

Identification of specific DNA methylation sites on the Y-chromosome as biomarker in prostate cancer

Lushuai Yao^{1,2,*}, Shancheng Ren^{3,*}, Minjie Zhang^{1,2,*}, Fengxia Du¹, Yasheng Zhu³, Hui Yu^{1,2}, Chenyu Zhang¹, Xiaohua Li¹, Caiyun Yang¹, Huixian Liu¹, Dong Wang^{1,2}, Hao Meng^{1,2}, Shuang Chang^{1,2}, Xiao Han^{1,2}, Yinghao Sun³ and Yingli Sun¹

¹ Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China

² University of Chinese Academy of Sciences, Beijing, China

³ Department of Urology, Shanghai Changhai Hospital, Second Military Medical University, Shanghai, China

* These authors have contribute equally to this work

Correspondence to: Yingli Sun, email: sunyl@big.ac.cn

Yinghao Sun, email: sunyh@medmail.com.cn

Keywords: DNA methylation, prostate cancer, Y-chromosome, biomarker

Received: April 01, 2015

Accepted: September 26, 2015

Published: October 16, 2015

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

As a diagnostic biomarker, prostate special antigen (PSA) tests always generate false positive results and lead to unnecessary and/or repeat biopsies. Therefore, there is an urgent need for developing more sensitive, specific diagnostic biomarkers. We epigenotyped methylated sites in cancer tissues and adjacent normal tissues from 66 patients. In comparison with normal adjacent tissues, we observed that there were 6 aberrant methylation sites in prostate cancer tissues on the Y-chromosome. We further performed pyrosequencing using urine of PCa patients and we identified one methylated site (cg05163709) as a potential biomarker. We evaluated the predictive capacity of the aberrant methylated sites using the area under receiver operating characteristic (ROC) curve (AUC). The ROC analysis showed a higher AUC for cg05163709 (0.915) than prostate-specific antigen (PSA, 0.769). These results indicated that aberrant DNA methylation of cg05163709 on the Y-chromosome could serve as a potential diagnostic biomarker with high sensitivity and specificity.

INTRODUCTION

Prostate cancer (PCa) is the second cancer-related killer for men in the Western world [1]. Early diagnosis is crucial for PCa diagnosis and treatment. Prostate-specific antigen (PSA) test, which has been used in the early diagnosis, lacks specificity and sensitivity, which lead to unnecessary biopsies [2-4]. Therefore, there is an urgent need to develop sensitive, non-invasive and cost-effective early detection biomarkers.

Alteration of DNA methylation is often identified in cancer and is associated with cancer initiation and progression [5, 6]. A number of studies have identified some novel specific DNA methylation sites as biomarkers of PCa. For example, hypermethylation of some important genes, such as *GSTP1*, *PITX2* and the *GABRE~miR-*

452~miR-224, has shown promise as biomarkers of prostate cancer [7, 8].

The Y-chromosome is the sex-determining chromosome in many species [9, 10]. Loss of it can result in some diseases such as hematologic disease, acute myelogenous leukemia and myelodysplastic syndrome [11-14]. Therefore the Y-chromosome plays a critical role in tumor suppression both in vivo and in vitro [15, 16]. In PCa, it is often detected that recurrent copy number loss at Yp11.2 was involved with the *TSPY* gene cluster [17]. However, whether the epigenetic modification of the Y-chromosome is associated with PCa progression is currently unclear.

In this study, we demonstrated the correlation between the DNA methylation of the Y-chromosome and PCa. Comparing PCa with adjacent normal tissues, we

screened 6 methylated sites which were conservative in adjacent normal tissues but varied remarkably in cancer tissues. Two of the six aberrant methylation sites were also detected in the urine of PCa patients using pyrosequencing. In addition, the receiver operating characteristic (ROC) analysis showed that the cg05163709 site was a promising diagnostic biomarker with high sensitivity and specificity.

RESULTS

Characterization of prostate cancer (PCa) tissues and adjacent normal (NA) tissues

We took 24 samples of blood (12 from PCa patients and 12 from healthy individuals) and 66 pairs of PCa and NA (Table S1). Clinical and pathological information

of the study cohort was outlined in Figure 1A. Two pathologist assessed all 66 pairs of PCa adjacent normal tissues using hematoxylin and eosin (H&E)-stained sections (Figure 1B) independently. Most patients were aged 60 to 80 years, and more than half of the patients' PSA concentrations were greater than 10ng/mL (Figure 1C). Chromatin immunoprecipitation (ChIP)-seq, genome sequencing, and DNA methylation microarray were performed to determine the correlation between the DNA methylation pattern of Y-chromosome and PCa (Figure 1D).

H3K4me3 modification on Y-chromosome showed no significant change between PCa and NA.

We performed ChIP-seq of H3K4me3 in cancer tissues and normal adjacent tissues of paired samples from

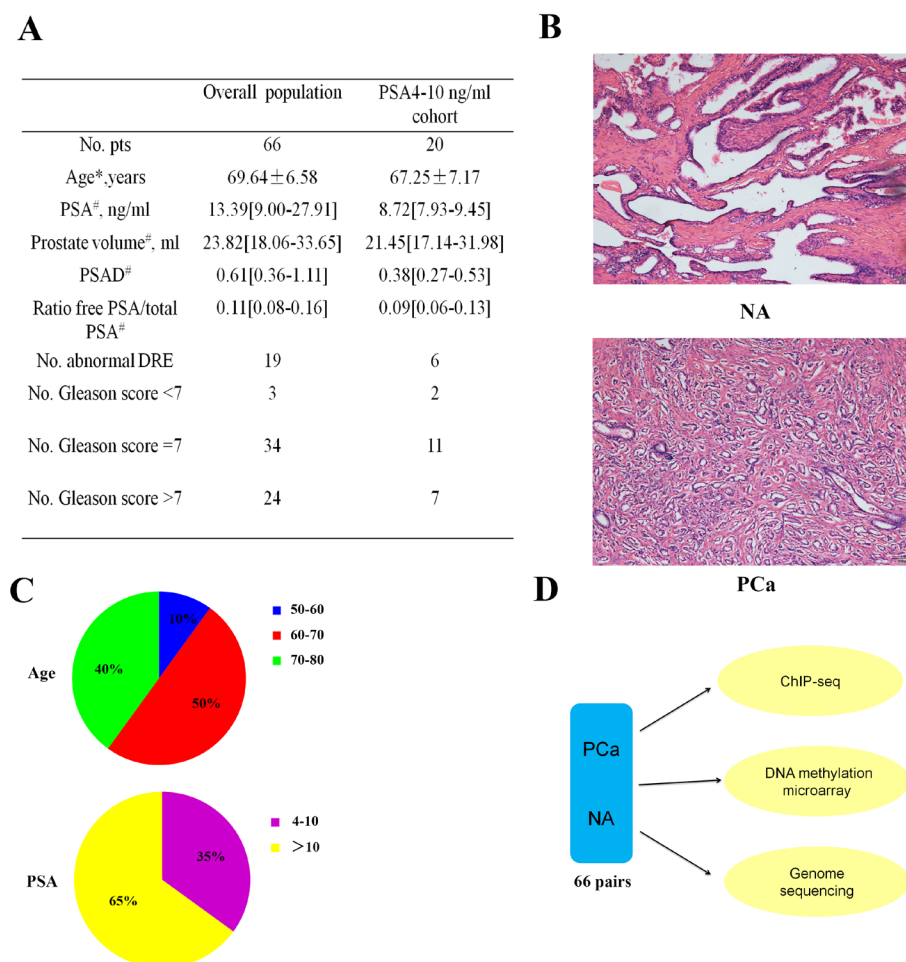


Figure 1: Characteristics of the prostate cancer (PCa) samples. **A.** Clinical and pathological information on the study cohort. (*Age is reported as means ± SD. #These data are reported as medians [interquartile ranges - IQRs]. #Clinically significant PCa according to the Epstein criteria: clinical stage ≥ T2, Gleason score ≥ 7, PSAD > 0.15, and > 33% positive cores on biopsy. Abbreviations: DRE = digital rectal examination; PCa = prostate cancer; PSA = prostate-specific antigen; PSAD = PSA density (serum PSA/prostate volume). **B.** Representative normal adjacent (NA, upper) and PCa tissues (PCa, lower) H&E sections from one patient. Scale bar: 100 μm. **C.** Distributions of the age (years) and PSA concentrations (ng/ml). **D.** Workflow for determining the relationship between the Y-chromosome and PCa.

5 patients. We found no significant difference of H3K4me3 modification between PCa and NA (Figure 2A and 2B). The results indicated that H3K4me3 modification on Y-chromosome could not serve as a potential biomarker.

The methylation levels of certain sites on the Y-chromosome changed remarkably in PCa.

Then we did DNA methylation array using the Illumina 450K methylation microarray platform to detect the methylation level of the Y-chromosome. In order to find out whether aberrant methylation occurred on the Y-chromosome in PCa tissues, we compared the DNA methylation level between cancer tissues and adjacent normal tissues of paired samples from 66 patients. We found that the methylation levels of a number of sites were obviously different in PCa (Figure S1). Through principal component analysis of the methylation levels of all of the tested sites on the Y-chromosome, we found that adjacent normal tissues clustered into a class, indicating that their methylation levels were similar with each other, whereas the PCa tissues were heterogeneous (Figure 3A). Taken

together, these results demonstrated that the methylation levels of the Y-chromosome changed remarkably in PCa. Ultimately, we identified 37 differentially methylated sites (Wilcoxon rank-sum-test, $p < 0.01$; false-discovery rate [FDR]-adjusted $p < 0.01$; $|\Delta\beta\text{-value}| \geq 0.2$), using the Illumina Methylation Analyzer (IMA) package in the R statistical language (Figure 3B, Table S2).

6 of these 37 differential methylation sites were conservative in adjacent normal tissues.

To determine whether the methylation level of these differential methylation sites were only presented variation in cancer tissues while conservative in adjacent normal tissues, we analyzed the DNA methylation level of the Y-chromosome of the 66 samples of adjacent normal tissues for contrast. After filtering 35 sites containing missing values in NA, 381 sites were remained. Finally, we identified 75 conservative sites ($SD < 0.25$) (Figure 4A). Among them, 58.67% were hypomethylated ($\beta\text{-value} \leq 0.25$), and 44.33% were hypermethylated ($\beta\text{-value} \geq 0.75$) (Figure 4B). Notably, we also found that 52.0%

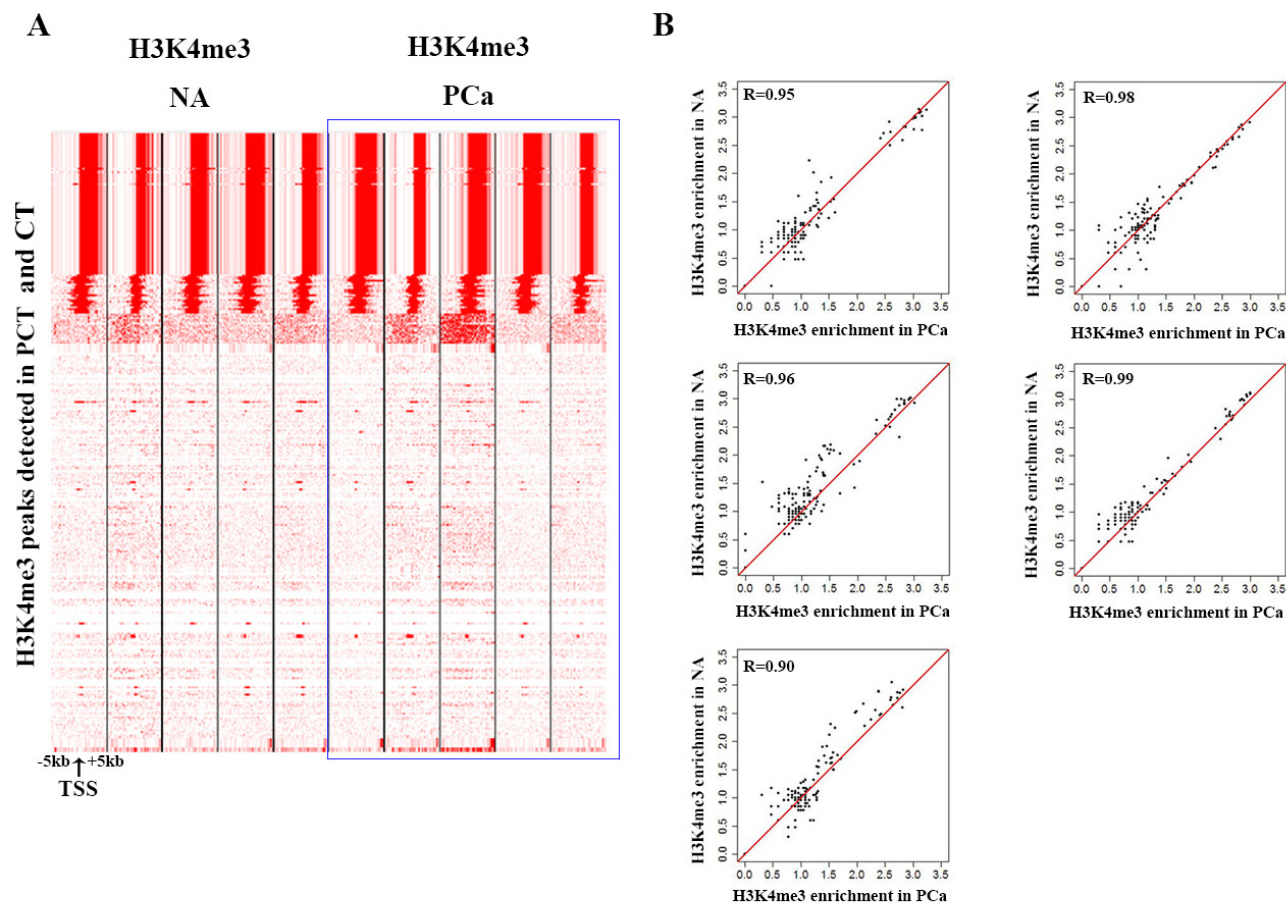


Figure 2: H3K4me3 modification is stable on the Y-chromosome. A. Heatmap displaying the density of aligned ChIP-seq reads for H3K4me3 within a 10-kb window. B. Pearson correlation coefficients between adjacent normal tissues and PCa of five pairs of samples.

of the sites were located in CpG islands (Figure 4C) and 44.3% in promoter regions (Figure 4D).

In addition, we compared the 37 differential methylation sites in cancer tissues to these 75 conservative sites. Eventually, we found that 6 sites (Target ID: cg03052502, cg04462340, cg05163709, cg05544622, cg14466580 and cg27539893) were in the intersection of them (Figure 5A). Interestingly, these sites were highly similar in NA but very different in PCa tissues (Figure S2). cg03052502, cg05163709 and cg05544622 were located in the promoters of *FAM197Y2*, *PRKY* and *TSPY4*, respectively. cg27539893 was located in the 3' untranslated region (UTR) of *TGIF2LY*. cg04462340 and cg14466580 were located in intergenic regions (Figure 5B).

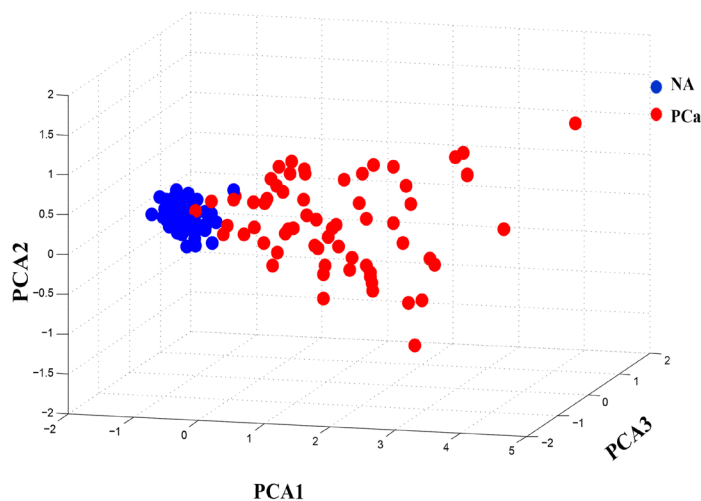
The mean β -value of cg05163709 was higher in PCa than in NA, whereas the mean β -value of other 5 sites were lower in PCa tissues (Figure 6A). All 6 sites

had marked differences in most of the cancer tissues compared with their corresponding NA (Figure S3). One or more of the methylated sites in the panel changed remarkably ($|\Delta\beta\text{-value}| \geq 0.2$) in 57 (86.36%) PCas (Figure 6B). Additionally, the changes in the methylation levels of all 6 sites were not associated with the patients' clinical characteristics (Table 1). These findings strongly suggested that the aberrant methylation levels of these 6 selected sites on the Y-chromosome could be the potential candidate biomarker of PCa.

Aberrant methylation of the 6 selected sites is not caused by single-nucleotide polymorphisms (SNPs)

DNA sequence mutations can influence the methylation levels around mutant sites [18, 19]. To

A



B

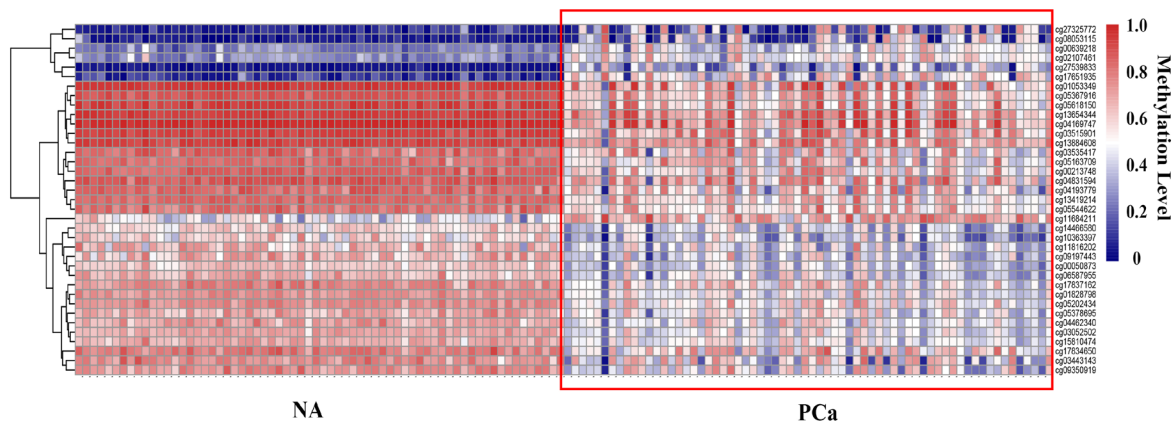


Figure 3: DNA methylation level in prostate cancer (PCa) is clearly different from adjacent normal (NA) tissues. A. The principal component analysis of the methylation pattern on the Y-chromosome between PCa and adjacent normal tissues. **B.** Heat map of the DNA methylation levels of the 37 methylated sites that changed markedly in PCa ($|\Delta\beta\text{-value}| \geq 0.2$, $p < 0.01$, FDR-adjusted $p < 0.01$).

investigate whether the changes in the DNA methylation levels of the 6 sites were caused by base pair mutations, we performed whole-genome sequencing of all 66 pairs of PCa tissues and adjacent normal tissues. We found no obvious mutations on the Y-chromosome in PCa tissues, consistent with previous work [20]. We found only 30 pair base mutations on the Y-chromosome in total, and each mutation occurred in only one sample (Table S3). None of the 6 selected methylated sites was mutated in any tissue sample. This result indicated that the genome sequence of the Y-chromosome was stable in PCa and that the aberrant methylation of the 6 sites was not caused by

base mutations.

Aberrant methylation of certain sites could be detected in urine of PCa patients with high specificity and sensitivity

In order to develop non-invasive biomarker for prostate cancer, we next investigated whether the changes in DNA methylation could be detected in urine sample. We selected 3 sites (cg05163709, cg27539833 and cg14466580) to perform this analysis.

We prospectively collected 135 urine sediment

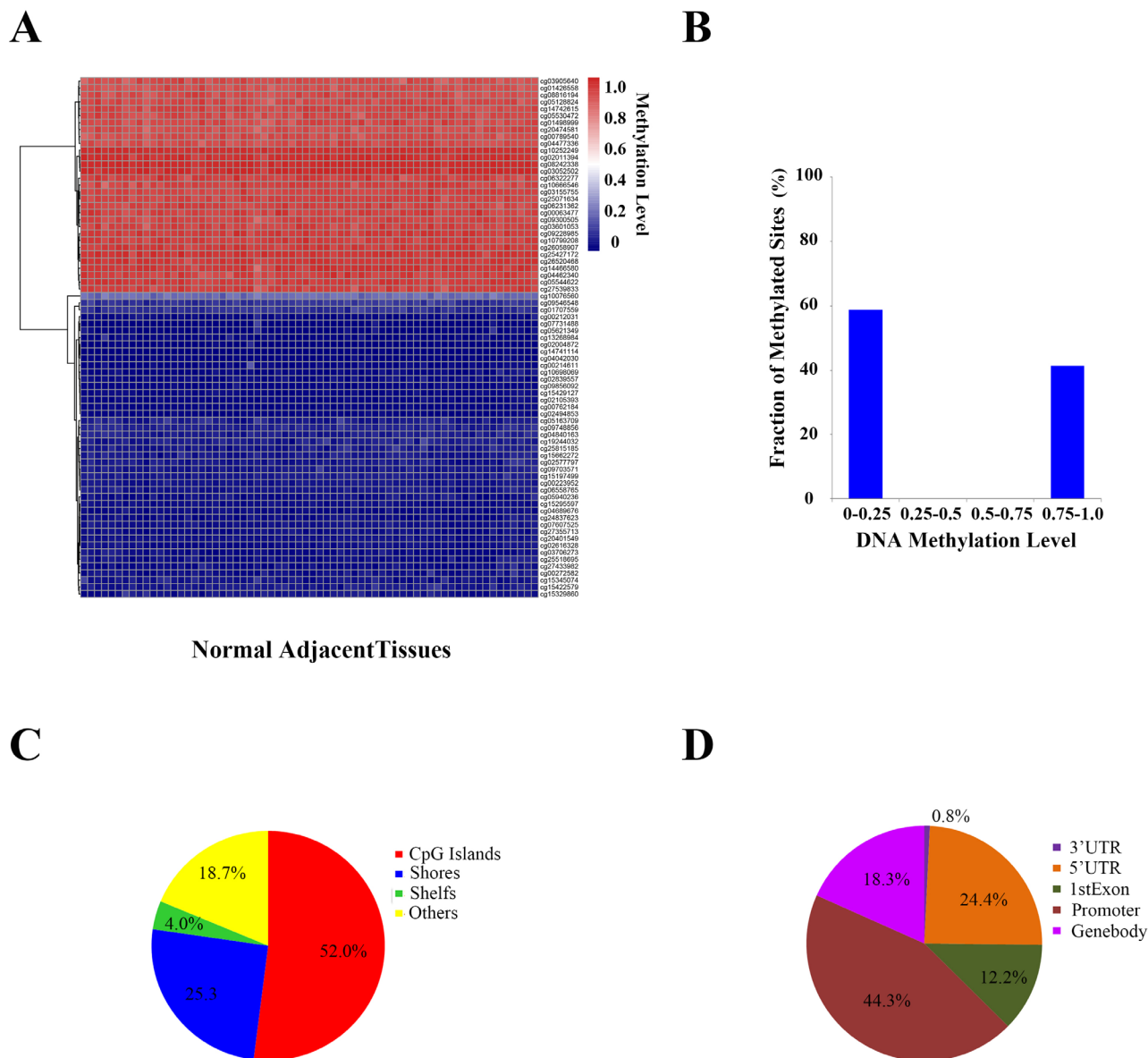


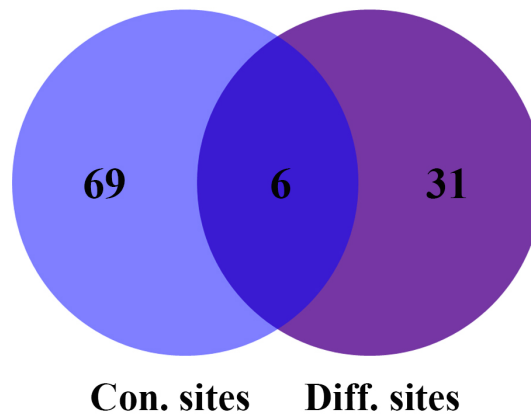
Figure 4: DNA methylation patterns on the Y-chromosome adjacent normal (NA) tissues. **A.** DNA methylation level of the conservative methylated sites on the Y-chromosome in each prostate and adjacent normal tissue. **B.** Ratios of the conservative methylated sites (β -value < 0.25 means low methylation values, β -value > 0.75 means heavy methylation values, and $0.25 \leq \beta$ -value ≤ 0.75 means intermediate methylation values). **C.** Characteristics of the conservative methylated sites. **D.** Location distribution of the conservative methylated sites.

samples; the basic clinical and pathological information from these patients was summarized in Table S4. We measured the DNA methylation levels of cg05163709 CpG site through pyrosequencing from 135 (76+59) samples, and cg27539833 CpG site from 76 samples. The methylation level of cg14466580 was not detected.

As shown in Figure 7A, the methylation levels of cg05163709 and cg27539833 were significantly higher and lower, respectively, in samples with positive biopsy outcomes than in samples with negative results. The areas

under the curve (AUCs) were 0.915 for cg05163709 and 0.769 for PSA. A cg05163709 methylation-level cut-off of 3 provided the optimal balance between sensitivity (94.6%) and specificity (78.3%). Additionally, the diagnostic performance of cg05163709 was statistically superior to PSA (Figure 7B). Although the AUC for cg27539833 also reached 0.729, it was not superior to serum PSA alone (Figure 7C). Taken together, these results indicated that the methylation level of cg05163709 could serve as a potential diagnostic biomarker.

A



B

Target ID	P-Value	Adjust Pval	Beta-Difference	Mean_PCa	Mean_NA	UCSC_REF GENE_ME	UCSC_RE FGENE_GROUP	MAPINFO
cg03052502	4.68E-20	4.82E-19	-0.22197	0.750654	0.972621	FAM197Y 2	TSS1500	9193029
cg04462340	2.30E-23	3.65E-22	-0.20581	0.718943	0.924757			13911503
cg05163709	9.96E-28	5.42E-26	0.284045	0.345909	0.061864	PRKY	TSS1500	7141248
cg05544622	2.47E-19	1.96E-18	-0.20167	0.734588	0.936261	TSPY1; TSPY4	TSS1500; Body	9303646
cg14466580	1.50E-20	1.73E-19	-0.28341	0.663387	0.946797			10037020
cg27539833	9.32E-23	1.27E-21	-0.25101	0.67675	0.927762	TGIF2LY	3'UTR	3447947

Figure 5: 6 aberrant methylated sites were conservative in prostate tissues. A. The Venn diagram of conservative sites and different sites. **B.** Detailed information on the six methylated sites.

Table 1: Association of DNA methylation level change status and clinical characteristics of 66 prostate cancer patients

Clinical characteristics	All		Cg05163709		Cg14466580		Cg03052502		Cg27539833		Cg05544622		Cg04462340		
	Total	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Age, y															
>65	45	39	6	31	14	28	17	25	20	27	18	24	21	24	21
≤65	15	14	1	13	2	10	5	10	5	9	6	9	6	8	7
PSA, ng/ml															
4-10	21	19	2	15	6	13	8	13	8	16	5	11	10	8	13
>10	39	33	6	29	10	25	14	22	17	23	16	22	17	22	17
PCa volume, ml															
≤30	36	31	5	28	8	24	12	23	13	24	12	21	15	17	19
30-50	17	15	2	12	5	10	7	8	9	9	8	10	7	10	7
≥50	7	5	2	4	3	4	3	4	3	3	4	2	5	2	5
Gleason score															
<7	3	2	1	0	3	2	1	2	1	1	2	2	1	1	2
=7	34	29	5	25	9	22	12	20	14	20	14	18	16	15	19
>7	24	22	2	20	4	15	9	14	10	16	8	14	10	15	9
Clinical stage															
I	42	35	7	27	15	28	14	26	16	26	16	26	16	22	20
II	19	18	1	16	3	12	7	11	8	9	10	8	11	7	12
III	5	5	0	4	1	2	3	2	3	3	2	2	3	3	2

*Yes means the difference value of the DNA methylation level between prostate cancer tissues and normal adjacent tissues be equal or great than 0.2.

*Staging information was not available for some patients.

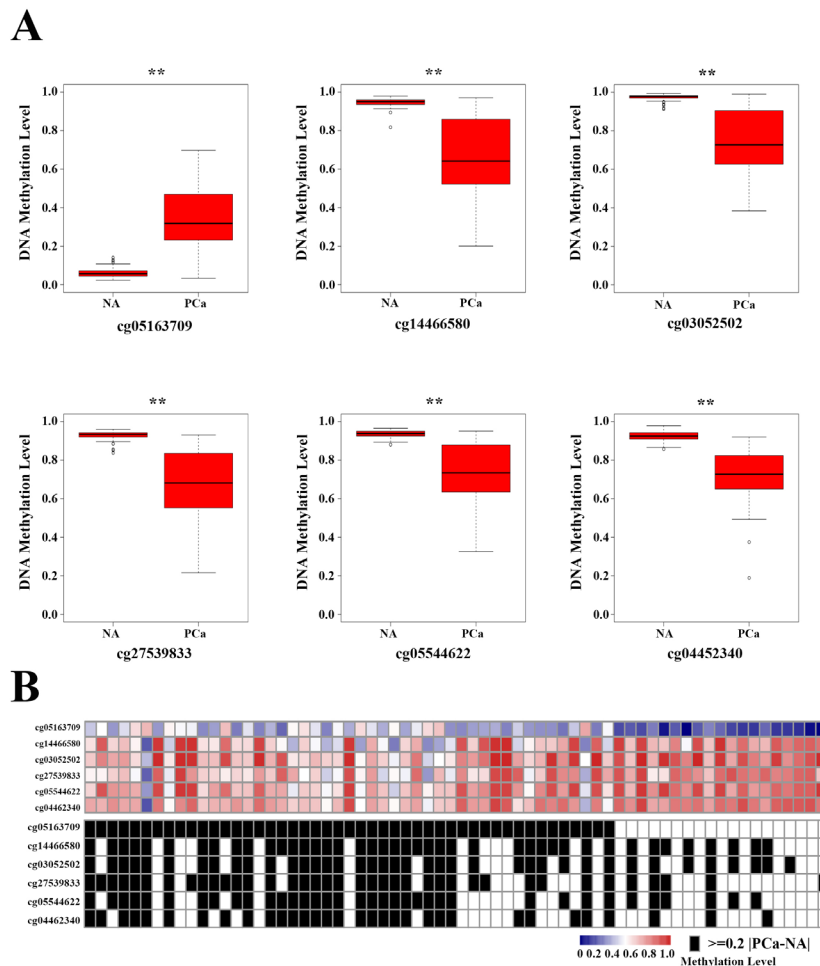


Figure 6: DNA methylated site panel in 66 prostate cancer (PCa) patients. A. The methylation levels of the six methylated sites in adjacent normal tissues and PCa. ****** $p < 0.01$. **B.** Heat map of the selected six methylated sites in PCa. Black boxes in the bottom heatmap indicate PCa patients with obvious differences in the methylated sites ($|\Delta\beta\text{-value}| \geq 0.2$, $p < 0.01$, FDR-adjusted $p < 0.01$ vs controls).

DISCUSSION

Currently, PSA is the most comprehensively used biomarker for PCa. However, the lack of sensitivity and specificity can lead to unnecessary and/or repeat biopsies [21]. Therefore, there is an urgent need for developing more sensitive, specific biomarkers of PCa.

Aberrant DNA methylation could be served as promising cancer biomarker [22-24]. Studies about DNA-based urine biomarkers mostly focused on hypermethylation of gene panels, including GSTP1, RARB and APC, with AUCs of approximately 0.72 [25, 26]. The Y-chromosome is critical for sex determination and sex organ development, including the prostate. Prostate cancer is often accompanied with the loss of Y-chromosome [17]. However, there is no report that links DNA methylation on the Y chromosome to potential biomarkers of PCa, and whether the DNA methylation on the Y-chromosome contributes to PCa development and progression is unclear.

In our current work, we aimed to identify the correlation between PCa and epigenetic modification of the Y-chromosome. First, we performed H3K4me3 ChIP-

seq, but found no significant change between cancer tissues and adjacent normal tissues. Then, we focused on the variations in DNA methylation of the Y-chromosome. Compared with adjacent normal tissues, we found 6 methylated sites (Target ID: cg03052502, cg04462340, cg05163709, cg05544622, cg14466580, cg27539893) that were conservative in adjacent normal tissues but changed markedly in tumor tissues. These changes were highly PCa-specific, and were not age or geographically related. Two sites (cg05163709, cg27539893) were selected for further testing by pyrosequencing using PCa patient urine DNA samples. AUC-ROC analysis showed that the methylation level of cg05163709 generated an AUC of 0.915, with high sensitivity and specificity. One advantage of this method is non-invasive because prostate cells can be detected in the urine of men after a digital rectal examination (DRE). The major drawback of our study is the relatively small sample size. Taken together, our results show that the methylation level of urine cg05163709 is a promising diagnostic biomarker, and our work provides the basis for the validation of this novel biomarker in large cohort-based prospective clinical studies.

The incidence of PCa has considerable racial

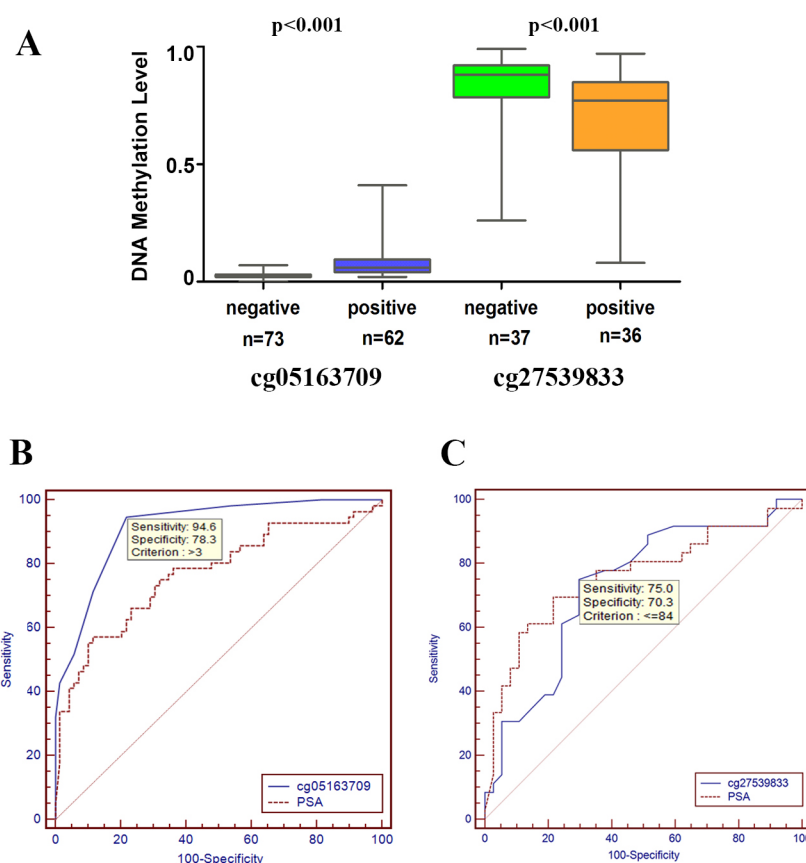


Figure 7: Receiver operating characteristic (ROC) curve analysis for evaluating the diagnostic performance of the methylation levels of cg05163709 and cg27539833. A. Comparison of DNA methylation levels of cg05163709 and cg27539833 between positive and negative biopsies. **B.-C.** Comparison of the diagnostic performance of the methylation level of cg05163709 (B), the methylation level of cg27539833 (C) and PSA.

diversity. It would be interesting to test this biomarker in other ethnic populations. In addition, we would investigate the mechanism of these variations of DNA methylation in our future studies. Our work also provides the possibility of identifying and developing DNA methylation-related biomarkers in other types of cancer.

MATERIALS AND METHODS

Sample collection and characterization

Sixty-six pairs of cancer and adjacent normal tissues from Chinese patients with PCa were supplied by Shanghai Changhai Hospital. DNA was extracted from the samples using a QIAamp DNA Mini Kit (Cat. No.51306, QIAGEN, Hilden, Germany). All of the urine samples were collected from patients who were scheduled for prostate biopsy because of elevated serum PSA (≥ 4 ng/ml) and/or suspicious DRE in Shanghai Changhai Hospital from August 2014 to December 2014. We also collected their clinical and pathological information. The urine sediment samples were collected as previously reported [27]. DNA was extracted from the urine sediment samples using a QIAamp DNA Mini Kit (Cat.No.51306). We obtained written informed consent from all of the patients. The research protocol was approved by the Ethics Committees of Shanghai Changhai Hospital.

Data generation

DNA samples were bisulfate converted using a Zymo EZ DNA Methylation Kit (Zymo Research). After the DNA was amplified and enzymatically fragmented, the fragments were purified with ethanol precipitation and hybridized to the Infinium Human Methylation 450 BeadChip array (Illumina Inc., San Diego, CA, USA). After hybridization overnight, extension, staining and washing were performed successively. Eventually, the BeadChips were imaged using an iScan system (Illumina, Inc.).

Data processing

The raw files were processed for quality control and normalization. Then, the data were analyzed using the official Illumina Methylation Analysis Algorithms. Further we corrected and normalized the deviation caused by different fluorescent markers and probes. Finally, we obtained the methylation level of each tested site as a β -value. β -values ranged from 0 to 1, representing completely unmethylated to fully methylated sites, respectively.

Statistical analysis of methylation levels of functional regions

The IMA (Illumina Methylation Analyzer) package in R language was used to calculate the methylation level of each methylated site. The methylation level of each site is represented by its ratio with the mean methylation level of the tested sites.

Statistical analysis of the diagnostic performance of cg05163709 and cg27539833

Wilcoxon rank-sum test or the Mann-Whitney U test was used to compare the continuous variables that were or were not normally distributed, respectively. The Kruskal-Wallis test was applied to compare more than two groups of variables that were not normally distributed. The categorical variables were compared using Pearson's chi-square test. We used the AUC-ROC to evaluate the diagnostic performance of each methylation site. The comparison of the AUCs for different sites was performed using the z-test method [28, 29]. Statistical analysis was performed with MedCalc statistical software version 10.4.7.0 (MedCalc Software bvba, Mariakerke, Belgium).

Software for data visualization

The 'heatmap' package of R was used for hierarchical clustering. The methylation pattern of the Y-chromosome was loaded into IGV for visualization.

ChIP and ChIP sequencing

ChIP experiments were performed according to the procedure described previously with minimal modification [30]. DNA fragments were purified with the QIAquick PCR purification kit (Qiagen 28106). Sequencing was performed on the Illumina HiSeq 2000 platform. Peaks were called using MACS v1.4.2 [31]. Heatmap counts were generated using seqMINER [32].

Genome sequencing

Paired-end reads were produced using the Hiseq2000 system. Using the Burrows-Wheeler Alignment tool algorithm, the uniquely alignable reads on NCBI37/hg19 were retained for downstream analysis [33]. SNPs were collected by "mpileup" command line in SAMTools as well as UnifiedGenotyper in GATK [34, 35]. Quality recalibration and local realignment were performed in the GATK pipeline before variation calling.

Pyrosequencing

Templates used for pyrosequencing were prepared by bisulfite modified DNA. Each PCR was performed in a 40ul volume containing 0.6ul of each primer, 4ul 10XPCR buffer, 0.5ul QIAGEN hotstart Taq, 31.5ul distilled water, and 2.0ul template DNA (treated by bisulfite, EpiTect Bisulfite Kit, Qiagen). Reactions were incubated at 95°C for 3 min, followed by 50 amplification cycles (95°C for 15s, 54°C for 30s, 72°C for 30s), and then a final elongation step at 72°C for 5 min, with the temperature then held at 4°C. Amplicons were confirmed by agarose gel electrophoresis and purified using a QIAquick Gel Extraction Kit (Qiagen). After mixed with 40ul sequencing buffer (contained 0.5uM sequencing primer), degeneration was performed at 80°C for 2 min. Pyrosequencing was performed using PyroMark ID sequencer.

Primer used in pyrosequencing.

Cg05163709 Forward:

GGAAAGGGGTGATTAATATTTAGTTA

Reverse: 5'-BIOTIN-

CAACCTAATAAAAACTATACAAACACAT

Sequencing primer:

ATAAGTATGTTTAAATTATTGTTTAAAG

Cg27539833 Forward:

GGAATAGTTTGTAAAGAAAAAGGTTAAGAT

Reverse: 5'-BIOTIN-AATTTACCACAATACACA

AAAACTAACTACTTA

Sequencing primer:

AGATTTTGTAGTTTTTTGTCGTTA

ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China (973 Program, Grant No. 2012CB518302, and Grant No. 2013CB911001), the National Natural Science Foundation of China (Grant No. 91019024) and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA01040407).

CONFLICTS OF INTEREST

There is no conflict of interest.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D. Global cancer statistics. *CA: a cancer journal for clinicians*. 2011; 61:69-90.
2. Loeb S, Gashti SN and Catalona WJ. Exclusion of inflammation in the differential diagnosis of an elevated prostate-specific antigen (PSA). *Urol Oncol-Semin Ori*. 2009; 27:64-66.
3. Andriole GL, Grubb RL, Buys SS, Chia D, Church TR, Fouad MN, Gelmann EP, Kvale PA, Reding DJ, Weissfeld JL, Yokochi LA, Crawford ED, O'Brien B, Clapp JD, Rathmell JM, Riley TL, et al. Mortality Results from a Randomized Prostate-Cancer Screening Trial. *New Engl J Med*. 2009; 360:1310-1319.
4. Schroeder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Maattanen L, Bangma CH, Aus G, et al. Screening and Prostate-Cancer Mortality in a Randomized European Study. *New Engl J Med*. 2009; 360:1320-1328.
5. Azhikina TL and Sverdlov ED. Study of tissue-specific CpG methylation of DNA in extended genomic loci. *Biochemistry Biokhimiia*. 2005; 70:596-603.
6. Kim JH, Dhanasekaran SM, Prensner JR, Cao XH, Robinson D, Kalyana-Sundaram S, Huang C, Shankar S, Jing XJ, Iyer M, Hu M, Sam L, Grasso C, Maher CA, Palanisamy N, Mehra R, et al. Deep sequencing reveals distinct patterns of DNA methylation in prostate cancer. *Genome Res*. 2011; 21:1028-1041.
7. Goering W, Kloth M and Schulz WA. DNA methylation changes in prostate cancer. *Methods Mol Biol*. 2012; 863:47-66.
8. Strand SH, Orntoft TF and Sorensen KD. Prognostic DNA methylation markers for prostate cancer. *International journal of molecular sciences*. 2014; 15:16544-16576.
9. Kent-First M. The Y chromosome and its role in testis differentiation and spermatogenesis. *Seminars in reproductive medicine*. 2000; 18:67-80.
10. Serio M. (1984). *Sexual differentiation : basic and clinical aspects*. (New York: Raven Press).
11. Wiktor A, Rybicki BA, Piao ZS, Shurafa M, Barthel B, Maeda K and Van Dyke DL. Clinical significance of Y chromosome loss in hematologic disease. *Genes, chromosomes & cancer*. 2000; 27:11-16.
12. Wong AK, Fang B, Zhang L, Guo X, Lee S and Schreck R. Loss of the Y chromosome: an age-related or clonal phenomenon in acute myelogenous leukemia/myelodysplastic syndrome? *Archives of pathology & laboratory medicine*. 2008; 132:1329-1332.
13. Zhang LJ, Shin ES, Yu ZX and Li SB. Molecular genetic evidence of Y chromosome loss in male patients with hematological disorders. *Chinese medical journal*. 2007; 120:2002-2005.
14. Bianchi NO. Y chromosome structural and functional changes in human malignant diseases. *Mutation research*. 2009; 682:21-27.
15. Vijayakumar S, Garcia D, Hensel CH, Banerjee M, Bracht T, Xiang R, Kagan J and Naylor SL. The human Y chromosome suppresses the tumorigenicity of PC-3, a human prostate cancer cell line, in athymic nude mice. *Genes, chromosomes & cancer*. 2005; 44:365-372.
16. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati

- S, Pakalapati G, Sandgren J, de Stahl TD, Zaghlool A, Giedraitis V, Lannfelt L, Score J, Cross NCP, Absher D, Janson ET, Lindgren CM, et al. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nature genetics*. 2014; 46:624-628.
17. Vijayakumar S, Hall DC, Reveles XT, Troyer DA, Thompson IM, Garcia D, Xiang RH, Leach RJ, Johnson-Pais TL and Naylor SL. Detection of recurrent copy number loss at Yp11.2 involving TSPY gene cluster in prostate cancer using array-based comparative genomic hybridization. *Cancer Res*. 2006; 66:4055-4064.
 18. Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN, Plomin R and Mill J. Allelic Skewing of DNA Methylation Is Widespread across the Genome. *Am J Hum Genet*. 2010; 86:196-212.
 19. Shoemaker R, Deng J, Wang W and Zhang K. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res*. 2010; 20:883-889.
 20. Xu JF, Mo ZN, Ye DW, Wang ML, Liu F, Jin GF, Xu CL, Wang X, Shao Q, Chen ZW, Tao ZH, Qi J, Zhou FJ, Wang Z, Fu YW, He DL, et al. Genome-wide association study in Chinese men identifies two new prostate cancer risk loci at 9q31.2 and 19q13.4. *Nature genetics*. 2012; 44:1231-1235.
 21. Djavan B, Zlotta A, Remzi M, Ghawidel K, Basharkhah A, Schulman CC and Marberger M. Optimal predictors of prostate cancer on repeat prostate biopsy: a prospective study of 1,051 men. *The Journal of urology*. 2000; 163:1144-1148; discussion 1148-1149.
 22. Chiam K, Ricciardelli C and Bianco-Miotto T. Epigenetic biomarkers in prostate cancer: Current and future uses. *Cancer letters*. 2014; 342:248-256.
 23. Vestergaard EM, Nexø E, Tørring N, Borre M, Orntoft TF and Sørensen KD. Promoter hypomethylation and upregulation of trefoil factors in prostate cancer. *International journal of cancer Journal international du cancer*. 2010; 127:1857-1865.
 24. Sørensen KD, Wild PJ, Mortezavi A, Adolf K, Tørring N, Heeboll S, Ulhoi BP, Ottosen P, Sulser T, Hermanns T, Moch H, Borre M, Orntoft TF and Dyrskjot L. Genetic and Epigenetic SLC18A2 Silencing in Prostate Cancer Is an Independent Adverse Predictor of Biochemical Recurrence after Radical Prostatectomy. *Clinical Cancer Research*. 2009; 15:1400-1410.
 25. Roobol MJ, Haese A and Bjartell A. Tumour markers in prostate cancer III: biomarkers in urine. *Acta Oncol*. 2011; 50 Suppl 1:85-89.
 26. Baden J, Green G, Painter J, Curtin K, Markiewicz J, Jones J, Astacio T, Canning S, Quijano J, Guinto W, Leibovich BC, Nelson JB, Vargo J, Wang Y and Wuxiong C. Multicenter evaluation of an investigational prostate cancer methylation assay. *The Journal of urology*. 2009; 182:1186-1193.
 27. Wang F, Ren S, Chen R, Lu J, Shi X, Zhu Y, Zhang W, Jing T, Zhang C, Shen J, Xu C, Wang H, Wang Y, Liu B, Li Y, Fang Z, et al. Development and prospective multicenter evaluation of the long noncoding RNA MALAT-1 as a diagnostic urinary biomarker for prostate cancer. *Oncotarget*. 2014; 5:11091-11102. PMID: 25526029.
 28. Hanley JA and McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982; 143:29-36.
 29. Hanley JA and McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983; 148:839-843.
 30. Yu J, Mani RS, Cao Q, Brenner CJ, Cao X, Wang X, Wu L, Li J, Hu M, Gong Y, Cheng H, Laxman B, Vellaichamy A, Shankar S, Li Y, Dhanasekaran SM, et al. An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell*. 2010; 17:443-454.
 31. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W and Liu XS. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008; 9:R137.
 32. Ye T, Krebs AR, Choukallah MA, Keime C, Plewniak F, Davidson I and Tora L. seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res*. 2011; 39:e35.
 33. Li H and Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25:1754-1760.
 34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25:2078-2079.
 35. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*. 2011; 43: 491-498.