; last accessed 20-04-2016. http://zmf.umm.uni-heidelberg.de/apps/zmf/mir walk2/
22. microRNA.org, Memorial Sloan-Kettering Cancer Center [database]; last accessed 20-04-2016. http://www.microrna.org
23. PicTar, New York University [database]; last accessed 20-04-2016. http://pictar.mdc-berlin.de/

Specialized software (stand-alone; by the order of citation):

1. Feature Extraction, version 10.7.3.1, Agilent Technologies, USA.
2. GeneSpring™ GX, version 13.1.1, Agilent Technologies, USA.
3. Cytoscape, version 3.2.1, National Institute of General Medical Sciences, USA.
4. Qlucore Omics Explorer, version 3.0, trial license, Sweden.
5. Methyl Primer Express® Software, version 1.0, Applied Biosystems™, Thermo Fisher Scientific, USA.
6. MxPro, version 4.0, Agilent Technologies, USA.
7. 2100 Expert, version B.02.08.SI648, Agilent Technologies.
8. GenEx, version 6.0.1, MultiD Analyses, Sweden.

SUPPLEMENTARY MATERIAL

All sources cited in supplements are listed in References section.

Table S1. Currently used Tumor, Node, Metastasis staging system (TNM) for classification of prostate cancer (adapted from [59]).

Primary tumor (T)		
TX	Primary tumor cannot be assessed	
T0	No evidence of primary tumor	
T1	Clinically inapparent tumor not palpable or visible by imaging	
	T1a	Tumor incidental histological finding in 5% or less of tissue resected
	T1b	Tumor incidental histological finding in more than 5% of tissue resected
	T1c	Tumor identified by needle biopsy (e.g. because of elevated prostate-specific antigen (PSA) level)
T2	Tumor confined within the prostate	
	T2a	Tumor involves one half of one lobe or less
	T2b	Tumor involves more than half of one lobe, but not both lobes
	T2c	Tumor involves both lobes
T3	Tumor extends through the prostatic capsule	
	T3a	Extracapsular extension (unilateral or bilateral) including microscopic bladder neck involvement
	T3b	Tumor invades seminal vesicle(s)
T4	Tumor is fixed or invades adjacent structures other than seminal vesicles: external sphincter, rectum, levator muscles, and/or pelvic wall	
Regional lymph nodes (N)		
NX	Regional lymph nodes cannot be assessed	
N0	No regional lymph node metastasis	
N1	Regional lymph node metastasis	
Distant metastasis (M)		
M0	No distant metastasis	
M1	Distant metastasis	
	M1a	Non-regional lymph node(s)
	M1b	Bone(s)
	M1c	Other site(s)

Table S2. Clinical-pathological and molecular characteristics of the analysis groups.

Parameter	All cases			Methylation analysis in FFPE tissues[1]		Methylation analysis in snap-frozen tissues			Methylation analysis in urine[1]		Gene expression analysis[2]		miRNA expression analysis	
Group composition	PCa[3] (N = 311)	NPT (N = 72)	BPH[4] (N = 35)	PCa (N = 149)	adjacent NPT (N = 37)	PCa (N = 129)	NPT (N = 35)	BPH (N = 17)	PCa (N = 253)	BPH (N = 32)	PCa (N = 81)	NPT (N = 25)	PCa (N = 41)	NPT (N = 12)
Tumor stage, N														
≤pT2	234	-	-	114	-	84	-	-	196	-	51	-	28	-
≥pT3	77	-	-	35	-	45	-	-	57	-	30	-	13	-
ISUP grade group (Gleason score), N														
I (3 + 3)	132	-	-	74	-	32	-	-	123	-	13	-	13	-
II (3 + 4)	119	-	-	50	-	70	-	-	89	-	48	-	22	-
III (4 + 3)	26	-	-	12	-	19	-	-	12	-	14	-	3	-
II or III (7)	11	-	-	5	-	0	-	-	11	-	0	-	0	-
IV (8)	6	-	-	4	-	3	-	-	4	-	2	-	2	-
V (9)	5	-	-	2	-	3	-	-	2	-	3	-	0	-
Unknown	12	-	-	2	-	2	-	-	12	-	1	-	1	-
Tumor cellularity of tissues, N														
90-100%	69	-	-	-	-	63	-	-	-	-	48	-	19	-
70-89%	30	-	-	-	-	30	-	-	-	-	24	-	5	-
50-69%	36	-	-	-	-	36	-	-	-	-	8	-	12	-
30-49%	5	-	-	-	-	0	-	-	-	-	0	-	5	-
1-3%	-	3	-	-	-	-	3	-	-	-	-	0	-	0
0%	-	37	-	-	-	-	32	-	-	-	-	25	-	12
Undefined	149[5]	37[5]	-	149[5]	37[5]	-	-	-	-	-	-	-		
BCR status, N														
Yes (mean time to BCR, mo)	66 (21)[6]	23 (20)[6]	-	36 (22)	12 (23)	32 (18)	12 (17)	-	52 (20)	-	30 (19)	7 (20)	11 (23)	2 (8)
No (mean follow-up, mo)	216 (44)	48 (54)	-	99 (43)	24 (46)	86 (39)	23 (57)	-	176 (39)	-	49 (42)	17 (58)	29 (50)	10 (48)
Unknown	29	1	-	14	1	11	0	-	25	-	2	1	1	0

Table S2. Continued.

Parameter	All cases			Methylation analysis in FFPE tissues[1]		Methylation analysis in snap-frozen tissues			Methylation analysis in urine[1]		Gene expression analysis[2]		miRNA expression analysis	
Group composition	PCa[3] (N = 311)	NPT (N = 72)	BPH[4] (N = 35)	PCa (N = 149)	adjacent NPT (N = 37)	PCa (N = 129)	NPT (N = 35)	BPH (N = 17)	PCa (N = 253)	BPH (N = 32)	PCa (N = 81)	NPT (N = 25)	PCa (N = 41)	NPT (N = 12)
PSA, ng/mL														
Mean±SD (range)	9.0±9.4 [0.1; 84.2]	8.0±6.8 [2.5; 44.0]	8.9±12.2 [0.8; 69.8]	8.8±8.6 [0.9; 69.8]	7.6±4.9 [3.0; 24.0]	10.9±11.6 [2.5; 84.2]	9.3±8.5 [2.5; 44.0]	7.3±6.6 [0.8; 28.1]	8.7±9.4 [0.1; 84.2]	7.3±6.6 [0.8; 28.1]	10.8±9.4 [2.5; 44.0]	10.0±9.5 [2.5; 44.0]	11.6±14.7 [2.8; 84.2]	4.6±1.6 [2.6; 7.9]
Unknown	7	1	2	3	0	2	0	0	6	0	2	1	1	0
Prostate mass, g														
Mean±SD (range)	52±25 [16; 197]	50±19 [16; 104]	-	54±25 [23; 186]	47±15 [26; 95]	48±17 [16; 123]	51±21 [16; 104]	-	53±26 [16; 197]	-	48±17 [16; 123]	48±20 [16; 104]	52±19.8 [26; 126]	58±27 [28; 104]
Unknown	4	1	-	1	1	0	0	-	3	-	0	0	1	0
TMPRSS2-ERG fusion transcript, N														
Yes	118	-	-	46	-	77	-	-	83	-	46	-	24	-
No	71	-	-	36	-	44	-	-	55	-	26	-	17	-
Unknown	122	-	-	67	-	8	-	-	115	-	9	-	0	-
Age, years														
Mean±SD (range)	62±7 [41; 82]	61±7 [46; 77]	72±8 [58; 83]	62±7 [43; 77]	61±6 [48; 77]	61±8 [41; 82]	62±7 [46; 74]	70±8 [59; 80]	62±7 [42; 82]	72±8 [58; 83]	61±8 [41;82]	61±6 [48; 74]	62±8 [48; 73]	63±4 [57; 70]

BCR – biochemical disease recurrence, FFPE – formalin-fixed paraffin-embedded samples, PCa – prostate cancer, NPT – noncancerous prostate tissue, BPH – benign prostatic hyperplasia, PSA – prostate-specific antigen, SD – standard deviation.

[1]BCR-free survival data last updated in July 2013 was used for the analysis in the first stage of the study.

[2]The same sub-set of BPH samples as in "Methylation analysis in snap-frozen tissues" was also included in gene expression analysis.

[3]Tissue samples were available for 248 cases in total

[4]Tissue samples were available for 17 cases in total

[5]Tumor cellularity was not quantitatively evaluated in FFPE samples. For some cases, both snap-frozen and FFPE tissue samples were available.

[6]Incomplete survival data for 2 cases.

Table S3. Primers used for methylation-specific PCR, location of amplicons, and amplification conditions.

Gene symbol	Primer pair ID	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Product size, nt	Amplicon location from TSS	Primer annealing T, °C	Number of PCR cycles	Ref.
Protein-coding genes								
RASSF1	M	GGGTTTTGCGAGAGCGCG	GCTAACAAACGCGAACCG	169	-73/+96	61	39	[313]
	U	GGTTTTGTGAGAGTGTGTTTAG	CACTAACAAACACAAACCAAAC	169	-72/+97	63		
RARB	M	TCGAGAACGCGAGCGATTCG	GACCAATCCAACCGAAACGA	146	+106/+251	60	39	[314]
	U	TTGAGAATGTGAGTGATTTGA	AACCAATCCAACCAAAACAA	146	+106/+251			
GSTP1	M	TTCGGGGTGTAGCGGTCGTC	GCCCCAATACTAAATCACGACG	91	+78/+168	60	37	[315]
	U	GATGTTTGGGGTGTAGTGGTTGTT	CCACCCCAATACTAAATCACAACA	97	+74/+170			
MGMT	M	ATTTGGTGAGTGTTTGGGTCGTTTC	AAAACGCACCTAAAACTCGCCC	159	+103/+261	67	37	[316]
	U	ATTTGGTGAGTGTTTGGGTTGTTTT	AAAACACACCTAAAACTCACCC	159	+103/+261	64		
DAPK	M	GGATAGTCGGATCGAGTTAACGTC	CCCTCCCAAACGCCGA	98	+44/+141	66	36	[315]
	U	GGAGGATAGTTGGATTGAGTTAATGTT	CAAATCCCTCCCAAACACCAA	106	+41/+146			
p14	M	GTGTTAAAGGGCGGCGTAGC	AAAACCCTCACTCGCGACGA	122	+167/+288	66	36	[317]
	U	TTTTTGGTGTTAAAGGGTGGTGTAGT	CACAAAAACCCTCACTCACAACAA	132	+161/+292			
p16	M	TTATTAGAGGGTGGGGCGGATCGC	GACCCCGAACCGCGACCGTAA	150	+227/+376	64	36	[318]
	U	TTATTAGAGGGTGGGGTGGATTGT	CAACCCCAAACCACAACCATAA	151	+227/+377			
PRKCB	M	TAAGCGTAGTTGGACGAGC	AAAACGACGACCGCTACTAC	124	-7/+117	56	36	This study
	U	TGTTAAGTGTAGTTGGATGAGT	AAAACAACAACCACTACTACACC	127	-10/+117			
KCTD8	M	TTTTTATTGTCGTCGTCGTATC	CTCCGCGTACTCCTAACG	169	+45/+213	58	37	This study
	U	GTTTTTTTATTGTTGTTGTTGTATT	ACCCTCCACATACTCCTAACA	175	+42/+216			
FILIP1L	M	TACGGTTCGTTTATACGGTC	CGACCTATAAACGTTACGTCA	160	-151/+9	57	36	This study
	U	GGAATTATGGTTTGTTTATATGGTT	CCCAACCTATAAACATTACATCAC	167	-156/+11			
EPAS1	M	ATATATTCGCGTCGGTGTTC	CGCTCGCGAATATAAAACTC	118	-100/+18	56-57	37-38	This study
	U	TGTTTATATATTTGTGTTGGTGTTTG	CCACTCACAAATATAAAACTCCC	124	-105/+19			
CCDC181/ C1orf114	M	CGGTATTTCGCGAGTTTTTATAAC	CGAAAACGACAAAAATCTACG	164	-166/-3	57	35	This study
	U	TAGTGGTATTTTGTGAGTTTTTATAAT	ACAAAAACAACAAAAATCTACACA	168	-169/-2			

Table S3. Continued.

Gene symbol	Primer pair ID	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Product size, nt	Amplicon location from TSS	Primer annealing T, °C	Number of PCR cycles	Ref.
Protein-coding genes								
NEK9	M	TTAATAGATTCGAGAGGTCGTATC	CGCGAAACTCGATAATAACTC	253	-210/+43	57	38	This study
	U	GGGTTAATAGATTTGAGAGGTTGTATT	CCACAAAACTCAATAATAACTCCT	257	-213/+44			
NAALAD2	M	TATTTATTATGTTCGGGTTATTGC	CCTACTACATTCGCGAACTTC	244	-116/+128	58	35	This study
	U	GTTATTTATTATGTTTGGGTTATTGT	CCTACTACATTCACAAACTTCAA	246	-118/+128			
ZMIZ1	M	TCGTTTCGAAAATTTTTTAAATC	AACTCCCGAAACGCTATC	246	-188/+58	55	38	This study
	U	TGTAGTTTGTTTTGAAAATTTTTTAAA	AACTCCCAAAACACTATCACC	252	-194/+58			
ADAMTS12	M	GAGTTCGGGAGGAAGATGTATC	ACAACGACTACAAAACTACCCG	241	-195/+46	62	35	This study
	U	GAGTTTGGGAGGAAGATGTATT	AAACAACAACTACAAAACTACCA	243	-195/+48			
CD44	M	TCGTTGAGTTTGGCGTAGATC	ACTACCGCCGAATCCGCG	89	+480/+568	58	35	This study
	U	GTGTTGTTGAGTTTGGTGTAGATT	CAAAAAAACTACCACCAAATCCACA	99	+477/+575			
miRNA host genes								
MIR155HG/ BIC/ MIRHG2	M	GGTCGTACGTTCGTAGGC	ACGAAAACGCGAAACTAAAAT	178	-36/+142	56	35	This study
	U	AAGGGTTGTATGTTTGTAGGT	AAACAAAAACACAAAACTAAAATC	183	-39/+144			
COPZ2	M	TTCGAGGAAAGGGAGGTTAC	GTAACCAAACCTCGAACCG	136	-82/+54	56	37	This study
	U	GTTTTGAGGAAAGGGAGGTTAT	ACATAACCAAACCTCAAACCAC	140	-84/+56			
MIR137HG	M	TTCGTAAGGACGGTTGTTC	CAAACGTTTTCGTAAACGAA	124	-4417/-4294	55	38	This study
	U	GTTTTTGTAAGGATGGTTGTTT	AACCAAACATTTTCATAAACAAA	130	-4420/-4288			
GIPR/PGQTL2	M	CGTTTAGCGTAGTTTTAGGGTAATC	AACCGTATACACGCGAAACT	221	+914/+1134	56	38	This study
	U	TTTTGTTTAGTGTAGTTTTAGGGTAATT	AAACCATATACACACAAAACTCC	225	+911/+1135			
MIR31HG	M	TGCGGGAACGTTTATTTTC	CGAACGTAAAACCTACGAACC	127	-113/+14	58	36-37	This study
	U	AAGGTGTGGGAATGTTTATTTTT	CAAACATAAAACCTACAAACCCC	131	-117/+14			

TSS – transcription start site, M/ U – primers specific for methylated/ unmethylated sequence.

Table S4. Primers used for quantitative methylation-specific PCR and location of amplicons.

Gene symbol	Primer ID	Primer sequence (5'-3')	Amplicon size, bp	Amplicon location from TSS	Reference
ACTB	Fw	TGGTGATGGAGGAGGTTTAGTAAGT	133	-1629/-1497	[319]
	Rv	AACCAATAAAACCTACTCCTCCCTTAA			
	Probe	FAM-ACCACCACCCAACACACAATAACAAACAC A-BHQ-1			
RASSF1	Fw	GCGTTGAAGTCGGGGTTC	75	+45/+119	[319]
	Rv	CCCGTACTTCGCTAACTTTAAACG			
	Probe	FAM-ACAAACGCGAACCGAACGAAACCA-BHQ-1			
RARB	Fw	GGGATTAGAATTTTTTATGCGAGTTGT	92	+63/+154	[314]
	Rv	TACCCCGACGATACCCAAAC			
	Probe	FAM-TGTCGAGAACGCGAGCGATTCG-BHQ-1			
GSTP1-1	Fw	AGTTGCGCGGCGATTTC	140	+29/+168	[195]
	Rv	GCCCCAATACTAAATCACGACG			
	Probe	FAM-CGGTCGACGTTCGGGGTGTAGCG-BHQ-1			
GSTP1-2	Fw	GTCGGCGTCGTGATTTAGTATTG	100	+142/+241	[320]
	Rv	AAACTACGACGACGAAACTCCAA			
	Probe	FAM-AAACCTCGCGACCTCCGAACCTTATAAAA-BHQ-1			
CCDC181	Fw	GCGGTATTTCGCGAGTTTTTAT	131	-167/-37	This study
	Rv	TATCCTCAAACCACCGACC			
	Probe	FAM-AGTATCGGGATGGGTGTCGGGA-BHQ-1			
PRKCB	Fw	CGTAGTTGGGGTTAGCGGTG	145	-28/+117	This study
	Rv	AAAACGACGACCGCTACTACA			
	Probe	JOE-TTAGAGTCGGCGTAGGGGAAGCG-BHQ-1			
ADAMTS12	Fw	CGGGAGGAAGATGTATCGAGC	138	-190/-53	This study
	Rv	TCAACTAACAATATCCGCTTTCG			
	Probe	Cy5-TTTCGTTTTGGTTTATTTTATATTTCG-BHQ-3			

Two primer sets for *GSTP1* were tested at the initial stage of the study, data obtained using *GSTP1-1* primer set was used in the analysis.
TSS – transcription start site, Fw/ Rv – forward/ reverse primer.

Table S5. The list of genes analyzed in the present study.

No.	Gene symbol	Official gene name[1]	Genomic coordinates in GRCh38.p7 (strand)	Type of analysis
1	*RARB*	*Retinoic acid receptor beta*	Chr3:24829323..25597932 (+)	Methylation
2	*GSTP1*	*Glutathione S-transferase pi 1*	Chr11:67583595..67586653 (+)	Methylation
3	*RASSF1*	*Ras association domain family member 1*	Chr3:50329786..50340936 (-)	Methylation
4	*MGMT*	*O-6-methylguanine-DNA methyltransferase*	Chr10:129467184..129770983 (+)	Methylation
5	*DAPK1*	*Death associated protein kinase 1*	Chr9:87497228..87708634 (+)	Methylation
6	*$p14^{ARF}$/ CDKN2A*[2]	*Cyclin dependent kinase inhibitor 2A (transcript variant 4, encoding p14arf)*	Chr9:21967752..21995043 (-)	Methylation
7	*$p16^{INK4a}$/ CDKN2A*[2]	*Cyclin dependent kinase inhibitor 2A (transcript variant 1, encoding p16ink4a protein)*	Chr9:21967752..21995043 (-)	Methylation
8	*ACTB*	*Actin beta*	Chr7: 5527148..5530601 (-)	Methylation (EC)
9	*PRKCB*	*Protein kinase C beta*	Chr16:23835979..24220611 (+)	Methylation, gene expression
10	*CCDC181/ C1orf114*	*Coiled-coil domain containing 181*	Chr1:169394870..169462221 (-)	Methylation, gene expression
11	*ADAMTS12*	*ADAM metallopeptidase with thrombospondin type 1 motif 12*	Chr5:33523535..33892180 (-)	Methylation, gene expression
12	*ZMIZ1*	*Zinc finger MIZ-type containing 1*	Chr10:79068994..79316528 (+)	Methylation, gene expression
13	*NAALAD2*	*N-acetylated alpha-linked acidic dipeptidase 2*	Chr11:90131694..90193577 (+)	Methylation, gene expression
14	*HPRT1*	*Hypoxanthine phosphoribosyltransferase 1*	ChrX: 134460145..134500668 (+)	Gene expression (EC)
15	*FILIP1L*	*Filamin A interacting protein 1 like*	Chr3:99833144..100114513 (-)	Methylation
16	*KCTD8*	*Potassium channel tetramerization domain containing 8*	Chr4:44173903..44448885 (-)	Methylation
17	*EPAS1*	*Endothelial PAS domain protein 1*	Chr2:46297402..46386703 (+)	Methylation
18	*NEK9*	*NIMA related kinase 9*	Chr14:75082115..75127075 (-)	Methylation
19	*CD44*	*CD44 molecule (Indian blood group)*	Chr11:35138870..35232402 (+)	Methylation
20	*MIR137HG*	*MIR137 host gene*	Chr1:97988000..98049693 (-)	Methylation, miRNA expression
21	*MIR155HG/ BIC/ MIRHG2*	*MIR155 host gene*	Chr21:25562145..25575168 (+)	Methylation, miRNA expression
22	*COPZ2*	*Coatomer protein complex subunit zeta 2*	Chr17:48026167..48048091 (-)	Methylation, miRNA expression
23	*GIPR/ PGQTL2*	*Gastric inhibitory polypeptide receptor*	Chr19:45668187..45683724 (+)	Methylation
24	*MIR31HG*	*MIR31 host gene*	Chr9:21454268..21559698 (-)	Methylation, miRNA expression
25	*DNMT1*	*DNA methyltransferase 1*	Chr19:10133344..10195135 (-)	Gene expression
26	*KDM1A/ LSD1*	*Lysine demethylase 1A*	Chr1:23019443..23083691 (+)	Gene expression
27	*KDM5B/ JARID1B*	*Lysine demethylase 5B*	Chr1:202725185..202809470 (-)	Gene expression

EC – endogenous control.

[1]The provided gene names are approved by HUGO (Human Genome Organization) Genome Nomenclature Committee.

[2]Transcripts 1 and 4, encoding p16INK4a and p14ARF proteins, have distinct first exons which contain the translation start codon, and share a common second exon, which is translated in different reading frames.

Table S6. Associations of promoter methylation frequencies of tumor suppressor genes with clinical-pathological patients' characteristics and *TMPRSS2-ERG* fusion status in prostate tumor tissues and urine from patients diagnosed with the malignancy.

Tumor suppressor gene	Discrete variables					
	Tumor stage		Gleason score		*TMPRSS2-ERG* status	
	pT3 vs. ≤pT2, %	P	≥7 vs. 6, %	P	Yes vs. no, %	P
Tissue						
RARB	77.1 vs. 61.4	0.1063	69.1 vs. 59.5	0.2937	71.7 vs. 69.4	>0.9999
RASSF1	68.6 vs. 49.6	0.0544	58.8 vs. 50.7	0.3980	67.4 vs. 66.7	>0.9999
GSTP1	74.3 vs. 50.9	**0.0190**	66.2 vs. 43.2	**0.0072**	54.4 vs. 52.8	>0.9999
MGMT	34.3 vs. 22.8	0.1881	25.0 vs. 24.3	>0.9999	23.9 vs. 19.4	0.7892
DAPK1	20.0 vs. 7.9	0.0596	11.8 vs. 9.5	0.7864	8.7 vs. 0.0	0.1270
p14	8.6 vs. 6.1	0.7000	10.3 vs. 4.1	0.1947	6.5 vs. 5.6	>0.9999
p16	5.7 vs. 8.8	0.7326	7.4 vs. 9.5	0.7670	8.7 vs. 8.3	0.6364
Urine						
RARB	24.6 vs. 30.6	0.4126	31.4 vs. 26.8	0.4794	30.1 vs. 25.5	0.5698
RASSF1	52.6 vs. 42.3	0.1768	45.8 vs. 43.1	0.6988	51.8 vs. 56.4	0.6067
GSTP1	17.5 vs. 9.2	0.0929	16.9 vs. 5.7	**0.0073**	12.1 vs. 20.0	0.2314

Tumor suppressor gene	Continuous variables					
	PSA		Prostate mass		Age	
	$Z_{ad.}$	P	$Z_{ad.}$	P	$Z_{ad.}$	P
Tissue						
RARB	-0.36	0.7213	-2.57	**0.0101**	-1.54	0.0928
RASSF1	2.58	**0.0098**	-1.46	0.1451	0.07	0.9416
GSTP1	1.48	0.1392	-0.48	0.6325	-0.52	0.6325
MGMT	1.85	0.0648	1.34	0.1805	1.49	0.1374
DAPK1	-0.19	0.8509	0.63	0.5307	0.30	0.7611
p14	-0.36	0.7215	-0.53	0.5953	-1.27	0.2031
p16	-0.15	0.8839	-1.05	0.2951	0.50	0.6149
Urine						
RARB	-0.77	0.4399	0.21	0.8300	-0.82	0.4114
RASSF1	-0.15	0.8810	0.36	0.7192	1.54	0.1224
GSTP1	1.87	0.0612	-1.36	0.1740	-0.84	0.3985

ISUP – International Society of Urological Pathology, PSA – prostate-specific antigen, Z_{ad} – Mann-Whitney's Z adjusted parameter. Significant P-values are in bold.

Table S7. Cox regression analysis of biochemical prostate cancer recurrence including selected clinical-pathological variables and gene methylation markers analyzed in tumor tissue and urine.

No.	Covariates	HR (95% CI)	Covariate's P-value	Model's P-value
Univariate models (tissue, all cases)				
1	pT (3 vs. ≤2)	3.91 (2.29; 6.66)	**<0.0001**	**<0.0001**
2	Gleason score (≥7 vs. 6)	2.51 (1.54; 4.10)	**0.0002**	**0.0005**
3	Prostate mass (cont.)	1.00 (0.99; 1.01)	0.4916	0.5050
4	PSA (cont.)	1.02 (1.00; 1.04)	0.0544	0.0945
5	*TMPRSS2-ERG* (yes vs. no)	1.06 (0.50; 2.25)	0.8716	0.8709
6	Age (cont.)	1.03 (0.99; 1.07)	0.1454	0.1403
7	*RASSF1* (M vs. U)	2.27 (1.12; 4.63)	**0.0242**	**0.0194**
8	*DAPK1* (M vs. U)	2.55 (1.11; 5.84)	**0.0276**	**0.0454**
9	*RASSF1* and/ or *DAPK1* (at least one M vs. both U)	2.20 (1.06; 4.54)	**0.0348**	**0.0268**
Multivariate models (Gleason 6 cases only)				
Tissue				
10	pT (3 vs. ≤2)	15.46 (3.35; 71.23)	**0.0005**	**0.0007***
	RASSF1 (M vs. U)	5.81 (1.08; 31.22)	**0.0415**	
Urine				
11	pT (3 vs. ≤2)	10.59 (2.49; 45.08)	**0.0015**	**0.0003***
	RASSF1 (M vs. U)	8.78 (2.53; 30.46)	**0.0007**	

M/U – methylated/ unmethylated promoter status, cont. – continuous variable, *TMPRSS2-ERG* – fusion transcript status, HR – hazard ratio, CI – confidence intervals, pT – tumor stage, PSA – prostate-specific antigen. Significant P-values are in bold.
*Backward entering of covariates.

Table S8. Genes showing the most significant differences of methylation levels comparing prostate tumors vs. noncancerous tissues. Values are given for the most significant probe per gene/ annotated locus.

No.	Gene/ locus	Probe location	Methylation	FC	P-value
1	*MAGI2*	Intragenic	Up	1.5	<0.0001
2	*PNMA6A*	Intragenic	Down	1.6	<0.0001
3	*CCDC181*	Promoter	Up	1.5	<0.0001
4	*C8orf33*	Intragenic	Down	1.5	<0.0001
5	*NAALAD2*	Promoter	Up	1.5	<0.0001
6	*FGD5*	Promoter	Up	1.6	<0.0001
7	*MTA1*	Intragenic	Down	1.5	0.0001
8	*HOXA7*	Intragenic	Up	1.6	0.0001
9	*ABCD1*	Intragenic	Down	1.5	0.0001
10	*GPR149*	Intragenic	Up	1.5	0.0001
11	*CEBPG*	Intragenic	Down	1.5	0.0001
12	*KCNH7*	Intragenic	Up	1.6	0.0001
13	*CDKN2A*	Intragenic	Up	1.5	0.0001
14	*POU4F2*	Intragenic	Up	1.5	0.0001
15	*PCSK1*	Intragenic	Up	1.5	0.0001
16	*ADAMTS12*	Intragenic	Up	1.5	0.0001
17	*SLC6A17*	Intragenic	Up	1.6	0.0001
18	*VSTM2B*	Intragenic	Up	1.5	0.0001
19	*EFCAB4A*	Intragenic	Up	1.5	0.0002
20	*BTBD6*	Intragenic	Down	1.5	0.0002
21	*SPON2*	Intragenic	Down	1.5	0.0002
22	*BDNF*	Promoter	Up	1.5	0.0002
23	*CALCR*	Promoter	Up	1.6	0.0002
24	*HCN1*	Intragenic	Up	1.5	0.0002
25	*GNAS*	Promoter	Down	1.7	0.0002
26	*TERT*	Intragenic	Down	1.6	0.0002
27	*FGF5*	Intragenic	Up	1.5	0.0002
28	*STK40*	Intragenic	Down	1.7	0.0002
29	*RHOU*	Promoter	Down	1.5	0.0002
30	*RASAL1*	Intragenic	Down	1.5	0.0002
31	*TRMT2B*	Intragenic	Down	1.8	0.0002
32	*RAI1*	Intragenic	Up	1.5	0.0002
33	*PARD6G*	Promoter	Down	1.5	0.0002
34	*ARHGAP36*	Intragenic	Up	1.5	0.0002
35	*FLJ44511*	Intragenic	Up	1.5	0.0002
36	*SNTG1*	Promoter	Up	1.6	0.0002
37	*PCGF3*	Intragenic	Down	1.5	0.0002
38	*SHANK3*	Intragenic	Down	1.6	0.0003
39	*FUT6*	Intragenic	Down	1.5	0.0003
40	*MGRN1*	Intragenic	Down	1.5	0.0003
41	*AGAP3*	Intragenic	Down	1.5	0.0003
42	*TAL1*	Promoter	Up	1.7	0.0003
43	*IRF2BP1*	Intragenic	Down	1.6	0.0003
44	*TNRC18*	Promoter	Up	1.5	0.0003
45	*KCNJ6*	Intragenic	Up	1.5	0.0003
46	*SPATA6*	Intragenic	Up	1.5	0.0004
47	*CDC34*	Intragenic	Down	1.6	0.0004
48	*ZNF865*	Intragenic	Down	1.5	0.0004
49	*FASN*	Intragenic	Down	1.6	0.0004
50	*TLX2*	Intragenic	Up	1.5	0.0004

FC – absolute fold change value.

Table S9. Genes showing the most significant differences of methylation levels comparing biochemical disease recurrence (BCR)-positive *vs.* BCR-negative cases. Values are given for the most significant probe per gene/ annotated locus.

No.	Gene/ locus	Probe location	Methylation	FC	P-value
1	*COL23A1*	Intragenic	Up	1.3	<0.0001
2	*HOXC9*	Promoter	Up	1.2	<0.0001
3	*MATN4*	Promoter	Up	1.2	<0.0001
4	*POU4F3*	Intragenic	Up	1.3	<0.0001
5	*MECOM*	Promoter	Down	1.2	<0.0001
6	*CAMTA1*	Intragenic	Up	1.4	<0.0001
7	*C7orf51*	Intragenic	Up	1.3	0.0001
8	*BMP8A*	Intragenic	Up	1.2	0.0001
9	*RAPGEFL1*	Promoter	Up	1.3	0.0001
10	*BARHL1*	Promoter	Up	1.2	0.0001
11	*C1orf170*	Promoter	Up	1.5	0.0001
12	*TMEM233*	Intragenic	Up	1.2	0.0001
13	*C1D*	Intragenic	Down	1.2	0.0002
14	*NFATC1*	Promoter	Down	1.3	0.0002
15	*CLDN5*	Intragenic	Up	1.3	0.0002
16	*PLK1*	Intragenic	Up	1.3	0.0002
17	*CACNG7*	Intragenic	Up	1.3	0.0002
18	*GSX2*	Intragenic	Up	1.2	0.0002
19	*C19orf20*	Intragenic	Up	1.3	0.0002
20	*EPHA10*	Intragenic	Up	1.3	0.0002
21	*SPRY4*	Intragenic	Down	1.2	0.0002
22	*WNT10B*	Intragenic	Up	1.3	0.0002
23	*TRIM54*	Intragenic	Up	1.2	0.0002
24	*SV2A*	Intragenic	Up	1.4	0.0003
25	*HOPX*	Intragenic	Up	1.2	0.0003
26	*RSPH9*	Intragenic	Up	1.3	0.0003
27	*HOXB4*	Intragenic	Up	1.3	0.0003
28	*SYP*	Intragenic	Up	1.2	0.0003
29	*C2orf88*	Promoter	Up	1.2	0.0003
30	*CALY*	Intragenic	Up	1.3	0.0003
31	*FAM43B*	Intragenic	Up	1.2	0.0003
32	*IGFBP7*	Intragenic	Up	1.4	0.0003
33	*PLCXD1*	Promoter	Down	1.3	0.0004
34	*WBP2*	Promoter	Down	1.2	0.0004
35	*SCNN1B*	Promoter	Up	1.2	0.0004
36	*LOC400043*	Intragenic	Up	1.2	0.0004
37	*ZNF778*	Intragenic	Up	1.2	0.0004
38	*MMP25*	Intragenic	Up	1.3	0.0004
39	*LOC100287042*	Intragenic	Down	1.3	0.0004
40	*NRXN2*	Intragenic	Up	1.4	0.0005
41	*EXOC3L2*	Intragenic	Up	1.3	0.0005
42	*INSM1*	Promoter	Up	1.3	0.0005
43	*SEPT9*	Promoter	Up	1.3	0.0005
44	*C1orf190*	Promoter	Up	1.2	0.0005
45	*SPOCK2*	Intragenic	Up	1.2	0.0005
46	*ALLC*	Promoter	Up	1.6	0.0005
47	*MIR152*	Promoter	Up	1.4	0.0005
48	*FAM100A*	Promoter	Down	1.2	0.0005
49	*SPTBN4*	Promoter	Up	1.3	0.0005
50	*ZNF471*	Promoter	Down	1.3	0.0005

FC – absolute fold change value.

Table S10. MiRNA host genes showing the most significant differences of methylation levels comparing prostate tumors *vs.* noncancerous tissues. Values are given for the most significant probe per gene.

No.	Gene/ locus	pre-miRNA	Methylation	FC	P-value
1	*HOXC4*	mir-615	up	1.3	<0.0001
2	*ARMC9*	mir-4777	down	1.2	<0.0001
3	*LINGO2*	mir-873, mir-876	up	1.6	<0.0001
4	*PCDH15*	mir-548f-1	up	1.5	<0.0001
5	*ADGRB2*	mir-4254	up	1.3	<0.0001
6	*LOC101929710*	mir-583	up	1.3	0.0001
7	*MIR9-3HG*	mir-9-3	up	1.4	0.0001
8	*ANKRD34C-AS1*	mir-184	up	1.3	0.0001
9	*MIR124-2HG*	mir-124-2	up	1.3	0.0001
10	*PACSIN3*	mir-6745	up	1.3	0.0001
11	*HOXD3*	mir-10b	up	1.3	0.0001
12	*CADM2*	mir-5688	up	1.4	0.0001
13	*C1orf61*	mir-91-1	up	1.3	0.0001
14	*TAF4*	mir-3195	down	1.2	0.0001
15	*RGS6*	mir-7843	up	1.4	0.0001
16	*AEN*	mir-7-2, mir-1179, mir-3529	up	1.3	0.0001
17	*CHRM2*	mir-490	up	1.3	0.0001
18	*OSBP2*	mir-3200	up	1.2	0.0001
19	*KIFC3*	mir-6772	up	1.2	0.0001
20	*DCC*	mir-4528	up	1.3	0.0001
21	*RPL28*	mir-6805	up	1.4	0.0002
22	*CALCR*	mir-489, mir-653	up	1.6	0.0002
23	*CTNNA2*	mir-4264, mir-8080	up	1.4	0.0002
24	*SCUBE2*	mir-5691	up	1.4	0.0002
25	*LRP1B*	mir-7157	up	1.3	0.0002
26	*NOTCH3*	mir-6795	down	1.3	0.0002
27	*FBRSL1*	mir-6763	down	2.0	0.0003
28	*SREBF2*	mir-33a	down	1.5	0.0003
29	*TTC28*	mir-5739	up	1.3	0.0003
30	*MIR31HG*	mir-31	up	1.3	0.0003
31	*MRE11A*	mir-548l	up	1.2	0.0004
32	*DSCAM*	mir-4760	up	1.2	0.0004
33	*PAX5*	mir4476, mir-4540	up	1.3	0.0004
34	*BCL3*	mir-8085	down	1.4	0.0004
35	*SFRP1*	mir-548ao	up	1.2	0.0004
36	*CPNE4*	mir-5704	up	1.4	0.0004
37	*FOXF2*	mir-6720	up	1.3	0.0005
38	*MIR137HG*	mir-137, mir-2682	up	1.5	0.0005
39	*PKD1*	mir-1225, mir-4516, mir-4616, mir-6511b	down	1.5	0.0005
40	*MEIS2*	mir-8063	up	1.3	0.0005
41	*MAP7D2*	mir-23c	up	1.2	0.0005
42	*SLC12A7*	mir-4635	down	1.6	0.0005
43	*CAMK1D*	mir-548q, mir-4480, mir-4481	down	1.3	0.0005
44	*IGF2*	mir-483	up	1.3	0.0005
45	*TRPM3*	mir-204	up	1.3	0.0005
46	*TENM3*	mir-1305	up	1.4	0.0005
47	*ATF5*	mir-4751	down	1.2	0.0005
48	*PTPRN2*	mir-153-2, mir-595	up	1.4	0.0006
49	*JAK1*	mir-101-1, mir-3671	down	1.3	0.0006
50	*KCNT2*	mir-4735	up	1.3	0.0006

FC – absolute fold change value.

Table S11. MiRNA host genes showing the most significant differences of methylation levels comparing biochemical disease recurrence (BCR)-positive *vs.* BCR-negative prostate cancer cases. Values are given for the most significant probe per gene.

No.	Gene/ locus	pre-miRNA	Methylation	FC	P-value
1	*RAPGEFL1*	mir-6867	up	1.3	0.0001
2	*HOXB3*	mir-10a	up	1.3	0.0003
3	*SEPT9*	mir-4316	up	1.3	0.0005
4	*COPZ2*	mir-152	up	1.4	0.0005
5	*PRKCB*	mir-1273h	up	1.3	0.0006
6	*TAF4*	mir-3195	down	1.5	0.0007
7	*USP15*	mir-6125	down	1.2	0.0008
8	*ESRP2*	mir-6773	down	1.2	0.0010
9	*TRIM28*	mir-6807	up	1.2	0.0011
10	*HDAC4*	mir-2467, mir-4269, mir-4440, mir-4441	down	1.3	0.0011
11	*S1PR2*	mir-4322	up	1.2	0.0012
12	*ESPN*	mir-4252	up	1.3	0.0018
13	*CELSR3*	mir-4793	up	1.3	0.0019
14	*GRID1*	mir-346	down	1.3	0.0021
15	*CACNG8*	mir-935	up	1.2	0.0022
16	*ARL10*	mir-1271	up	1.3	0.0023
17	*IGF2*	mir-483	down	1.4	0.0023
18	*COL18A1*	mir-6815	down	1.2	0.0030
19	*NPR1*	mir-8083	up	1.2	0.0032
20	*FSTL4*	mir-1289	down	1.2	0.0033
21	*CPSF6*	mir-1279	down	1.4	0.0033
22	*ROBO1*	mir-3923	down	1.4	0.0034
23	*MAST1*	mir-6794	up	1.3	0.0035
24	*LOC100130691*	mir-6512	up	1.4	0.0035
25	*RBPMS2*	mir-1272	up	1.3	0.0036
26	*TENM3*	mir-1305	down	1.4	0.0036
27	*H19*	mir-675	down	1.2	0.0037
28	*GNG7*	mir-7850	up	1.2	0.0039
29	*SNCB*	mir-4281	up	1.3	0.0042
30	*PTPRN2*	mir-153-2, mir-595	down	1.3	0.0042
31	*GRK5*	mir-4681	down	1.3	0.0045
32	*MEST*	mir-335	down	1.3	0.0050
33	*ELMO1*	mir-1200	down	1.4	0.0051
34	*OPLAH*	mir-6846	up	1.2	0.0058
35	*EIF3CL*	mir-6862	up	1.3	0.0064
36	*IGF1R*	mir-4714	down	1.2	0.0065
37	*DLEU2*	mir-15a, mir-16-1, mir-3613	down	1.4	0.0068
38	*BZRAP1-AS1*	mir-142, mir-4736	down	1.2	0.0073
39	*KDM2B*	mir-7107	up	1.3	0.0074
40	*ANKRD28*	mir-3134	down	1.3	0.0085
41	*STXBP1*	mir-3911	up	1.3	0.0086
42	*SLC7A5*	mir-6775	down	1.2	0.0087
43	*GABRE*	mir-224, mir-452	down	1.3	0.0087
44	*FOXF2*	mir-6720	down	1.2	0.0089
45	*FOXP2*	mir-3666	down	1.2	0.0090
46	*RYR2*	mir-4428	down	1.3	0.0091
47	*NCOR2*	mir-6880	down	1.2	0.0093
48	*LASP1*	mir-6779	down	1.2	0.0102
49	*ANK1*	mir-486-1, mir-486-2	up	1.3	0.0103
50	*MAP7D2*	mir-23c	up	1.4	0.0103

FC – absolute fold change value.

Table S12. Associations of protein-coding gene promoter methylation with continuous clinical-pathological variables in prostate tumors.

Protein-coding gene	PSA		Prostate mass		Age	
	Z_{ad}	P	Z_{ad}	P	Z_{ad}	P
PRKCB	1.51	0.1305	-0.77	0.4433	0.19	0.8515
CCDC181	-0.02	0.9828	0.43	0.6705	1.00	0.3193
ADAMTS12	0.91	0.3621	-0.42	0.6743	1.20	0.2320
ZMIZ1	2.06	**0.0390**	-0.14	0.8855	1.61	0.1076
FILIP1L	0.02	0.9814	0.04	0.9660	0.00	0.9974
NAALAD2	1.64	0.1015	-0.34	0.7303	0.52	0.6010
CD44	-1.16	0.2469	0.24	0.8095	1.95	0.0510
KCTD8	0.54	0.5859	-0.99	0.3244	0.48	0.6332

PSA – prostate-specific antigen, Z_{ad} – Mann-Whitney's Z adjusted parameter.
Significant P-values are in bold.

Table S13. Associations of protein-coding gene expression with clinical-pathological variables and *TMPRSS2-ERG* fusion status in prostate tumors.

Protein-coding gene	Discrete variables					
	Tumor stage pT		ISUP grade group		*TMPRSS2-ERG* status	
	Z_{ad}	P	H	P	Z_{ad}	P
PRKCB	-0.45	0.6528	2.33	0.5075	0.17	0.8674
CCDC181	-0.65	0.5154	9.58	**0.0225**	2.47	**0.0136**
ADAMTS12	-0.35	0.7248	2.70	0.4396	0.97	0.3323
ZMIZ1	-0.29	0.7692	3.04	0.3856	-0.09	0.9313
NAALAD2	-1.07	0.2864	10.63	**0.0139**	-0.03	0.9742
Protein-coding gene	Continuous variables					
	PSA		Prostate mass		Age	
	R_S	P	R_S	P	R_S	P
PRKCB	-0.17	0.1362	0.00	0.9815	0.13	0.2456
CCDC181	-0.20	0.0755	-0.01	0.9482	0.15	0.1698
ADAMTS12	-0.12	0.2958	0.05	0.6676	0.01	0.9334
ZMIZ1	0.11	0.3433	0.03	0.8052	0.02	0.8385
NAALAD2	-0.19	0.0972	-0.03	0.8003	0.27	**0.0153**

ISUP – International Society of Urological Pathology, PSA – prostate-specific antigen, Z_{ad} – Mann-Whitney's Z adjusted parameter, H – Kruskal-Wallis's H parameter, R_S – Spearman's correlation coefficient.
Significant P-values are in bold.

Table S14. Univariate Cox proportional hazard analysis of molecular and/or clinical-pathological variables in Lithuanian and The Cancer Genome Atlas (TCGA) prostate cancer cohorts. For simplicity, only models' P-values are provided for TCGA data.

No.	Covariates	Lithuanian cohort (BCR-free)			TCGA cohort (BCR-free)	TCGA cohort (disease-free)
		HR (95% CI)	Covariate's P-value	Model's P-value	Model's P-value	Model's P-value
Promoter methylation						
1	*PRKCB* (M *vs.* U/ cont.)	5.45 (1.31; 22.66)	**0.0203**	**0.0028**	0.0795	0.2302
2	*CCDC181* (M *vs.* U/ cont.)	3.40 (0.47; 24.67)	0.2294	0.1411	0.0702	0.7190
3	*ADAMTS12* (M *vs.* U/ cont.)	>1000	0.9499	**0.0003**	0.0624	0.3107
4	*ZMIZ1* (M *vs.* U/ cont.)	3.10 (0.75; 12.92)	0.1215	0.0670	0.6455	0.3578
5	*FILIP1L* (M *vs.* U/ cont.)	1.90 (0.58; 6.24)	0.2924	0.2517	0.1497	0.0561
6	*NAALAD2* (M *vs.* U/ cont.)	7.81 (1.87; 32.69)	**0.0051**	**0.0002**	0.0911	0.1695
7	*KCTD8* (M *vs.* U/ cont.)	1.41 (0.65; 3.03)	0.3865	0.3990	0.9625	0.8106
8	*CD44* (M *vs.* U/ cont.)	1.00 (0.48; 2.07)	0.9973	0.9957	0.1054	0.3811
9	mir-155 (M *vs.* U/ cont.)	2.19 (0.98; 4.91)	0.0574	**0.0447**	0.2564	0.3231
10	mir-152 (M *vs.* U/ cont.)	2.63 (1.22; 5.67)	**0.0142**	**0.0236**	0.4708	0.7989
11	mir-137 (M *vs.* U/ cont.)	1.48 (0.74; 2.96)	0.2753	0.2806	0.0928	**0.0122**
12	mir-31 (M *vs.* U/ cont.)	3.93 (1.51; 10.23)	**0.0052**	**0.0168**	0.4990	**0.0188**
Gene expression						
13	*PRKCB* (cont.)	0.60 (0.27; 1.34)	0.2118	0.1836	**0.0378**	0.3147
14	*CCDC181* (cont.)	0.00 (0.00; 28.14)	0.1329	0.1111	0.1856	0.3035
15	*ADAMTS12* (cont.)	2.26 (0.02; 262.82)	0.7375	0.7387	**0.0321**	0.2170
16	*ZMIZ1* (cont.)	0.94 (0.82; 1.09)	0.4090	0.3960	0.0548	0.0703
17	*NAALAD2* (cont.)	0.00 (0.00; 0.17)	**0.0205**	**0.0059**	0.7077	0.9964
18	*DNMT1* (cont.)	0.09 (0.00; 7.16)	0.2842	0.2648	0.2181	**0.0033**
19	*KDM1A* (cont.)	0.95 (0.84; 1.06)	0.3664	0.3555	0.0894	0.5037
20	*KDM5B* (cont.)	0.96 (0.75; 1.22)	0.7331	0.7309	**0.0093**	0.9132
Other molecular and clinical-pathological variables						
21	*TMPRSS2-ERG* (yes *vs.* no)	0.70 (0.34; 1.44)	0.3328	0.3375	n.a.	n.a.
22	pT (3 *vs.* 2)	4.32 (2.08; 8.94)	**0.0001**	**0.0001**	**0.0001**	**0.0002**
23	ISUP grade group (5 groups)	3.23 (2.09; 4.99)	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
24	Prostate mass (cont.)	1.01 (1.00; 1.03)	0.1221	0.1461	n.a.	n.a.
25	PSA (cont.)	1.02 (1.00; 1.04)	0.0621	0.1048	0.4881	0.4723
26	Age (cont.)	1.00 (0.95; 1.05)	0.9470	0.9462	0.4655	0.6256

M/U – methylated/ unmethylated promoter status, cont. – continuous variable, *TMPRSS2-ERG* – fusion transcript status, pT – tumor stage, ISUP – International Society of Urological Pathology, PSA – prostate-specific antigen, HR – hazard ratio, CI – confidence intervals. Significant P-values are in bold.

*Forward entering of covariates.

Table S15. Multivariate Cox proportional hazard analysis of protein-coding gene methylation and other molecular and/or clinical-pathological variables in Lithuanian and The Cancer Genome Atlas (TCGA) prostate cancer cohorts. Selected combinations are shown. For simplicity, only models' P-values are provided for TCGA data.

No.	Covariates	Lithuanian cohort (BCR-free)			TCGA cohort (BCR-free)	TCGA cohort (disease-free)
		HR (95% CI)	Covariate's P-value	Model's P-value	Model's P-value	Model's P-value
Combinations of gene promoter methylation						
1	*PRKCB* (M *vs.* U/ cont.)	3.02 (0.73; 12.54)	0.1306	**0.0003**	0.1353	0.4563
	ADAMTS12 (M *vs.* U/ cont.)	>1000	0.9516			
2	*PRKCB* (M *vs.* U/ cont.)	3.22 (0.76; 13.69)	0.1157	**0.0002**	0.1314	0.3260
	NAALAD2 (M *vs.* U/ cont.)	5.77 (1.34; 24.90)	**0.0195**			
3	*ADAMTS12* (M *vs.* U/ cont.)	>1000	0.9528	**0.0001**	0.1132	0.3580
	NAALAD2 (M *vs.* U/ cont.)	4.44 (1.05; 18.73)	**0.0434**			
4	*PRKCB* (M *vs.* U/ cont.)	2.83 (0.68; 11.80)	0.1542	**0.0001**	0.2001	0.5197
	ADAMTS12 (M *vs.* U/ cont.)	>1000	0.9528			
	NAALAD2 (M *vs.* U/ cont.)	4.33 (1.02; 18.45)	**0.0488**			
Combinations with gene expression						
5	*NAALAD2* (M *vs.* U/ cont.)	5.10 (1.18; 22.08)	**0.0303**	**0.0007**	0.2267	0.3891
	NAALAD2 (expression, cont.)	0.00 (0.00; 2.70)	0.0828			
Combinations with fusion transcript status						
6	*TMPRSS2-ERG* (yes *vs.* no)	1.61 (0.80; 3.27)	0.1872	**0.0113**	n.a.	n.a.
	PRKCB (M *vs.* U/ cont.)	5.30 (1.26; 22.25)	**0.0233**			
7	*TMPRSS2-ERG* (yes *vs.* no)	1.97 (0.96; 4.04)	0.0646	**0.0003**	n.a.	n.a.
	ADAMTS12 (M *vs.* U/ cont.)	>1000	0.9484			
8	*TMPRSS2-ERG* (yes *vs.* no)	1.48 (0.73; 3.00)	0.2748	**0.0003**	n.a.	n.a.
	NAALAD2 (M *vs.* U/ cont.)	8.57 (2.04; 35.92)	0.0035			
Combinations with clinical-pathological variables						
9	pT (3 *vs.* 2)	3.77 (1.82; 7.84)	**0.0004**	**<0.0001***	**0.0002**	**0.0007**
	PRKCB (M *vs.* U/ cont.)	4.36 (1.04; 18.25)	**0.0450**			
10	pT (3 *vs.* 2)	3.85 (1.86; 7.98)	**0.0003**	**<0.0001***	**0.0001**	**0.0005**
	NAALAD2 (M *vs.* U/ cont.)	6.87 (1.63; 28.84)	**0.0089**			
11	ISUP grade group (5 groups)	2.80 (1.80; 4.35)	**<0.0001**	**<0.0001***	**<0.0001**	**<0.0001**
	NAALAD2 (M *vs.* U/ cont.)	5.51 (1.29; 23.53)	**0.0219**			
12	PSA (cont.)	1.02 (1.00; 1.05)	**0.0463**	**0.0031***	0.5813	0.6454
	PRKCB (M *vs.* U/ cont.)	5.34 (1.28; 22.29)	**0.0221**			
	ZMIZ1 (M *vs.* U/ cont.)	5.78 (0.79; 42.25)	0.0855			

Table S15. Continued.

No.	Covariates	Lithuanian cohort (BCR-free)			TCGA cohort (BCR-free)	TCGA cohort (disease-free)
		HR (95% CI)	Covariate's P-value	Model's P-value	Model's P-value	Model's P-value
Combinations with clinical-pathological variables						
13	PSA (cont.)	1.02 (1.00; 1.04)	0.0996	**0.0163**	0.7824	0.5044
	ZMIZ1 (M *vs.* U/ cont.)	5.78 (0.79; 42.25)	0.0855			
14	Prostate mass (cont.)	1.02 (1.00; 1.04)	0.0505	**0.0027**	n.a.	n.a.
	PRKCB (M *vs.* U/ cont.)	5.65 (1.35; 23.58)	**0.0180**			
15	Prostate mass (cont.)	1.02 (1.00; 1.04)	**0.0491**	**0.0003**	n.a.	n.a.
	ADAMTS12 (M *vs.* U/ cont.)	>1000	0.9502			
16	Prostate mass (cont.)	1.02 (1.00; 1.04)	**0.0462**	**0.0096**	n.a.	n.a.
	ZMIZ1 (M *vs.* U/ cont.)	6.86 (0.94; 50.15)	0.0593			
17	Prostate mass (cont.)	1.02 (1.00; 1.04)	0.0932	**0.0003**	n.a.	n.a.
	NAALAD2 (M *vs.* U/ cont.)	7.56 (1.80; 31.66)	**0.0059**			

M/U – methylated/ unmethylated promoter status, *TMPRSS2-ERG* – fusion transcript status, pT – tumor stage, ISUP – International Society of Urological Pathology, PSA – prostate-specific antigen, HR – hazard ratio, CI – confidence intervals. Significant P-values are in bold, n.a. – not available/ not analyzed.

*Forward entering of covariates.

Table S16. Associations of miRNA host gene promoter methylation status with continuous clinical-pathological parameters in prostate tumors.

miRNA host gene	PSA		Prostate mass		Age	
	Z_{ad}	P-value	Z_{ad}	P-value	Z_{ad}	P-value
mir-155	1.05	0.2945	1.40	0.1615	1.65	0.0996
mir-152	2.21	**0.0275**	1.16	0.2458	0.55	0.5794
mir-137	-0.64	0.5207	0.05	0.9608	-0.30	0.7673
mir-31	0.09	0.9295	2.02	**0.0430**	1.13	0.2583

PSA – prostate-specific antigen, Z_{ad} – Mann-Whitney's Z adjusted parameter.
Significant P-values are in bold.

Table S17. Numbers of genes targeted by selected miRNAs as determined by the online tools with a focus on epigenetic regulation-associated targets (EFs) [260].

Numbers of identified targets*						
Source	miR-155-5p		miR-152-3p		miR-137	
	Total # of targeted EFs	Total # of targets	Total # of targeted EFs	Total # of targets	Total # of targeted EFs	Total # of targets
miRDB [321] (http://mirdb.org)	25	311	36	559	44	544
TargetScan 7.1 [322] (http://www.targetscan.org)	48	552	60	795	92	1305
miRTarBase [323] (http://mirtarbase.mbc.nctu.edu.tw)	64	920	9	131	9	122
miRWalk 2.0 [324] (http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/)	85	1384	28	783	57	1060
microRNA.org [325] (http://www.microrna.org)	158	5445	138	7389	148	5487
PicTar [326] (http://pictar.mdc-berlin.de)	9	184	29	340	28	357

*Database search was performed in April 2016.

Table S18. Genes defined as epigenetic regulation-associated targets (EFs) [260] of the selected miRNAs that were determined by the online tools.

Source*	miR-155-5p	miR-152-3p	miR-137
Determined by at least 4 of 6 databases	*AICDA, ARID2, BRD1, CDC73, CHAF1A, CHD7, CHD8, CHD9, YWHAE, YWHAZ, JARID2, KDM2A, KDM3A,* ***KDM5B****, MASTL, PHC2, PSIP1, RCOR1, SAP30L, SATB1, SIRT1, SMARCA4, SP1, TADA2B, TAF5L, TLE4*	*BRPF1, CHD7, CHD9, CLOCK, CUL5,* ***DNMT1****, GADD45A, INO80, YWHAB, JARID2, KDM2B, LBR, MLLT10, PPARGC1A, PRKAA1, PRKAG2, RPS6KA5, SIRT7, SKP1, TAF4, UBE2D1, ZNF217*	*ARID4B, BAZ1A, BAZ2A, BRD1, CHD7, CHD9, CTBP1, DR1, E2F6, HCFC2, HLTF, HMGN3, JDP2,* ***KDM1A****, KDM2A, KDM4A,* ***KDM5B****, MBD6, MYO1C, NCOA2, NCOA3, NFYB, PPARGC1A, PRDM1, PRKAB1, SYNCRIP, TAF12, TDRD7, TRIM33, UBE2H, ZNF217, ZNF516*
Determined by at least 2 of 6 databases	*ACTR8, ANKRD32, ARID4A, ARNTL, ATAD2B, AURKA, AURKB, BCORL1, BRPF3, BRWD3, C11orf30, CDK2, CDK5, CHAF1B, CHRAC1, CHTOP, CUL4B, DDB2, DEK, DNAJC2,* ***DNMT1****, DR1, EYA2, EPC2, EXOSC2, EZH1, GATAD2B, HCFC2, YEATS2, ING3, KANSL1, KAT2A,* ***KDM1A****, MAP3K7, MBTD1, MECP2, MOV10, MSH6, MSL3, NASP, NFYC, PAK2, PBRM1, PHF14, POLE3, PPM1G, PRMT2, RAD51, RING1, RNF2, RPS6KA3, RPS6KA5, RSF1, SATB2, SETD7, SETD8, SF3B1, SIN3A, SMARCAD1, SMARCD2, SMARCE1, SRSF1, SUZ12, TRIM24, UBE2D3, UBE2H, USP44, WDR82, WHSC1L1, ZBTB16, ZMYM2, ZMYND11*	*BAZ2B, CDK1, CHD1, CHUK, EYA3, KAT7, KMT2A, NAP1L1, NCOA1, PHF20, SYNCRIP, TBL1XR1, TET2, UBE2D3, WHSC1, BAZ2A, BRWD3, EP300, GATAD2B, HCFC2, IKZF1, ING2, INO80C, KANSL1, KDM6B, MAX, MECP2, MTA2, PADI1, PHC3, PHIP, RAG1, RYBP, RLIM, RPS6KA4, SCML2, SMARCD1, USP7*	*ASH1L, AURKA, BCOR, C11orf30, CHD1, CUL3, CUL4A, EPC2, ERBB4, EZH2, FOXP1, FOXP2, IKZF1, ING2, KAT2B, KAT5, KMT2A, L3MBTL4, MAP3K7, MAX, MBD5, MRGBP, MTA3, NAP1L2, NCOA1, NFYC, NIPBL, PAK2, PCGF5, PHF20, PHF20L1, PHIP, PPP4R2, PRDM2, PRKAA1, PRKAB2, PRKCB, PRMT7, RCC1, SAP18, SATB1, SATB2, SETD6, SIRT6, SMARCA5, SMEK1, SUPT7L, SUZ12, TADA2B, TAF5L, TBL1XR1, TLE4, TSSK6, UBN1, USP36, USP46, ZMYM2, ZMYND11*
Determined by only 1 of 6 databases	*ACTL6A, ANP32B, ANP32E, APOBEC3G, ASF1A, ASXL3, BARD1, BAZ2A, BAZ2B, BRMS1L, BRWD1, CDC6, CDY1, CDK1, CDK7, CECR2, CHD6, CTCFL, CUL3, CUL5, DPPA3, EED, EYA1, EYA4, EID2, ENY2, EP300, ERCC6, EXOSC7, FOXP1, FOXP2, GFI1, HDAC2, HDAC4, HDAC6, HDAC9, HLTF, HMGN4, HMGN5, ING2, JADE1, JAK2, KAT2B, KAT6A, KDM4C, KDM5A, KDM6A, KDM7A, L3MBTL4, MBD2, MBD5, MBIP, MGA, MINA, MORF4L2, NAP1L2, NIPBL, PARP1, PCGF5, PHC3, PHF20L1, PHF8, PIWIL4, PPARGC1A, PPP2CA, PRDM13, PRDM2, PRDM5, PRKAA1, PRKAB2, PRKCB, PRKDC, RAD54B, RBBP7, RLIM, RUVBL1, SAP18, SCML2, SFMBT1, SHPRH, SMARCA1, SMARCAL1, SMYD2, SPOP, SRCAP, SUV39H2, TADA1, TAF12, TAF2, TAF7, TET2, TLE1, TOP2A, TRRAP, UBE2N, UBR2, UCHL5, UHRF2, UIMC1, USP15, VDR, WHSC1, ZNF711*	*ACTL6A, ANKRD32, ANP32E, APBB1, APOBEC3D, ARID1B, ARID2, ARID4B, ATF7IP, BAHD1, BPTF, BRCA2, BRD9, BUB1, CBX5, CDY1, CHD1L, CHD2, CHD4, CIT, CSNK2A1, CSRP2BP, CTCF, DEK, DNMT3B, DTX3L, ELP2, EPC2, EXOSC1, FAM175B, FBRSL1, FOXP1, FOXP2, HDAC9, HDGF, HIF1AN, HIRA, HMG20A, HMGN2, HP1BP3, YY1, ING3, JMJD1C, KAT2B, KDM4C, KDM5A, KDM5B, MASTL, MBD1, MBD6, MGA, MLLT1, MLLT6, MTA3, MTF2, NCOR1, NIPBL, NSD1, OGT, PARP1, PCGF5, PHF14, PHF20L1, PHF8, PPP2CA, PPP4C, PRDM16, PRDM2, PRDM4, PRKAG3, RAD51, RBBP5, RCOR3, SAP18, SATB1, SCML4, SET, SETD7, SETDB1, SF3B3, SFMBT1, SHPRH, SIRT6, SMARCA2, SMARCA4, SMARCD2, SP140, SPEN, SPOP, SUDS3, SUV420H1, TAF5L, TAF9, TFDP1, TLE1, TRIM27, TRIM33, TRRAP, TTK, UCHL5, UHRF2, USP12, USP15, USP3, ZBTB16, ZMYND8, ZNF532, ZNF711, ZZZ3*	*ANKRD32, APOBEC1, APOBEC2, ARID1B, ARID2, ASXL2, ATAD2B, ATF7IP, ATM, ATXN7, BCORL1, BRWD3, CDC6, CDC73, CDK3, CHD6, CREBBP, CSNK2A1, CUL5, DAXX, DNMT3A, DPF3, ELP3, EP300, ERCC6, EXOSC8, EXOSC9, EZH1, FAM175A, FOXP4, GTF3C4, HAT1, HDAC2, HDAC4, HELLS, HLCS, HMGN5, YEATS4, YY1, ING3, YWHAB, JADE1, KDM4C, KDM5D, KDM6A, KDM6B, KDM7A, MBD1, MBD2, MBTD1, MECP2, MYSM1, MTF2, NAT10, PADI4, PARG, PHC3, PRDM13, PRDM16, PRDM6, PRKAG1, PRKAG2, PRMT6, PSIP1, RAD54L2, RCOR3, RNF8, RPS6KA5, SETD1B, SETD3, SETD7, SF3B1, SFMBT2, SFPQ, SMARCA1, SMARCE1, SNAI2, SP1, SP100, SS18L2, STK4, SUDS3, SUV420H1, TAF7, TDRD3, TET2, TET3, TLE1, TLK2, TRIM24, UBE2B, UBE2D1, UBR7, USP15*

Targets selected for gene expression analysis are in bold.
*All 6 databases are listed in Table S17.

Table S19. Multivariate Cox proportional hazard analysis of miRNA host gene methylation and other molecular and/or clinical-pathological variables in Lithuanian and The Cancer Genome Atlas (TCGA) prostate cancer cohorts. Selected combinations with significant model's P-value are shown. For simplicity, only models' P-values are provided for TCGA data.

No.	Covariates	Lithuanian cohort (BCR-free)			TCGA cohort (BCR-free)	TCGA cohort (disease-free)
		HR (95% CI)	Covariate's P-value	Model's P-value	Model's P-value	Model's P-value
Combinations of miRNA host genes						
1	mir-155 (M *vs.* U/ cont.)	1.83 (0.79; 4.23)	0.1581	**0.0265**	0.5181	0.5917
	mir-152 (M *vs.* U/ cont.)	2.17 (0.98; 4.84)	0.0583			
2	mir-155 (M *vs.* U/ cont.)	1.89 (0.82; 4.35)	0.1347	**0.0172**	0.6854	0.0632
	mir-31 (M *vs.* U/ cont.)	3.15 (1.18; 8.42)	**0.0229**			
3	mir-152 (M *vs.* U/ cont.)	2.14 (0.95; 4.83)	0.0691	**0.0130**	0.5551	0.0512
	mir-31 (M *vs.* U/ cont.)	2.95 (1.07; 8.12)	**0.0368**			
4	mir-155 (M *vs.* U/ cont.)	1.67 (0.71; 3.92)	0.2417	**0.0175**	0.7529	0.1123
	mir-152 (M *vs.* U/ cont.)	1.87 (0.81; 4.29)	0.1419			
	mir-31 (M *vs.* U/ cont.)	2.62 (0.95; 7.24)	0.0648			
Combinations with fusion transcript status						
7	*TMPRSS2-ERG* (yes *vs.* no)	0.62 (0.30; 1.29)	0.2062	**0.0333**	n.a.	n.a.
	mir-155 (M *vs.* U/ cont.)	2.85 (1.14; 7.12)	**0.0258**			
8	*TMPRSS2-ERG* (yes *vs.* no)	0.67 (0.32; 1.41)	0.2961	**0.0151**	n.a.	n.a.
	mir-155 (M *vs.* U/ cont.)	2.37 (0.93; 6.06)	0.0732			
	mir-152 (M *vs.* U)	2.31 (1.03; 5.20)	**0.0441**			
Combinations with pathological variables						
9	pT (3 *vs.* 2)	4.14 (1.99; 8.61)	**0.0002**	**<0.0001***	**0.0003**	**0.0008**
	mir-152 (M *vs.* U/ cont.)	2.36 (1.09; 5.11)	**0.0302**			
10	pT (3 *vs.* 2)	4.30 (2.07; 8.93)	**0.0001**	**0.0002**	**0.0004**	**0.0006***
	mir-137 (M *vs.* U/ cont.)	1.43 (0.71; 2.88)	**0.3148**			
11	pT (3 *vs.* 2)	4.33 (2.09; 8.99)	**0.0001**	**<0.0001***	**0.0007**	**0.0007***
	mir-31 (M *vs.* U/ cont.)	3.94 (1.51; 10.30)	**0.0053**			

Table S19. Continued.

No.	Covariates	Lithuanian cohort (BCR-free)			TCGA cohort (BCR-free)	TCGA cohort (disease-free)
		HR (95% CI)	Covariate's P-value	Model's P-value	Model's P-value	Model's P-value
Combinations with pathological variables						
12	pT (3 *vs.* 2)	4.47 (2.13; 9.36)	**0.0001**	**<0.0001**	**0.0013**	**0.0007**
	mir-152 (M *vs.* U/ cont.)	2.17 1.00; 4.71)	0.0513			
	mir-31 (M *vs.* U/ cont.)	3.57 (1.36; 9.41)	**0.0103**			
13	ISUP grade group (5 groups)	3.30 (2.09; 5.21)	**<0.0001**	**<0.0001***	**<0.0001**	**<0.0001***
	mir-31 (M *vs.* U/ cont.)	3.23 (1.24; 8.43)	**0.0172**			
14	PSA (cont.)	1.02 (1.00; 1.04)	**0.0371**	**0.0176***	0.7860	0.1818
	mir-152 (M *vs.* U/ cont.)	2.74 (1.26; 5.96)	**0.0115**			

M/U – methylated/ unmethylated promoter status, *TMPRSS2-ERG* – fusion transcript status, pT – tumor stage, ISUP – International Society of Urological Pathology, PSA – prostate-specific antigen, HR – hazard ratio, CI – confidence intervals. Significant P-values are in bold.
*Forward entering of covariates

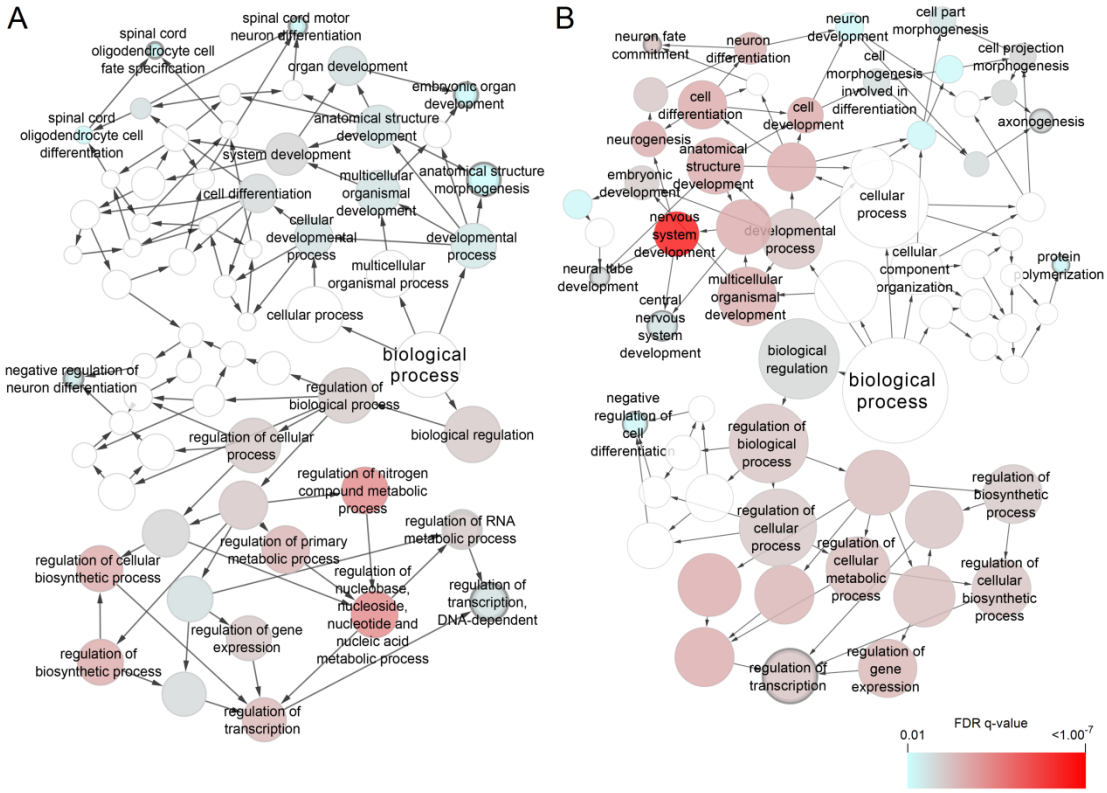


Figure S1. Networks of Gene Ontology (GO) terms for genes differentially methylated comparing prostate tumors and noncancerous tissues. A – for genes with methylation changes in promoter region; B – for genes with methylation changes in intragenic regions. Genes having fold change (FC) value of ≥1.5 were included in the analysis. The area of the node is proportional to the number of genes in the test set annotated to a particular GO term. The color scale is based on false discovery rate (FDR) adjusted P-value (q-value). White nodes are not significantly over-represented, they are included to show other nodes in the context of GO hierarchy. Nodes having only "parent" terms are depicted as bold circles.

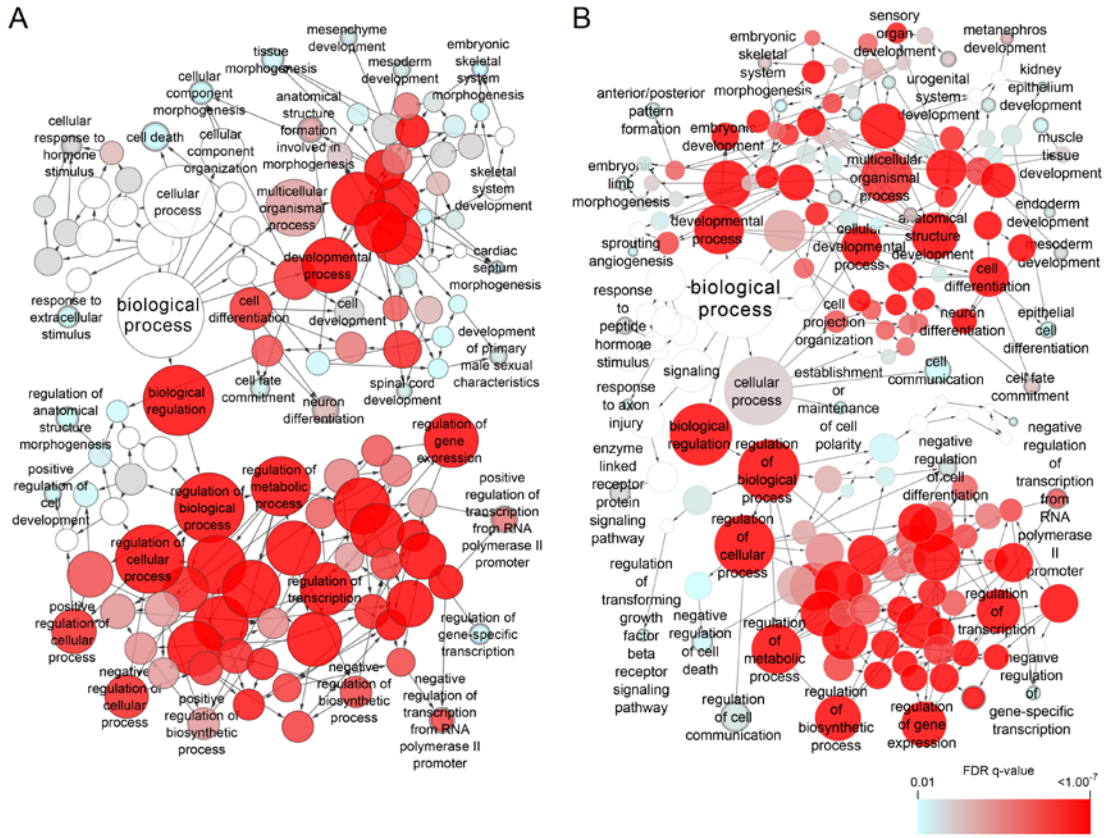


Figure S2. Networks of Gene Ontology (GO) terms for genes differentially methylated comparing biochemical disease recurrence (BCR)-positive and BCR-negative cases. A – for genes with methylation changes in promoter region; B – for genes with methylation changes in intragenic regions. Genes having fold change (FC) value of ≥1.2 were included in the analysis. The area of the node is proportional to the number of genes in the test set annotated to a particular GO term. The color scale is based on false discovery rate (FDR) adjusted P-value (q-value). White nodes are not significantly over-represented, they are included to show other nodes in the context of GO hierarchy. Nodes having only "parent" terms are depicted as bold circles.

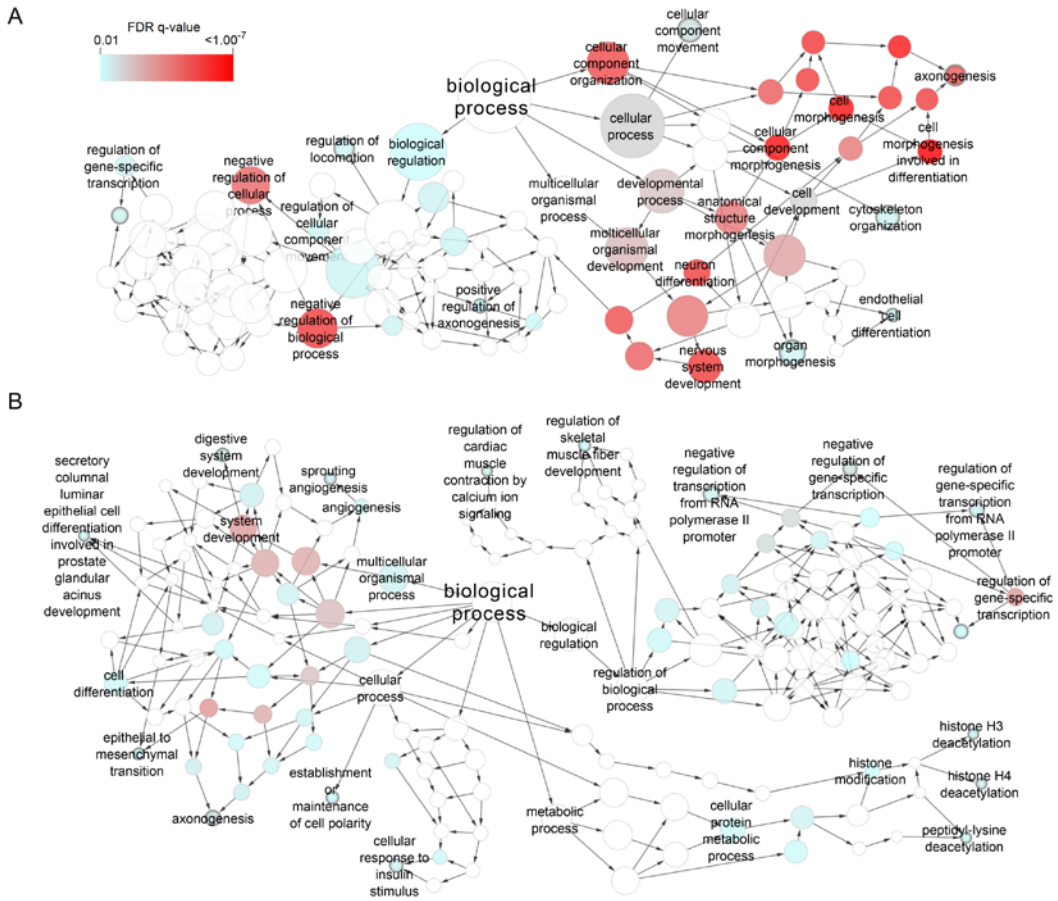


Figure S3. Networks of Gene Ontology (GO) terms for differentially methylated miRNA host genes in prostate cancer (PCa). A – for miRNA host genes with methylation differences between PCa and noncancerous tissues (NPT); B – for miRNA host genes with methylation changes between biochemical disease recurrence (BCR)-positive and BCR-negative cases. Genes having fold change (FC) value of ≥1.2 were included in the analysis. The area of the node is proportional to the number of genes in the test set annotated to a particular GO term. The color scale is based on false discovery rate (FDR) adjusted P-value (q-value). White nodes are not significantly over-represented, they are included to show other nodes in the context of GO hierarchy. Nodes having only "parent" terms are depicted as bold circles.

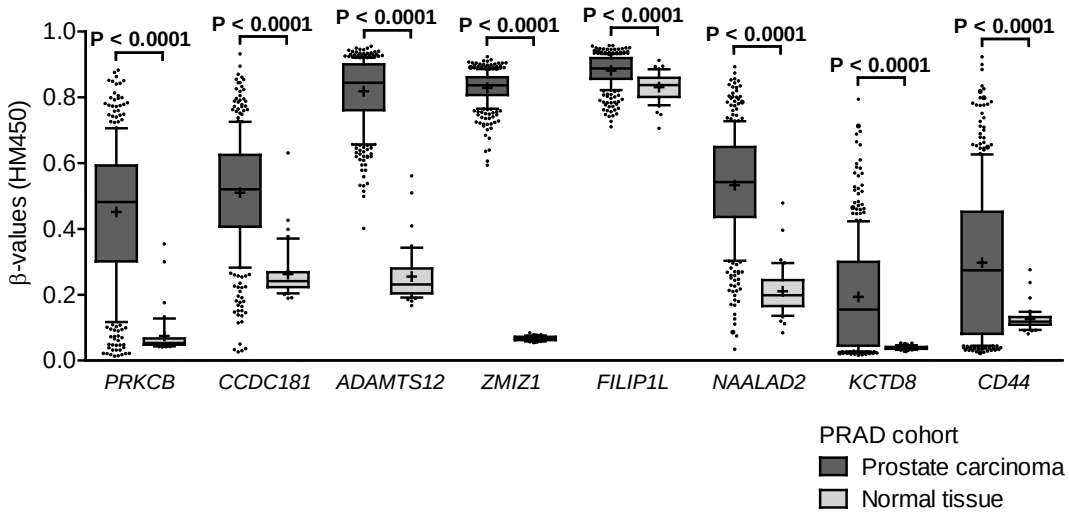


Figure S4. DNA methylation levels of genes *PRKCB*, *CCDC181*, *ADAMTS12*, *ZMIZ1*, *FILIP1L*, *NAALAD2*, *KCTD8*, and *CD44* in the prostate cancer dataset (PRAD) of The Cancer Genome Atlas (TCGA) [4]. Level 3 DNA methylation data, obtained using Illumina HumanMethylation450K (HM450) platform, was used to generate the plots. The box extends from the 25th to 75th percentiles; the line in the middle of the box is plotted at the median; the plus sign depicts the mean; the whiskers represent the 10-90% range; data values outside the range are marked as dots. Significant P-values are in bold.

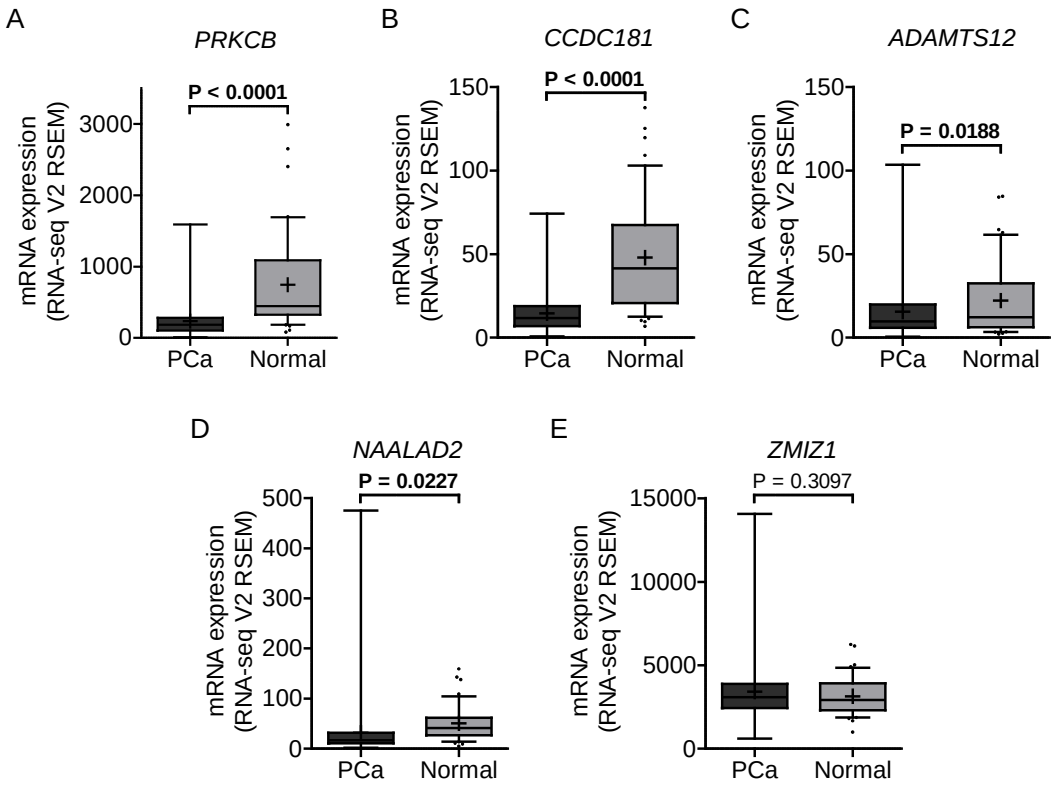


Figure S5. Relative expression levels of genes *PRKCB* (A), *CCDC181* (B), *ADAMTS12* (C), *NAALAD2* (D), and *ZMIZ1* (E) in the prostate cancer (PCa) cohort (PRAD) of The Cancer Genome Atlas (TCGA) [4]. Level 3 PRAD RNA-seq RSEM data were used to generate the plots. The box extends from the 25th to 75th percentiles; the line in the middle of the box is plotted at the median; the plus sign depicts the mean; the whiskers represent the 10-90% range; data values outside the range are marked as dots. Significant P-values are in bold.

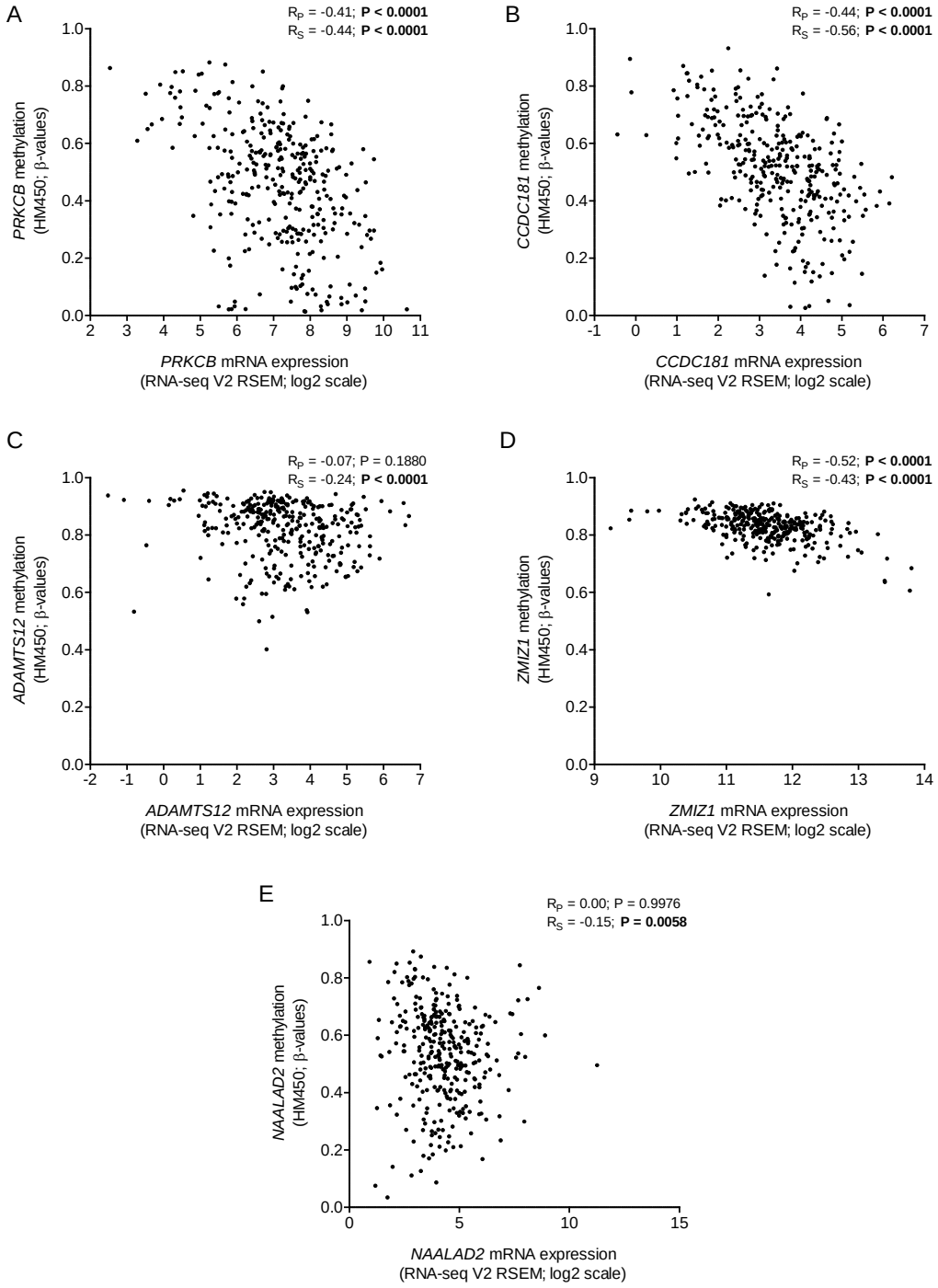


Figure S6. Correlations between promoter methylation and gene expression for *PRKCB* (A), *CCDC181* (B), *ADAMTS12* (C), *ZMIZ1* (D), and *NAALAD2* (E) in the prostate cancer (PRAD) cohort of The Cancer Genome Atlas (TCGA) [4]. Level 3 DNA methylation data, obtained using Illumina HumanMethylation450K (HM450) platform, and level 3 PRAD RNA-seq RSEM data were used to generate scatter plots. RNA-seq data is plotted on log2 scale. Pearson's R (R_P) and Spearman's R (R_S) correlation coefficients are provided with respective P-values. Significant P-values are in bold.

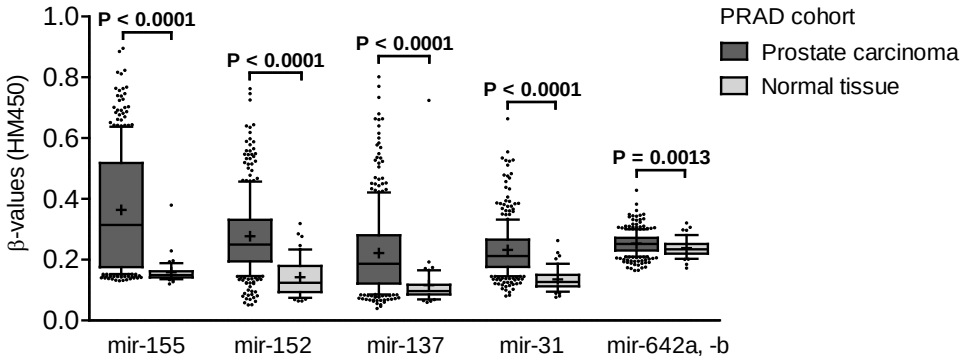


Figure S7. DNA methylation levels of mir-155, mir-152, mir-137, mir-31, and mir-642a, -b host gene promoter-associated loci in the prostate cancer cohort (PRAD) of The Cancer Genome Atlas (TCGA) [4]. Level 3 DNA methylation data, obtained using Illumina Human Methylation450K (HM450) platform, was used to generate the plots. The box extends from the 25th to 75th percentiles; the line in the middle of the box is plotted at the median; the plus sign depicts the mean; the whiskers represent the 10-90% range; data values outside the range are marked as dots. Significant P-values are in bold.

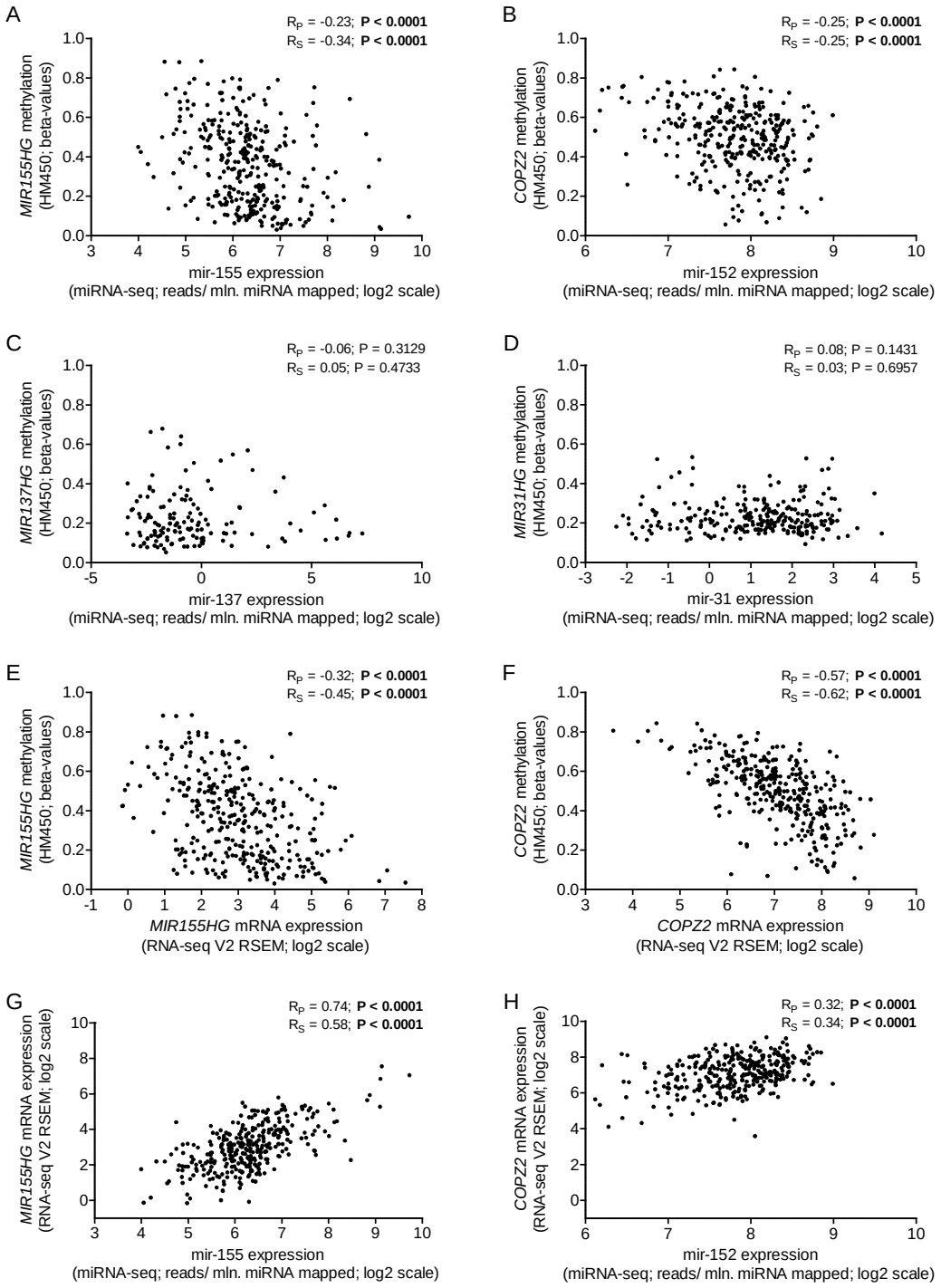


Figure S8. Expression of the selected miRNAs and their host genes in the prostate cancer cohort (PRAD) of The Cancer Genome Atlas (TCGA) [4]. A-D – correlations of miRNA host gene promoter methylation and expression of the respective miRNAs; E and F – correlations of miRNA host gene promoter methylation and expression of mRNAs; G and H – correlations of host gene mRNA expression and respective miRNAs. Level 3 DNA methylation data, obtained using Illumina HumanMethylation450K (HM450) platform, and level 3 PRAD RNA-seq and miRNA-seq RSEM data were used to generate scatter plots. R_P and R_S – Pearson's R and Spearman's R correlation coefficients, respectively. Significant P-values are in bold.

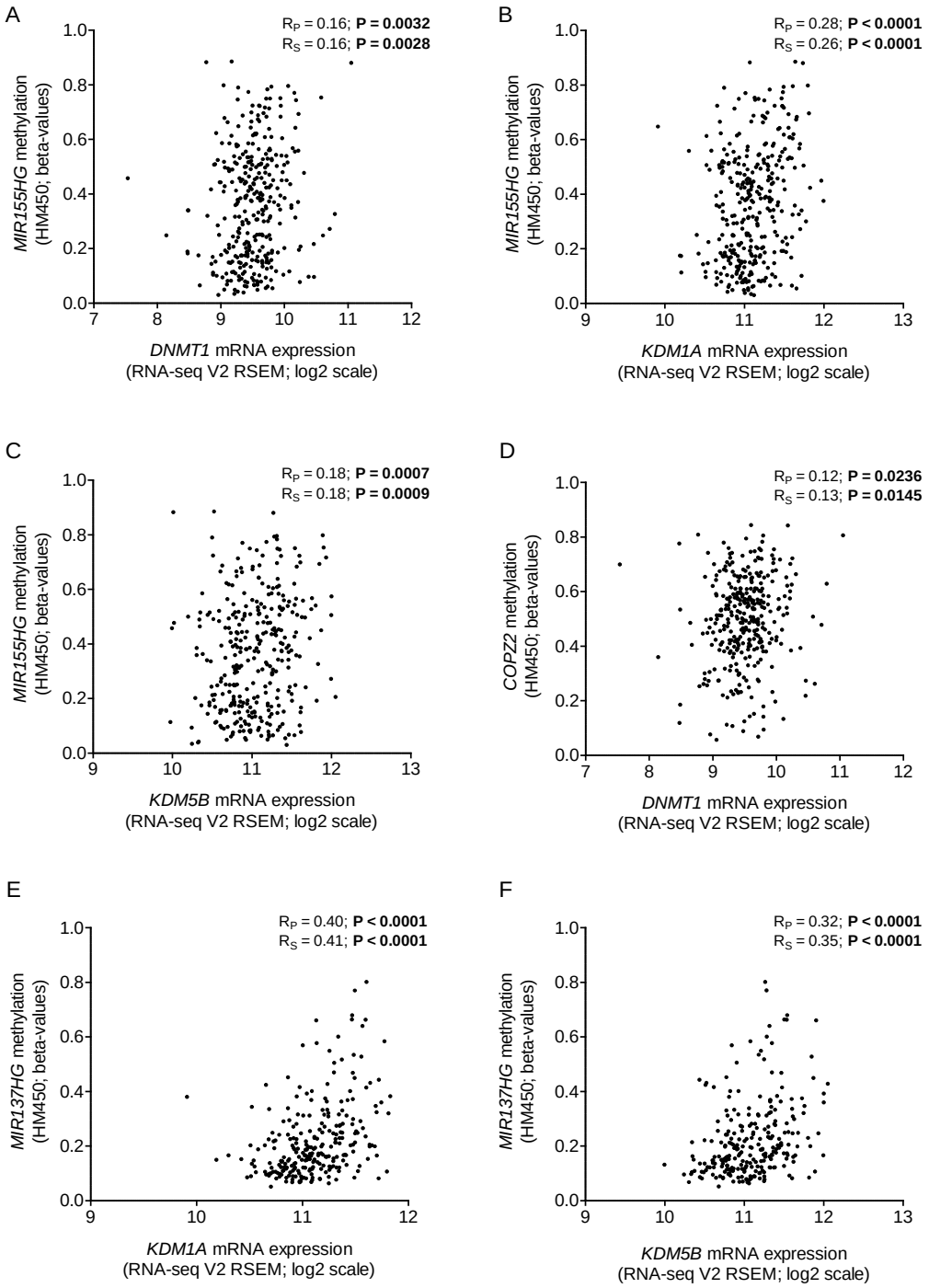


Figure S9. Correlations between miRNA host gene promoter methylation and target gene expression in the prostate cancer cohort (PRAD) of The Cancer Genome Atlas (TCGA) [4]. A-C – correlations between mir-155 host gene (*MIR155HG*) methylation and expression of *DNMT1*, *KDM1A*, and *KDM5B*; D – correlation between mir-152 host gene (*COPZ2*) methylation and expression of *DNMT1*; E and F – correlations between mir-137 host gene (*MIR137HG*) methylation and expression of *KDM1A* and *KDM5B*. Level 3 DNA methylation data, obtained using Illumina HumanMethylation450K (HM450) platform, and level 3 PRAD RNA-seq RSEM data were used to generate scatter plots. RNA-seq data is plotted on log2 scale. Pearson's R (R_P) and Spearman's R (R_S) correlation coefficients are provided with respective P-values. Significant P-values are in bold.