

Absolute quantitation of DNA methylation of 28 candidate genes in prostate cancer using pyrosequencing

Nataša Vasiljević^a, Keqiang Wu^a, Adam R. Brentnall^a, Dae Cheol Kim^a, Mangesh A. Thorat^a, Sakunthala C. Kudahetti^b, Xueying Mao^b, Liyan Xue^b, Yongwei Yu^c, Greg L. Shaw^d, Luis Beltran^d, Yong-Jie Lu^b, Daniel M. Berney^b, Jack Cuzick^a and Attila T. Lorincz^{a,*}

^aCancer Research UK Centre for Epidemiology, Mathematics and Statistics, Wolfson Institute of Preventive Medicine, Barts & the London School of Dentistry and Medicine, Queen Mary University of London, London, UK

^bMolecular Oncology and Imaging, Institute of Cancer, Barts & the London School of Dentistry and Medicine, Queen Mary University of London, London, UK

^cDepartment of Pathology, Changhai Hospital, The Second Military Medical University, Shanghai, China

^dDepartment of Histopathology, Whipps Cross Hospital, London, UK

Abstract. Aberrant DNA methylation plays a pivotal role in carcinogenesis and its mapping is likely to provide biomarkers for improved diagnostic and risk assessment in prostate cancer (PCa). We quantified and compared absolute methylation levels among 28 candidate genes in 48 PCa and 29 benign prostate hyperplasia (BPH) samples using the pyrosequencing (PSQ) method to identify genes with diagnostic and prognostic potential.

RARB, *HIN1*, *BCL2*, *GSTP1*, *CCND2*, *EGFR5*, *APC*, *RASSF1A*, *MDR1*, *NKX2-5*, *CDH13*, *DPYS*, *PTGS2*, *EDNRB*, *MAL*, *PDLIM4*, *HLAa*, *ESR1* and *TIG1* were highly methylated in PCa compared to BPH ($p < 0.001$), while *SERPINB5*, *CDH1*, *TWIST1*, *DAPK1*, *THRB*, *MCAM*, *SLIT2*, *CDKN2a* and *SFN* were not. *RARB* methylation above 21% completely distinguished PCa from BPH. Separation based on methylation level of *SFN*, *SLIT2* and *SERPINB5* distinguished low and high Gleason score cancers, e.g. *SFN* and *SERPINB5* together correctly classified 81% and 77% of high and low Gleason score cancers respectively. Several genes including *CDH1* previously reported as methylation markers in PCa were not confirmed in our study. Increasing age was positively associated with gene methylation ($p < 0.0001$).

Accurate quantitative measurement of gene methylation in PCa appears promising and further validation of genes like *RARB*, *HIN1*, *BCL2*, *APC* and *GSTP1* is warranted for diagnostic potential and *SFN*, *SLIT2* and *SERPINB5* for prognostic potential.

Keywords: Prostate cancer, BPH, DNA Methylation, pyrosequencing, biomarker

1. Introduction

Prostate cancer (PCa) is one of the most common malignancies affecting men and is a major public health problem likely to increase in magnitude with an increasingly aged population. Prostate specific antigen

(PSA) screening for PCa is widely used in the USA, which also has the highest reported incidence worldwide [1]. Despite poor specificity, PSA is valuable for early detection of PCa. The anticipated wider adoption of PSA screening in Europe will lead to increases in reported incidence of PCa, possibly reaching levels seen in the USA. While PSA screening decreases the absolute risk of death from PCa, the lives saved come at a price of invasive examinations and biopsy of healthy men, risk of over-diagnosis and over-treatment as well as increased health care burden [2,3]. It is estimated that to save one life, 1410 men need to be screened

*Corresponding author: Attila T. Lorincz PhD. Cancer Research UK Centre for Epidemiology, Mathematics and Statistics, Wolfson Institute of Preventive Medicine, Barts & the London School of Dentistry and Medicine, Queen Mary University of London, London EC1M 6BQ, UK. E-mail: a.lorincz@qmul.ac.uk.

resulting in 48 additional cases of unnecessarily diagnosed cancer [2]. One of the main reasons for this unsatisfactory situation is the biology of PCa, a large proportion of which is essentially harmless and will not result in significant morbidity or death if left untreated. Unfortunately, current PCa screening methods cannot identify the harmful cancers and thus the desired goal of an ideal screening remains unrealized. Therefore, a need for better screening, diagnosis, prognosis and treatment strategies and overall better biomarkers of prostate disease assessment is clearly evident.

Gene silencing and reactivation, effected in-part by DNA methylation or demethylation of CG dyads in CpG islands, is a key control mechanism in cellular differentiation, development, and disease [4–6]. Aberrant DNA methylation and the associated deregulation of gene expression are believed to play pivotal roles in cancer development and progression [7]. As these abnormalities can be measured, their careful mapping in candidate genes may identify novel biomarkers. One of the earliest [8] and often [9–17] reported abnormalities arising during prostate carcinogenesis is hypermethylation of the *GSTP1* gene detected in > 90% of cancers but not in normal tissues [18]. In contrast to *GSTP1*, the majority of data for other investigated genes come from isolated reports and small studies, e.g. *HIN1* [19] and *BCL2* [17]. Also, there are evident inconsistencies between the reports for genes such as *DAPK1* [12,13], *CDKN2A* [12,14,17], *CDHI* [11, 20] and *EDNRB* [12,20], consequently their methylation status and biomarker potential remains uncertain. A majority of previous studies employed methylation-specific polymerase chain reaction (MSP) [21] or quantitative real-time MSP [22] to assess methylation status and often artificially categorized the data into simplistic hypermethylated or hypomethylated categories. Despite common use, these methods have known disadvantages such as false-positive and false-negative results, as well as use of relative comparisons instead of absolute quantification.

Numerous studies report aberrant promoter and exon/intron methylation in a growing list of genes involved in tumor suppression, hormonal response, signaling, cell cycle control and tumor cell invasion. If carefully characterized in an appropriate set of steps involving credentialing, verification and validation as described by the Early Detection Research Network [23], some of these genes can fulfill the strong need for specialized individual biomarker assays. Our study was designed to select interesting markers for validation in large prospective cohorts with known survival

outcomes such as the Trans Atlantic Prostate Group (TAPG) cohort [24]. We selected candidate genes from the published literature to cover a wide range of cellular pathways since these genes are more likely to pass successfully through the validation pathway. Indeed such an approach has been very productive in the past, resulting in a widely used clinical assay Oncotype Dx for breast cancer. Therefore, we focused on three main groups. The first group consisted of genes that have repeatedly been shown as aberrantly methylated in various prostate malignancies – *GSTP1* [8–17], *RARB* [9, 13,16,17,25], *APC* [10,12,16,17,26], *RASSF1A* [9–12, 16,17,20], *CDHI* [11,12,20], *CDKN2A* [11,12,14,17], *TIG1* [16,17,25], *MDR1* [10,12,17], *EDNRB* [12,20], and *PTGS2* [10,12,15]. The second group consisted of genes that have been occasionally reported as differentially methylated in PCa – *CCND2* [16,26], *ESR1* [9,12], *DAPK1* [12,13], *SFN* [27], *DPYS* [28], *NKX2-5* [9,28], *PDLIM4* [15], *BCL2* [17], *HIN1* [19], *MCAM* [29], *EGFR5* [28], and *CDHI3* [11]. The third group comprised genes that have displayed interesting potential in cancers other than prostate – *MAL* [30], *THRB* [31], *SLIT2* [32], *HLAa* [33], *TWIST1* [34], and *SERPINB5* [35].

To quantify DNA methylation in the most accurate available way, we chose to use PCR followed by pyrosequencing (PSQ). This is a sensitive and highly reproducible method [36] that is uniquely suited for DNA methylation analyses of clinical specimens yielding small amounts of DNA. Unlike most other methods, it provides absolute quantitative information on bases at each interrogated CG site. The assay design readily allows the interrogation of different parts of the gene(s) of interest as well as inclusion of internal controls to address inaccuracies resulting from incomplete bisulfite conversion. In addition, the PSQ method can be used to analyze up to a hundred genes in a large number of specimens using automated workstations and obtain data from limited amounts of DNA which would not be possible with high-throughput microarrays.

We investigated methylation of the 28 genes in a well-annotated clinical specimen set comprising benign prostatic hyperplasia (BPH) and PCa (Table 1). The main aims were to identify genes where deviation in methylation patterns can reliably differentiate between benign and malignant prostate tissues as well as to explore associations between gene methylation and Gleason score as a surrogate indicator of disease outcome. In addition, as there have been only a few published reports using the PSQ method, our gene selection approach also offered an opportunity to test the

Table 1
Clinical and pathological characteristics of the 29 BPH
and 48 PCa studied

		BPH	PCa
Age (years)	Range	60–87	39–78
	Average	76	65*
Gleason score	6		9
	7		23
	8		7
	9–10		9
PSA (ng/ml)	< 4	3	0
	4–9.9	3	12
	10–20	1	10
	> 20	0	13
	Unknown	22	13
Source of tissue	TURP	29	6
	Prostatectomy		37
	Unknown		5

*3 missing values.

reproducibility of the method on the more important of the current candidate genes as well as to compare to the data obtained with methods producing relative measurements.

2. Materials and methods

2.1. Human prostate tissue specimens

The study set included fresh frozen prostate tissue from 77 patients of which 48 were diagnosed with cancer and 29 as BPH. Specimens were collected either after radical prostatectomy, transurethral resection of the prostate (TURP) or TURP in cancer patients (channel TURP) (Table 1). Specimens were collected from three different sites, Changhai Hospital in Shanghai, China, Whipps Cross Hospital, London and St Bartholomew's Hospital, London during the period 1996–2008. All specimens were centrally reviewed to confirm diagnosis by expert genitourinary pathologists (DB, YY). Gleason grading was performed by modern standardized criteria [37].

Informed consent was obtained from all patients. UK national approval was obtained from the Northern Multi-Research Ethics Committee, followed by local ethics committee approval from each of the collaborating hospital trusts. Ethical approval from Changhai Hospital Ethics Committee was obtained for the Chinese specimens.

2.2. DNA extraction and bisulfite conversion

Genomic DNA was extracted from 2–3 10 μ m slices of the fresh frozen material using QIAamp DNA Mini

Kit (Qiagen Inc., Hilden, Germany) and quantified by UV absorption, typically yielding in total > 1 μ g of DNA per specimen. 120–300 ng of DNA was used in the bisulfite conversion reactions where unmethylated cytosines were converted to uracil with the Epi-Tect Bisulfite kit (Qiagen) according to manufacturer's instructions. Briefly, DNA was mixed with water, DNA protect buffer and bisulfite mix and the conversion was run in a thermocycler (Biometra, Goettingen, Germany) at the recommended cycle conditions. Converted DNA was purified on a spin column and eluted twice into a total of 40 μ l Buffer EB.

2.3. PCR and pyrosequencing

Primer sets with one biotin-labelled primer were used to amplify the bisulfite converted DNA. New primers for each of the 28 genes (gene names follow the UCSC gene nomenclature system <http://genome.ucsc.edu/>) were designed using PyroMark Assay Design software version 2.0.1.15 (Qiagen); where possible primers were designed to keep amplicons short with lengths between 90 to 140 base pairs (bp) to facilitate later studies of formalin-fixed paraffin-embedded (FFPE) specimens. The size of the amplicons was restricted to a maximum of 210 bp. All primers were located in promoter or first exon CpG islands identified by Meth-Primer [38], depending on where the design of the assay allowed for optimal primers (Supplementary material Table S1). Due care was taken to avoid any primer overlapping CG dyads to prevent amplification biases. Median size of all amplicons was 104 bp. For genes, previously investigated by other methods, primers were positioned to investigate the same CGs or ones in close vicinity. For some genes e.g. *CDH1*, *GSTP1*, we examined two different sites within the CpG island separated by several hundred base pairs. To provide the internal control for total bisulfite conversion, a non-CG cytosine in the region for pyrosequencing was included where possible.

PCRs were performed using a converted DNA equivalent of 200 cells employing the PyroMark PCR kit (Qiagen). The cell genome-equivalents of DNA calculations assumed 6 pg DNA per diploid cell. Briefly, 12.5 μ l master mix, 2.5 μ l Coral red, 5 pmol of each primer, 7 μ l of water and 2 μ l sample were mixed for each reaction and run at thermal cycling conditions: 95°C for 15 min and then 45 cycles: 30 sec at 94°C; 30 sec at the optimized primer-specific annealing temperature (Supplementary material Table 1); 30 sec at 72°C and a final extension for 10 min at 72°C. The amplified

DNA was confirmed by electrophoresis in a 2% low melting point agarose gel (Sigma-Aldrich, Steinheim, Germany) in TBE buffer or by the QiaExel capillary electrophoresis instrument (Qiagen).

A standard pyrosequencing sample preparation protocol was applied [39]. 3 μ l streptavidin beads (GE Healthcare, Buckinghamshire, UK), 37 μ l PyroMark binding buffer (Qiagen), 20 μ l PCR product and 20 μ l water were mixed and incubated for 10 min on a shaking table at 1300 rpm. Using the Biotage Q96 Vacuum Workstation, amplicons were separated, denatured, washed and added to 45 μ l annealing buffer containing 0.33 μ M of pyrosequencing primer. Primer annealing was performed by incubating the samples at 80°C for 2 min and allowed to cool to room temperature prior to pyrosequencing. PyroGold reagents were used for the pyrosequencing reaction and the signal was analyzed using the PSQ 96MA system (Biotage, Uppsala, Sweden). Target CGs were evaluated by instrument software (PSQ96MA 2.1) which converts the pyrograms to numerical values for peak heights and calculates proportion of methylation at each base as a C/T ratio. All runs contained standard curves, which comprised a range of control methylated DNA (0%, 25%, 50%, 75%, and 100%) to allow standardized direct comparisons between different primer sets. For the standard curves a total of 300 ng of unmethylated (Qiagen) and hypermethylated DNA (Millipore, Billerica, MA, USA) were mixed to obtain the different ratios of DNA methylation and then bisulfite converted as described above.

2.4. Statistical analyses

The main analyses were based on mean values of all CG analyzed. The number of CGs analyzed varied between two to six in each gene as allowed by software-defined parameters. To limit numbers of assays run, and costs, genes that showed no potential in differentiating between BPH and cancer or between low and high Gleason score were investigated in fewer specimens.

Methylation differences between the tissues were examined by Mann-Whitney test. To account for the high number of genes tested on the same data, the Benjamin and Hochberg step-up procedure for controlling false discovery rate (FDR) was applied with FDR of 1% [40].

To explore the relationship between gene methylation and age, methylation was normalized by z-scores, where the raw methylation minus the sample mean was divided by the sample standard deviation. Association between methylation and age was explored by Spear-

man's test, while for methylation versus Gleason score, the Cuzick trend test was used [41]. For cases with a PSA measurement, the association with methylation was using Spearman's rank test. Further, Spearman coefficients, based on rank orderings of raw gene methylation in all cancers, were calculated to explore correlation in methylation between genes. The cut-offs chosen to present true-positive rates (TP, proportion of cancers correctly classified) and false-positive rates (FP, proportion of non-cancers incorrectly flagged) by gene were chosen using the same cost function for all genes – namely, to minimize FP – TP.

To help investigate methylation associated with high Gleason scores, a random forest classification algorithm [42] was applied. Plots were used to inspect the genes identified by the random forest, with classification boundaries added from linear discriminant analyses. Gleason score classification accuracy 95% confidence intervals (CI) were based on a non-parametric bootstrap method with 1,000 resamples. All statistical calculations were conducted using software R version 2.9.2 [43]. Rejection of the null hypothesis was assumed at an $\alpha < 0.05$.

3. Results

3.1. Descriptive statistics of candidate gene methylation

The reproducibility of the PSQ method was investigated at the outset by measuring methylation of *GSTP1* on three separate occasions. The mean methylation difference between highest and lowest reading for the same sample was 7% for BPH cases, and 13% for cancers; the Pearson correlation for runs 1 vs. 2 was 0.90 whereas between runs 1 and 3 was 0.97. This concordance was regarded as acceptable and all subsequent data were based on single measurements.

Methylation data were adjusted for primer bias through re-scaling each gene's methylation measurements by the median standard curve obtained for each primer set. The impact of applying these corrections to the genes had small effects on median methylation differences (Supplementary material Fig. S1) but allowed comparison across different genes. Of the 28 genes studied, methylation of 20 genes: *RARB*, *HIN1*, *BCL2*, *GSTP1*, *CCND2*, *EGFR5*, *APC*, *RASSF1A*, *MDR1*, *NKX2-5*, *CDH13*, *DPYS*, *PTGS2*, *EDNRB*, *MAL*, *PDLIM4*, *SERPINB5*, *HLAa*, *ESR1* and *TIG1* could distinguish prostate cancer from BPH tissue at

Table 2

The proportion of cancers correctly classified (true positive, TP) and the proportion of BPH incorrectly classified (false positive, FP) when the best cut off from the primer adjusted data was used. Genes with FP above 50% are excluded

Gene	Cut off (% methylation)	TP (%)	FP (%)
RARB	20.8	100	0
GSTP1	7.3	98	0
HIN1	26.5	98	0
APC	26.5	96	0
BCL2	10.4	96	3
CCND2	7.5	92	0
CHD13	27.5	88	0
EGFR5	35.5	96	14
NKX2-5	32.5	88	0
RASSF1A	14.8	92	7
DPYS	37.9	85	0
MDR1	26.8	85	3
PTGS2	13.8	79	0
EDNRB	33.5	77	0
MAL	3.8	85	14
PDLIM4	8	75	3
HLAa	6.11	74	14
TIG1	4.8	65	10
ESR1	38.5	62	7
SLIT2	33.5	56	14
CDKN2A	35.5	40	0
MCAM	1.4	45	18
SFN	96.5	15	3
THRB	27.5	26	21
CDH1	13.2	33	14
TWIST1	3.8	56	38

FDR of 1% (Fig. 1, Supplementary material Table S2). Cut-off levels were calculated to evaluate the diagnostic potential of methylation differences (Table 2). This allowed dichotomization of the data, where a cut-off of 21% methylation of *RARB* separated all cancers from BPH with 100% accuracy, i.e. TP = 100%, FP = 0%.

In BPH specimens, *EGFR5*, *DPYS*, *ESR1*, *MDR1*, *SERPINB5* and *SFN* displayed median methylation above 10% whereas most other genes were unmethylated (median methylation ~ 2%) (Fig. 1). In particular, *SERPINB5* and *SFN* were methylated to approximately 50% in BPH. Furthermore, *SERPINB5* was the only gene with significantly higher methylation in BPH than cancers ($p < 0.001$) (Fig. 1). *MCAM*, *CDKN2A*, *THRB*, *TWIST1*, *CDH1* and *DAPK1* were methylated below 10% in both BPH and PCa (Fig. 1, Supplementary material Table S2).

3.2. Association between clinical covariates and gene methylation

The relationship among gene methylation levels, age, Gleason score and PSA levels were explored in the PCa. There was a positive association of Gleason score

with age ($p < 0.001$) and PSA ($p = 0.0013$), though no association between PSA and age ($p = 0.22$) (data not shown).

There was a positive trend between age and standardized mean methylation values across all genes, akin to global methylation status, for each case (Pearson correlation 0.52, $p < 0.0001$; Fig. 2a). Furthermore, inspection of the distribution of p-values suggested a moderate effect of age common to the methylation of all genes, while Gleason score appeared to affect only subsets of genes (Fig. 2b). The methylation levels of *NKX2-5* and *APC* ($p = 0.009$), *TIG1*, *ESR1*, *GSTP1* ($p = 0.01$), *CDH13*, *EGFR5* ($p = 0.02$), *MCAM* ($p = 0.03$) and *SLIT2* ($p = 0.04$) showed a positive association with age (Supplementary material Table S3). The Cuzick trend test showed that the methylation of *SFN* ($p = 0.01$), *TIG1* ($p = 0.02$), *PDLIM4*, *APC* and *SERPINB5* ($p = 0.04$) were associated with Gleason score (Supplementary material Table S3). Moreover, according to random forest classification, high methylation of *SFN*, *SLIT2* and *SERPINB5* separated low from high Gleason score cancers. The linear discrimination boundaries described the structure found by the random forest classification (Fig. 3). Methylation level composite measure of *SFN* and *SERPINB5* correctly classified 81% (95% CI 56–91) of high Gleason scores while 23% (9–47) of low Gleason scores were misclassified. Similarly, methylation of *SFN* and *SLIT2*, detected 62% (47–81) of high and misclassified 12% (9–47) low Gleason scores while methylation of *SERPINB5* and *SLIT2* detected 62% (47–81) of high and misclassified 13% (3–31) of low Gleason scores (Fig. 3).

Methylation levels of 17 genes: *HIN1*, *TWIST1*, *GSTP1*, *RARB* ($p < 0.001$), *HLAa*, *BCL2*, *APC*, *PDLIM4*, *PTGS2*, *DPYS*, *CDH13* ($p < 0.01$) and *RASSF1A*, *MDR1*, *EGFR5*, *EDNRB*, *TIG1*, *CCND2* ($p < 0.05$) were positively associated with PSA (Supplementary material Table S3). Furthermore, methylation levels of most genes that could distinguish between BPH and cancer e.g. *RARB*, *APC*, *EGFR5*, *HIN1*, *RASSF1A*, *PTGS2* and *CDH13* were moderately correlated (Fig. 4).

4. Discussion

Other than PSA, with all its limitations, no generally accepted validated biomarkers are currently available for prognosis or therapeutic prediction in prostate cancer. Although several new markers such as PCA3 [44],

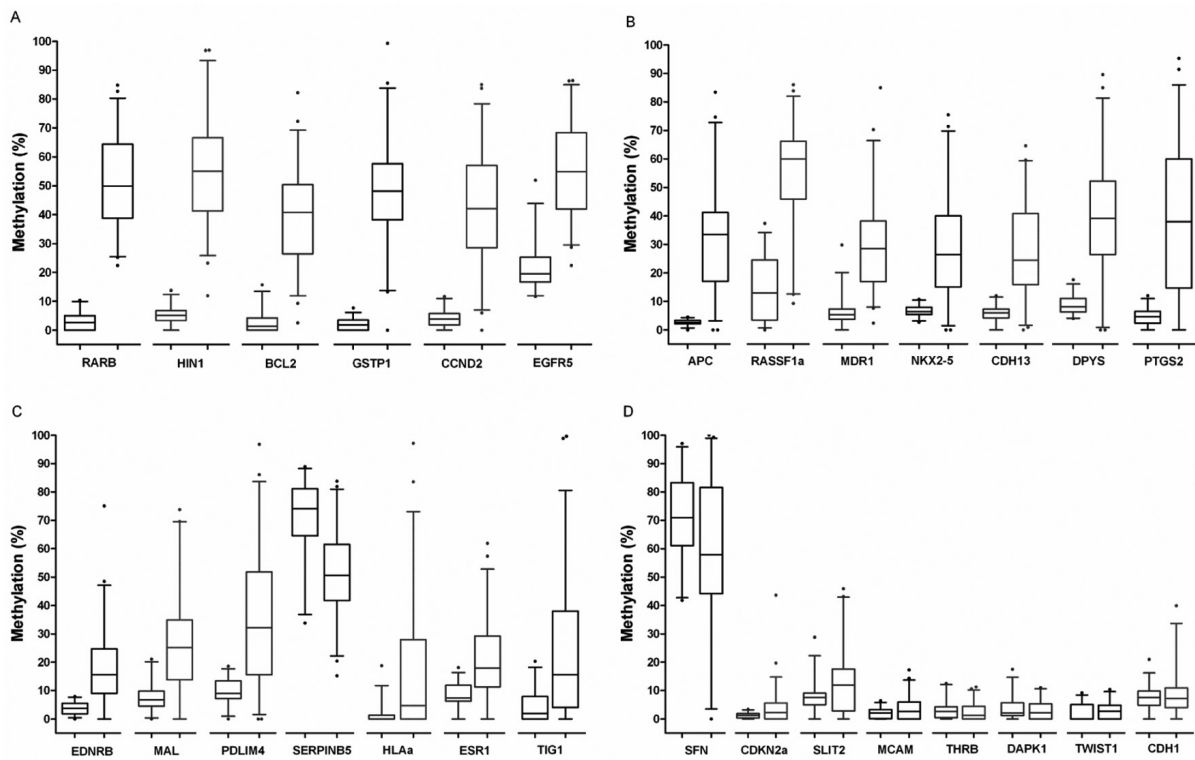


Fig. 1. The distribution of methylation measurements for a set of candidate genes in prostate cancer (PCa) and benign prostatic hyperplasia (BPH). The genes were rank ordered by Mann-Whitney statistics (Supplementary material, Table S2) throughout a-d. The primer unadjusted level of methylation of each gene is demonstrated as pairs of results with BPH corresponding to the boxplot on the left and PCa as the boxplot on the right. The differences in methylation of the tissues divided the genes in a, b and c) into those most likely to distinguish and d) not likely to distinguish BPH from PCa at FDR below 1%. Whiskers of the boxplot mark the 5th and 95th percentiles, the box 25th percentile, median and 75 percentile, while extreme values are shown by (●).

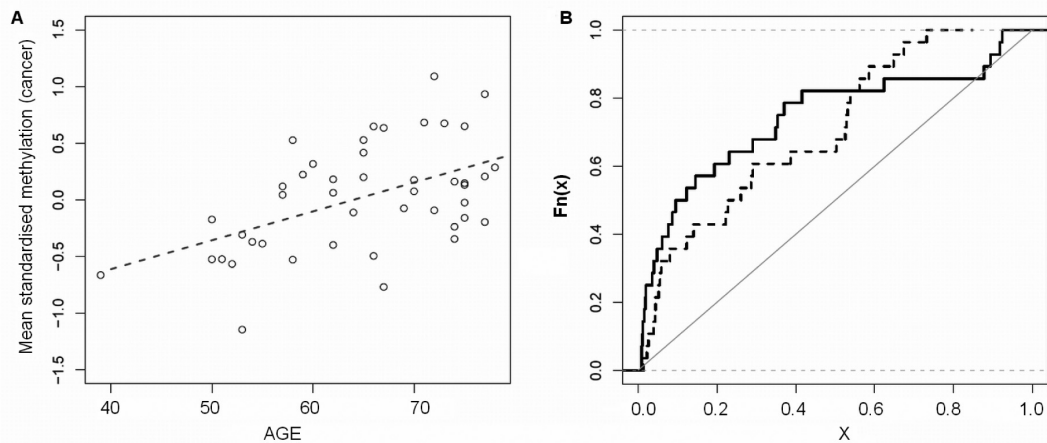


Fig. 2. a) Average methylation by age. Each point in the plot represents a case and corresponds to the average methylation across all the genes tested for that case. The level of methylation is on the y-axis and age on the x-axis. There appears to be a trend of increasing methylation with age as shown by the fitted line ($y = -1.6832 + 0.02631x$, where y and x refer to values of methylation level and age on respective axes). b) The empirical cumulative distribution of p values from tests of an association between methylation and age (solid line), and methylation and Gleason score (dashed line) Supplementary material, Table S3. The x-axis shows the p value and the y axis gives the empirical cumulative distribution $[Fn(x)]$. Under the null hypothesis of no association, a diagonal straight line is to be expected.

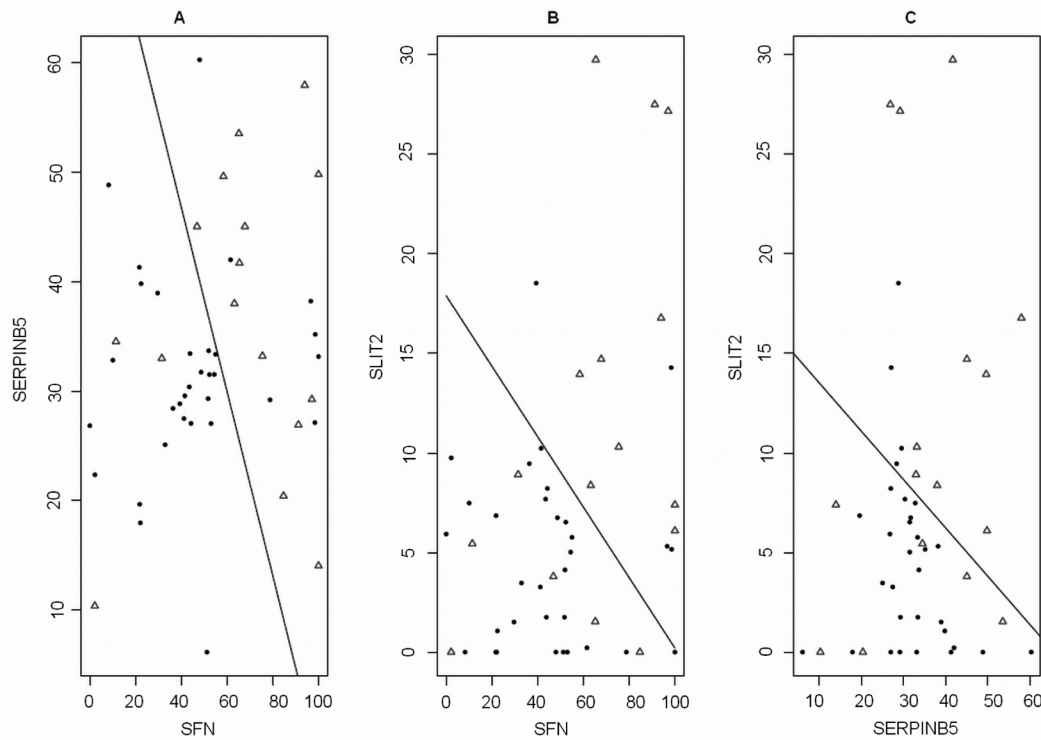


Fig. 3. Methylation values of high Gleason (> 7, triangle) and low Gleason score (circle) cases for three pairs of genes. The separating lines show linear discrimination boundaries that were fitted to classify high from low Gleason score. The aim was to separate these two categories as much as possible. The genes plotted are A) *SFN* and *SERPINB5*, B) *SFN* and *SLIT2*, C) *SERPINB5* and *SLIT2*.

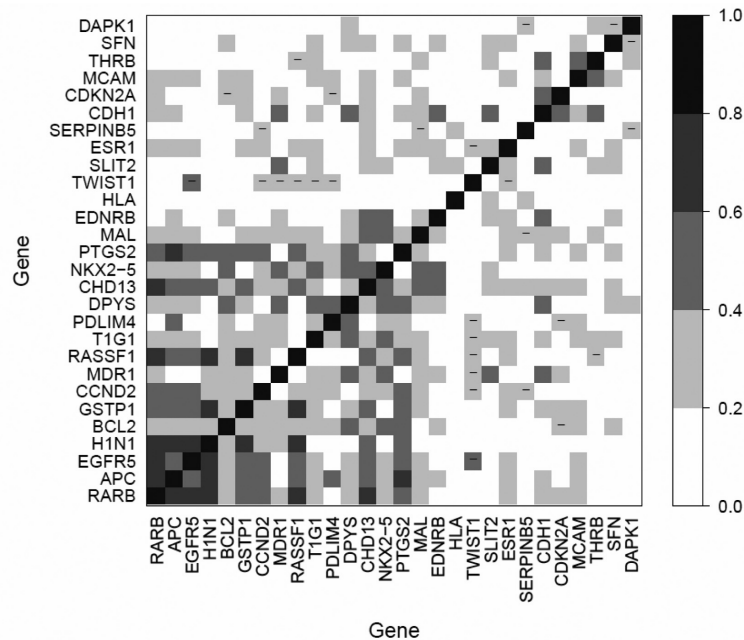


Fig. 4. Heatmap of Spearman correlation of methylation between each pair of genes. The shade depends on the absolute correlation; negative correlations are marked with (-) in the cell. Higher correlation indicates a more similar ranking of all cases by methylation than lower correlation.

TMPRSS-ERG [45], Ki-67 [46], HSP27 [47] and others are under consideration, they are not validated for widespread use and thus Gleason score and PSA in the context of other clinical information remain the mainstay of decision making in PCa [24].

Alterations in DNA methylation contribute to development and progression of PCa [7], and studying the methylation of genes involved in carcinogenesis is likely to reveal useful novel biomarkers that could eventually be validated for PCa. This is the first study to quantify methylation of a large number of candidate genes in a common set of prostate clinical specimens by means of a single highly accurate and validated DNA methylation assay. It demonstrates the methylation patterns of 28 genes with a clear aim to identify candidate methylation biomarkers with diagnostic and/or prognostic potential to be validated in large cohort studies. Such validation in the TAPG series comprising over 1000 FFPE PCa tissues is underway. Our robust PSQ approach is particularly suited to situations where the available specimen set for molecular studies is highly limited in either quality or amount. Assays were specifically designed for use in FFPE material limiting the amplicon size but covering a reasonable number of CG. In addition, our preliminary data show that methylation levels in frozen and matched FFPE specimens are equivalent (Supplementary material Table S4). To fine-tune the PSQ method, predefined quantitative controls were included within each experiment to allow the adjustment of all data and thus improve the accuracy of comparisons in levels of methylation between the investigated genes. Internal controls to check completeness of bisulfite conversion were also included in the assays. This is an important advantage of the PSQ method increasing the measurement accuracy compared to other commonly used methods. The absolute measurements obtained by PSQ were in good agreement to previous reports using relative measurements [48].

Twenty of the investigated genes, namely *RARB*, *HIN1*, *BCL2*, *GSTP1*, *CCND2*, *EGFR5*, *APC*, *RASSF1A*, *MDR1*, *NKX2-5*, *CDH13*, *DPYS*, *PTGS2*, *EDNRB*, *MAL*, *PDLIM4*, *SERPINB5*, *HLAa*, *ESR1* and *TIG1* were more highly methylated in cancers than BPH tissue (Fig. 1) while the risk of false discovery (FDR) was less than 1%. To the best of our knowledge, the methylation of *MAL*, *HLAa*, *SERPINB5*, *THRB*, *TWIST1* and *SLIT2* was demonstrated here for the first time in prostate tissue. While *THRB* and *TWIST1* were overall unmethylated, *HLAa* and *MAL* displayed low methylation with fair ability to discriminate between the tissues with median difference 15% and 7% respec-

tively (Fig. 1a). Methylation status of *HLAa* was associated to level of PSA (Supplemental material Table S3) but not age, Gleason score or methylation of other genes (Fig. 4). This merits further study as a diagnostic biomarker. The methylation of *SLIT2* was low, however, median methylation was elevated in cancers ~6% vs. ~2% in BPH and moreover methylation level of *SLIT2* could separate high and low Gleason score cancers indicating possible prognostic characteristics (Fig. 3).

Average methylation of *SERPINB5* and *SFN* was lower in cancers than in BPH, differences of 15% ($p < 0.001$) and 12% ($p = 0.05$) respectively (Supplementary material Table S2). Interestingly, for both genes in PCa the quantitatively higher levels of methylation were associated with high Gleason score (Fig. 3). *SFN* is regarded as a tumor suppressor gene and the encoded 14-3-3 σ protein interacts with cyclins to arrest the cell cycle when induced by the p53 pathway in response to DNA damage [49]. *SERPINB5* is another putative tumor suppressor gene, expressing the maspin protein which is involved in signal transduction and response to cellular stress [50]. Permanent silencing of these genes may render the cell defenseless towards accumulation of additional DNA damage during cellular stress and therefore play a role in progression of the malignancy to high nuclear grade. However, our data are not consistent with a simplistic tumor suppressor gene silencing mechanism as the basis for cancer progression. On the contrary the higher level of methylation in normal tissue and variable changes in level depending on Gleason score suggest complex regulation of methylation of these genes during progression. Higher methylation level of *SFN* in cancers than high-grade prostatic intraepithelial neoplasia (HGPIN) [27] as well as up-regulation of maspin expression in HGPIN and loss of its expression with progression has been previously observed [51] and underscores the prognostic potential of these genes. In addition, the methylation of *SFN* and *SERPINB5* were approximately ~50% in BPH, perhaps suggesting methylation of only one allele. Since the relatively small changes in DNA methylation levels of *SFN* and *SERPINB5* seemingly have significant effects on disease risk assessment, measurement by a very accurate method such as PSQ may be a necessity.

MCAM was unmethylated in the investigated specimens (Fig. 1d) despite that the aberrant methylation of the *MCAM* promoter in PCa was previously reported [29]. Investigation of different CG sites within the promoter by different methods could be the reason for this discrepancy.

Reports of methylation of *DAPK1*, *CDH1* and *CDKN2A* have been inconsistent. We observed equally low (median < 10%) methylation of these genes in BPH and cancer tissues, although non-significant differences of small magnitude were observed (Supplementary material Table S2, Fig. 1d). For *CDH1*, this observation was true for both promoter regions previously reported to show differences in methylation [52]. These inconsistencies may be due to amplification of non-significant differences by a semi-quantitative method resulting in skewed proportions of methylation. In addition, these genes appeared to have little association with Gleason score or PSA levels (Supplementary material Table S3) and since prognostic value has so far not been evident, we believe that these genes have poor value as biomarkers in PCa.

Baseline PSA data were available for only 35 cancers, nonetheless methylation of *HIN1* was positively associated with PSA ($p < 0.001$) with no evidence of a concurrent association with age or Gleason score (Supplementary data Table S3). Although, *HIN1* methylation has not yet been linked to disease prognosis, the protein's suggested role in preventing invasion [53] strengthens its possible role in cancer progression. Except for *TWIST1*, 16 of the genes showing difference between cancer and BPH were also positively associated with PSA ($p < 0.05$) (Supplementary data Table S3) further supporting their diagnostic potential.

Similar to numerous previous reports, we observed that gene methylation increased with age which was identified as the strongest covariate (Fig. 2). Therefore the need for appropriate adjustment for age or inclusion of age in statistical models in future DNA methylation studies cannot be overemphasized. A recent study using the PSQ method showed the correlation of age and methylation of *RARB*, *RASSF1a* and *GSTP1* in normal tissue [9].

This study should be regarded as a preliminary exploration and has a number of limitations including that the specimen set was not blinded and there were no comparisons to other well respected DNA methylation methods such as bisulfite sequencing. However, inclusion of quantitative controls in every experimental run and inclusion of non-CpG C controls wherever possible as well as the strong agreement in methylation levels of *GSTP1* detected in our study and by other investigators, using a variety of different methods, ensured reliability of our assays. Additional validation of our findings is important in order to establish the clinical utility of these biomarker sets.

In summary, PSQ is a useful and accurate method for studying methylation of genes. A great number

of genes are methylated in PCa and the methylation of many genes is correlated with each other (Fig. 4), in particular *RARB*, *APC*, *EGFR5*, *HIN1*, *RASSF1A*, *PTGS2* and *CDH13*, possibly indicating a concerted mechanism. While age displays a common effect on gene methylation, methylation of *SFN* and *SERPINB5* may be specific to progression of PCa and deserves further study. Overall, 20 of the 28 investigated genes were able to distinguish between BPH and cancer tissues, some with good and others, e.g. *RARB*, with perfect discrimination abilities. Our results further suggest that while methylation status of some genes demonstrates most promising characteristics for diagnosis, such as *RARB*, *HIN1*, *HLAa* and *GSTP1* others such as *SFN*, *SLIT2* and *SERPINB5* deserve further validation as prognostic markers.

Acknowledgments

This study was supported by grants from Cancer Research UK (Grant number C569/A10404), The Orchid Appeal Yearly (Grant number ONAG116R and ON-AG117R) and Qiagen Inc. ATL is an advisor to Qiagen Inc. and owns Qiagen stock. JC has given lectures sponsored by Qiagen on an occasional basis. No other conflicts were declared. Special thanks to Joachim Schorr and Dirk Loeffert for strong support in the early stages of the project. Sponsors had no role in study design, execution, data analyses or the content of the paper. Isabelle Bisson assisted in the collection of data and specimens, Dorota Scibior-Bentkowska assisted the completion of data acquisition.

Supplemental material

Supplementary data can be found at: http://www.wolfson.qmul.ac.uk/ccp/research/MolecularEpidemiology_Laboratory.html.

References

- [1] D.M. Parkin, S.L. Whelan, J. Ferlay et al., Cancer Incidence in Five Continents, *IARC Scientific Publications* **143** (1997).
- [2] F.H. Schroder, J. Hugosson, M.J. Roobol et al., Screening and prostate-cancer mortality in a randomized European study, *N Engl J Med* **360** (2009), 1320–1328.
- [3] A.L. Moore, P. Dimitropoulou, A. Lane et al., Population-based prostate-specific antigen testing in the UK leads to a stage migration of prostate cancer, *BJU Int* **104** (2009), 1592–1598.

- [4] E. Li, C. Beard, A.C. Forster et al., DNA methylation, genomic imprinting, and mammalian development, *Cold Spring Harb Symp Quant Biol* **58** (1993), 297–305.
- [5] T. Haaf, Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development, *Curr Top Microbiol Immunol* **310** (2006), 13–22.
- [6] C.P. Walsh, J.R. Chaillet and T.H. Bestor, Transcription of IAP endogenous retroviruses is constrained by cytosine methylation, *Nat Genet* **20** (1998), 116–117.
- [7] M. Esteller, P.G. Corn, S.B. Baylin et al., A gene hypermethylation profile of human cancer, *Cancer Res* **61** (2001), 3225–3229.
- [8] W.H. Lee, R.A. Morton, J.I. Epstein et al., Cytidine methylation of regulatory sequences near the pi-class glutathione S-transferase gene accompanies human prostatic carcinogenesis, *Proc Natl Acad Sci U S A* **91** (1994), 11733–11737.
- [9] B. Kwabi-Addo, W. Chung, L. Shen et al., Age-related DNA methylation changes in normal human prostate tissues, *Clin Cancer Res* **13** (2007), 3796–3802.
- [10] P.J. Bastian, J. Ellinger, A. Wellmann et al., Diagnostic and prognostic information in prostate cancer with the help of a small set of hypermethylated gene loci, *Clin Cancer Res* **11** (2005), 4097–4106.
- [11] R. Maruyama, S. Toyooka, K.O. Toyooka et al., Aberrant promoter methylation profile of prostate cancers and its relationship to clinicopathological features, *Clin Cancer Res* **8** (2002), 514–519.
- [12] S. Yegnasubramanian, J. Kowalski, M.L. Gonzalgo et al., Hypermethylation of CpG islands in primary and metastatic human prostate cancer, *Cancer Res* **64** (2004), 1975–1986.
- [13] M. Yamanaka, M. Watanabe, Y. Yamada et al., Altered methylation of multiple genes in carcinogenesis of the prostate, *Int J Cancer* **106** (2003), 382–387.
- [14] C. Jeronimo, R. Henrique, M.O. Hoque et al., A quantitative promoter methylation profile of prostate cancer, *Clin Cancer Res* **10** (2004), 8472–8478.
- [15] D.K. Vanaja, K.V. Ballman, B.W. Morlan et al., PDLIM4 repression by hypermethylation as a potential biomarker for prostate cancer, *Clin Cancer Res* **12** (2006), 1128–1136.
- [16] E. Rosenbaum, M.O. Hoque, Y. Cohen et al., Promoter hypermethylation as an independent prognostic factor for relapse in patients with prostate cancer following radical prostatectomy, *Clin Cancer Res* **11** (2005), 8321–8325.
- [17] N. Y. Cho, B. H. Kim, M. Choi 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* **211** (2007), 269–277.
- [18] W. G. Nelson, A. M. De Marzo and S. Yegnasubramanian, Epigenetic alterations in human prostate cancers, *Endocrinology* **150** (2009), 3991–4002.
- [19] I. Krop, A. Player, A. Tablante et al., Frequent HIN-1 promoter methylation and lack of expression in multiple human tumor types, *Mol Cancer Res* **2** (2004), 489–494.
- [20] R. Singal, L. Ferdinand, I.M. Reis et al., Methylation of multiple genes in prostate cancer and the relationship with clinicopathological features of disease, *Oncol Rep* **12** (2004), 631–637.
- [21] J. G. Herman, J. R. Graff, S. Myohanen et al., Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands, *Proc Natl Acad Sci U S A* **93** (1996), 9821–9826.
- [22] C.A. Eads, K.D. Danenberg, K. Kawakami et al., MethyLight: a high-throughput assay to measure DNA methylation, *Nucleic Acids Res* **28** (2000), E32.
- [23] D.F. Ransohoff, Rules of evidence for cancer molecular-marker discovery and validation, *Nat Rev Cancer* **4** (2004), 309–314.
- [24] J. Cuzick, G. Fisher, M.W. Kattan et al., Long-term outcome among men with conservatively treated localised prostate cancer, *Br J Cancer* **95** (2006), 1186–1194.
- [25] J. Zhang, L. Liu, G.P. Pfeifer, Methylation of the retinoid response gene TIG1 in prostate cancer correlates with methylation of the retinoic acid receptor beta gene, *Oncogene* **23** (2004), 2241–2249.
- [26] R. Henrique, F. R. Ribeiro, D. Fonseca et al., High promoter methylation levels of APC predict poor prognosis in sextant biopsies from prostate cancer patients, *Clin Cancer Res* **13** (2007), 6122–6129.
- [27] R. Henrique, C. Jeronimo, M.O. Hoque et al., Frequent 14-3-3 sigma promoter methylation in benign and malignant prostate lesions, *DNA Cell Biol* **24** (2005), 264–269.
- [28] W. Chung, B. Kwabi-Addo, M. Ittmann et al., Identification of novel tumor markers in prostate, colon and breast cancer by unbiased methylation profiling, *PLoS One* **3** (2008), e2079.
- [29] J.W. Liu, J.K. Nagpal, C. Jeronimo et al., Hypermethylation of MCAM gene is associated with advanced tumor stage in prostate cancer, *Prostate* **68** (2008), 418–426.
- [30] K.L. Ostrow, H.L. Park, M.O. Hoque et al., Pharmacologic unmasking of epigenetically silenced genes in breast cancer, *Clin Cancer Res* **15** (2009), 1184–1191.
- [31] Y. Ling, X. Xu, J. Hao et al., Aberrant methylation of the THRB gene in tissue and plasma of breast cancer patients, *Cancer Genet Cytogenet* **196**, 140–145.
- [32] A. Dallol, N.F. Da Silva, P. Viacava et al., SLIT2, a human homologue of the Drosophila Slit2 gene, has tumor suppressor activity and is frequently inactivated in lung and breast cancers, *Cancer Res* **62** (2002), 5874–5880.
- [33] L. Mora-Garcia Mde, A. Duenas-Gonzalez, J. Hernandez-Montes et al., Up-regulation of HLA class-I antigen expression and antigen-specific CTL response in cervical cancer cells by the demethylating agent hydralazine and the histone deacetylase inhibitor valproic acid, *J Transl Med* **4** (2006), 55.
- [34] Q. Feng, S. E. Hawes, J. E. Stern et al., Promoter hypermethylation of tumor suppressor genes in urine from patients with cervical neoplasia, *Cancer Epidemiol Biomarkers Prev* **16** (2007), 1178–1184.
- [35] B. W. Futscher, M. M. O'Meara, C. J. Kim et al., Aberrant methylation of the maspin promoter is an early event in human breast cancer, *Neoplasia* **6** (2004), 380–389.
- [36] S. Marsh, Pyrosequencing applications, *Methods Mol Biol* **373** (2007), 15–24.
- [37] J.I. Epstein, W.C. Allsbrook, Jr., M.B. Amin et al., The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma, *Am J Surg Pathol* **29** (2005), 1228–1242.
- [38] L.C. Li, R. Dahiya, MethPrimer: designing primers for methylation PCRs, *Bioinformatics* **18** (2002), 1427–1431.
- [39] J. Tost and I.G. Gut, DNA methylation analysis by pyrosequencing, *Nat Protoc* **2** (2007), 2265–2275.
- [40] Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society* **57** (1995), 289–300.
- [41] J. Cuzick, A Wilcoxon-type test for trend, *Stat Med* **4** (1985), 87–90.
- [42] L. Breiman, Random Forests, *Machine Learning* **45** (2001), 5–32.
- [43] R.D.C. Team. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing Vienna, Austria, 2009.

- [44] D. Hessels, G. W. Verhaegh, J. A. Schalken, et al., Applicability of biomarkers in the early diagnosis of prostate cancer, *Expert Rev Mol Diagn* **4** (2004), 513–526.
- [45] R. Mehra, S.A. Tomlins, J. Yu et al., Characterization of TMPRSS2-ETS gene aberrations in androgen-independent metastatic prostate cancer, *Cancer Res* **68** (2008), 3584–3590.
- [46] D.M. Berney, A. Gopalan, S. Kudahetti et al., Ki-67 and outcome in clinically localised prostate cancer: analysis of conservatively treated prostate cancer patients from the Trans-Atlantic Prostate Group study, *Br J Cancer* **100** (2009), 888–893.
- [47] C.S. Foster, A.R. Dodson, L. Ambroisine et al., Hsp-27 expression at diagnosis predicts poor clinical outcome in prostate cancer independent of ETS-gene rearrangement, *Br J Cancer* **101** (2009), 1137–1144.
- [48] P.J. Bastian, S. Yegnasubramanian, G.S. Palapattu et al., Molecular biomarker in prostate cancer: the role of CpG island hypermethylation, *Eur Urol* **46** (2004), 698–708.
- [49] M.H. Lee and G. Lozano, Regulation of the p53-MDM2 pathway by 14-3-3 sigma and other proteins, *Semin Cancer Biol* **16** (2006), 225–234.
- [50] S.S. Teoh, J.C. Whisstock and P.I. Bird, Maspin (SERPINB5) is an obligate intracellular serpin, *J Biol Chem* **285** (2010), 10862–10869.
- [51] C.R. Pierson, R. McGowen, D. Grignon et al., Maspin is up-regulated in premalignant prostate epithelia, *Prostate* **53** (2002), 255–262.
- [52] L.C. Li, P.R. Carroll and R. Dahiya, Epigenetic changes in prostate cancer: implication for diagnosis and treatment, *J Natl Cancer Inst* **97** (2005), 103–115.
- [53] I. Krop, M.T. Parker, N. Bloushtain-Qimron et al., HIN-1, an inhibitor of cell growth, invasion, and AKT activation, *Cancer Res* **65** (2005), 9659–9669.