European Journal of Human Genetics (2008) 16, 1521–1525

www.nature.com/ejhg

ARTICLE

Influence of *MUC1* genetic variation on prostate cancer risk and survival

Rona J Strawbridge*[,1,2], Monica Nister[1], Kerstin Brismar[2], Chunde Li[1] and Sara Lindström[3]

[1]*Department of Oncology Pathology, Karolinska Institutet/Karolinska University Hospital, Stockholm, Sweden;* [2]*Department of Molecular Medicine and Surgical Science, Karolinska Institutet/Karolinska University Hospital, Stockholm, Sweden* [3]*Department of Medical Epidemiology and Biostatistics, Karolinska Institutet/Karolinska University Hospital, Stockholm, Sweden*

Gene expression profiling has identified *MUC1* as being significantly overexpressed in prostate cancer with poor clinical outcome after radical surgery, but the molecular mechanisms are still unclear. In this paper, we examined whether the genetic variation in *MUC1* alters prostate cancer risk and progression. We identified five haplotype-tagging single-nucleotide polymorphisms that describe inherited genetic variation in and around *MUC1*. Individual single-nucleotide polymorphisms as well as haplotypes were tested for association with prostate cancer risk and prognosis in 2760 cases and 1722 controls from the Swedish population. We found no association between any single-nucleotide polymorphism or haplotype in the *MUC1* and risk of prostate cancer. Stratifying for disease severity or age of onset did not alter the results. Moreover, we observed no association with *MUC1* variation and prostate cancer-specific survival. Common variants in *MUC1* and the surrounding region are not associated with risk or prognosis of prostate cancer in Swedish men.
European Journal of Human Genetics (2008) **16,** 1521–1525; doi:10.1038/ejhg.2008.131; published online 16 July 2008

Keywords: *MUC1*; prostate cancer; haplotypes; SNPs; risk; survival

Introduction

Gene expression profiling has identified *mucin 1* (*MUC1*) as being significantly overexpressed in prostate cancer, particularly in metastasized disease.[1–3] Owing to its aberrant expression pattern during cancer progression, *MUC1* is an appealing diagnostic marker and a promising therapeutic target,[4,5] but the molecular mechanisms acting are still unclear.

The *MUC1* gene encodes a large cell surface molecule (MUC1) normally expressed only on the apical surfaces of the glandular epithelium;[2] however, transformation is associated with loss of polarity and ubiquitous MUC1 expression.[6] Several isoforms of the MUC1 protein have been described, some of which are produced by alternative mRNA splicing, whereas others are determined by a single-nucleotide polymorphism (SNP), rs4072037, located in exon 2.[7] The differential expressions of isoforms have been described in some cancers[8] including ovarian cancer,[9] gastric cancer[10] and breast cancer.[11]

We hypothesize that the inherited *MUC1* variation, including rs4072037, alters the probability of developing prostate cancer and/or prognosis. We selected five haplotype-tagging SNPs (htSNPs) using the method of haplotype tagging and genotyped them in 2826 prostate cancer patients and 1705 population-based controls originated from a Swedish prostate cancer case–control study (CAPS).

*Correspondence: Dr RJ Strawbridge, Department of Oncology Pathology Karolinska Institutet, Cancer Centrum Karolinska, R8:05, Karolinska University Hospital, Solna, Stockholm 171 76, Sweden.
Tel: +46 085 177 0585; Fax: +46 8 321 047.
E-mail: rona.strawbridge@ki.se
Received 8 February 2008; revised 22 May 2008; accepted 19 June 2008; published online 16 July 2008

Methods

Case–control study

The study population has been described in detail elsewhere.[12] Briefly, we identified and recruited prostate cancer cases from four of the six regional cancer registries in Sweden. The inclusion criterion was biopsy-confirmed (minimum 97% of patients) or cytologically verified adenocarcinoma of the prostate, diagnosed between July 2001 and October 2003. The pathology was carried out by the attending pathologist for each patient and reported to the regional cancer registries. Out of 3648 identified subjects with prostate cancer, 3161 (87%) agreed to participate. DNA samples from blood, TNM stage, Gleason grade and PSA levels at the time of diagnosis were available for 2893 subjects (92%).

Control subjects, who were recruited concurrently with case subjects, were randomly selected from the Swedish Population Registry and matched according to the expected age distribution of cases (groups of 5-year intervals) and geographic region. A total of 2149 out of 3153 control subjects (68%) who were invited subsequently agreed to participate in the study. DNA samples from blood were available for 1781 control subjects (83%). Serum PSA levels were measured for all control subjects but were not used as an exclusion variable.

At the time of this study, DNA was analysed for 2760 cases and 1722 controls. Table 1 presents the clinical characteristics of the study subjects. Each subject provided written informed consent. The study received institutional approval from the ethical board of Karolinska Institutet.

Follow-up

We collected information about prostate cancer-specific mortality for each case subject in CAPS.[13] Subjects were followed until 1 March 2007. The average follow-up time was 3.75 years (range: 0.04–5.9 years). A total of 499 (18%) individuals were deceased during the follow-up, and of those, 338 (12%) had prostate cancer classified as the underlying cause of death.

SNP selection and haplotype block definition

SNP selection was based on phase II data from the International HapMap project.[14] We included only SNPs with a minor allele frequency >5%. By including complete haplotype blocks as defined by Gabriel *et al*,[15] our target region spanned in total 40 kb covering both upstream and downstream of the gene as long as linkage disequilibrium (LD) was maintained. We used the tagSNPs software[16] to select htSNPs with the criteria $R^2_h > 0.95$, that is, the squared correlation between the estimated haplotype dosage explained by the selected SNPs and the true haplotype dosage. Figure 1 demonstrates the haplotype blocks of the region surrounding *MUC1*. The figure was generated using Haploview.[17]

Table 1 Clinical characteristics for the CAPS population

	Cases		*Controls*	
Characteristics	*N = 2760*	*%*	*N = 1722*	*%*
Age (years)				
≤59	631	22.9	277	16.1
60–69	1294	46.9	736	42.7
≥70	835	30.3	709	41.2
PSA levels[a], ng/ml				
< 4	208	7.7	1418	82.4
4–9.99	883	32.9	237	13.8
10–19.99	635	23.6	41	2.4
20–49.99	439	16.3	20	1.2
50–99.99	217	8.1	3	0.2
≥100	304	11.3	2	0.1
T stage				
T0/TX	77	2.8		
T1	1035	37.5		
T2	861	31.2		
T3	684	24.8		
T4	103	3.7		
N stage				
N0/NX	2667	96.6		
N1–N3	93	3.4		
M stage				
M0/MX	2497	90.5		
M1	263	9.5		
Gleason score				
≤4	98	3.9		
5	280	11.1		
6	944	37.4		
7	761	30.1		
8	242	9.6		
9	176	7.0		
10	24	1.0		
Differential grade				
GI/GX	1883	68.2		
GII	569	20.6		
GIII	308	11.2		
Prostate cancer stage[b]				
Localized	1583	57.4		
Advanced	1177	42.6		

[a]PSA was not available for all subjects.
[b]Case subjects were classified as advanced cases if they met at least one of the following criteria: T3/T4, N+, M+, Gleason score of 8–10 or PSA level ≥50 ng/ml. All cases were not classified.

SNP genotyping

Genotyping details have been described earlier.[18] Briefly, we used matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Sequenom Inc., San Diego, CA, USA).[19] PCR assays and associated extension reactions were designed using the SpectroDESIGNER software (Sequenom Inc). Primer sequences are available on request.

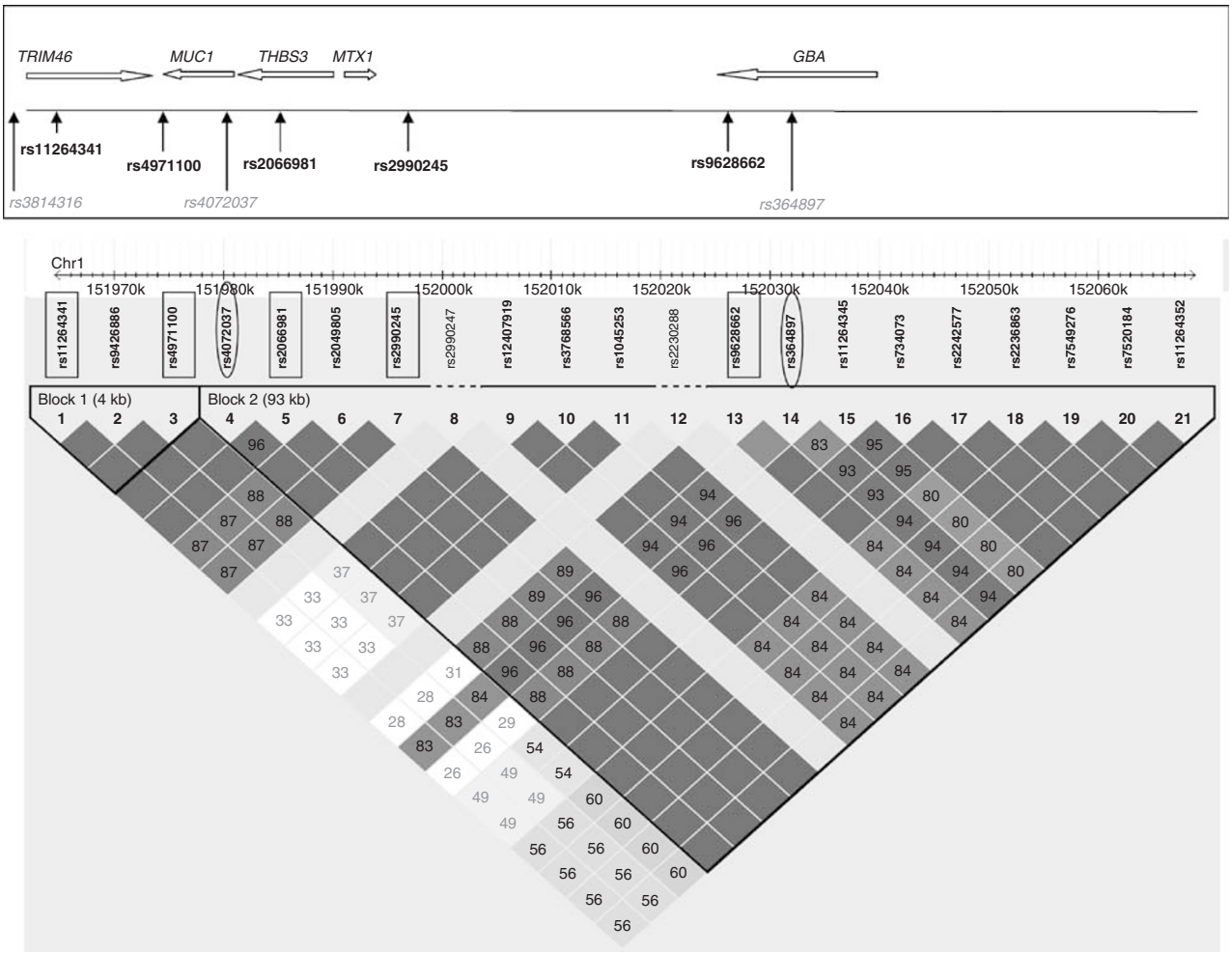

Figure 1 LD plot of the region surrounding *MUC1*. Successfully genotyped SNPs are indicated with rectangles; those that were not successfully assayed are indicated with ovals. The schematic diagram indicates the location and transcriptional direction (arrows) of other genes in the region. Block 2 covers the majority of *MUC1*, with the 3′ UTR extending into block 1. The numbers in the blocks refer to the LD score; red blocks indicate the complete LD and blue blocks indicate the negligible LD.

Statistical methods

We tested for Hardy–Weinberg equilibrium for each SNP using a replication method as implemented in the GENETICS package in the publicly available software R.[20] The association between prostate cancer risk and each SNP was assessed using a likelihood-ratio test of a covariate equal to the number of rare alleles (0, 1 or 2) based on an unconditional logistic regression model as implemented in R. We used the HAPLO.STATS package[21] in R to test the association between *MUC1* haplotypes and prostate cancer risk. Haplotypes with a frequency $<5\%$ were pooled together. We adjusted all analyses for age and geographical region.

The Follow-up began at the date of diagnosis and ended at the date of death or last follow-up (1 March 2007). A likelihood-ratio test of a covariate equal to the number of rare alleles (0, 1 or 2) based on the Cox proportional hazards model was used to test the association between SNP and prostate cancer-specific death. To estimate haplotypic effects on survival, we used the THESIAS software, which allows analysis of censored data using a standard Cox proportional hazards formulation.[22] Hazard ratios and corresponding confidence intervals were estimated for each haplotype by comparison to a reference haplotype chosen as the most frequent one. A likelihood-ratio test was used to perform a global test of association between haplotypes and prostate cancer death. Effects associated with rare haplotypes (frequency <0.05) were not estimated. All *P*-values are based on two-sided tests. All analyses were performed in R and Statistica (Statsoft, USA).

Results

Genotyping failed for two selected htSNPs, rs4072037 and rs3814316. To maintain a high coverage of the genetic

variation in the region, we genotyped two additional SNPs, rs11264341 and rs2990245. The LD blocks and selected htSNPs are demonstrated in Figure 1. rs364897 was monomorphic and thus not further analysed. Average genotyping success for the successfully analysed SNPs was 98.1% (range: 96.9–99.2%). The concordance rate between duplicated samples ($N=320$) was 100%. Among the controls, all SNPs were in Hardy–Weinberg equilibrium ($P>0.05$).

Association analysis

None of the five SNPs analysed demonstrated a significant difference in genotype frequencies (Table 2) between controls and cases, and thus none were associated with prostate cancer. Stratification for disease severity or age of onset did not alter the results. When grouped as sporadic, hereditary or familial cases, no associations were observed. Four htSNP haplotypes had frequencies above 5% in the controls, with a cumulative frequency above 94% (Table 3). A global test of association between *MUC1* haplotypes and prostate cancer risk was not significant (global $P=0.95$, Table 3); nor were tests for specific haplotypes significant (Table 3).

Survival analysis

We observed no association between single SNPs (Table 2) or haplotypes (global $P=0.94$, Table 3) with prostate cancer-specific survival. This strongly suggests that the genetic variation in the *MUC1* region does not alter the probability of developing lethal prostate cancer.

Discussion

This is the first study to investigate the genetic variation in *MUC1*, the products of which have potential as biomarkers for prostate cancer. These results also indicate that the genetic variation in a region of 40 kb surrounding *MUC1* is not associated with prostate cancer. As there are several other coding sequences in the vicinity of *MUC1*, other genes are potentially affected by the chosen SNPs (Figure 1). Of note, none of these other genes have previously been evaluated for the association with prostate cancer.

In addition to splicing/genetic determination of various isoforms influencing the MUC1 protein, post-translational modifications may also contribute to functional differences,[23] as is the case for breast cancer.[5] Therefore, it is feasible that the situation is similar in the prostate, with glycosylation (and other modification) patterns being altered between normal and tumour tissues. The level of MUC1 expression has been reported to be associated with prostate cancer death.[24] However, no information was given on the isoforms assessed, so it remains to be

Table 2 Association between *MUC1* SNPs and prostate cancer risk and survival in CAPS

				Association analysis		*Survival analysis*	
SNP	*Genotype*	*Controls (%)*	*Cases (%)*	*OR (95% CI)*	P	*HR (95% CI)*	P
rs11264341	CC	446 (29.6)	725 (30.1)	1.00 (0.91–1.09)	0.94	1.07 (0.92–1.26)	0.38
	CT	749 (49.6)	1,170 (48.6)				
	TT	314 (20.8)	510 (21.2)				
rs4971100	GG	491 (29.2)	794 (29.5)	1.00 (0.92–1.09)	0.96	1.08 (0.93–1.26)	0.33
	AG	834 (49.7)	1,326 (49.2)				
	AA	354 (21.1)	573 (21.3)				
rs2066981	TT	510 (29.9)	813 (29.8)	0.99 (0.91–1.08)	0.85	0.95 (0.82–1.11)	0.55
	CT	822 (48.2)	1,338 (49.0)				
	CC	372 (21.8)	580 (21.2)				
rs2990245	TT	405 (26.6)	641 (26.4)	1.01 (0.92–1.11)	0.78	0.99 (0.85–1.16)	0.94
	CT	754 (49.5)	1,196 (49.3)				
	CC	365 (24.0)	587 (24.2)				
rs9628662	TT	865 (50.6)	1,369 (50.0)	1.01 (0.92–1.11)	0.85	1.00 (0.85–1.18)	0.98
	GT	684 (40.0)	1,111 (40.6)				
	GG	161 (9.4)	258 (9.4)				

Prostate cancer risk was assessed with an unconditional logistic regression adjusted for age and geographical region. Hazard ratios and corresponding confidence intervals for survival analysis were performed with Cox regression.

Table 3 Association between *MUC1* haplotypes and prostate cancer risk and survival in CAPS

Haplotype					*Haplotype frequency (%)*		*Association analysis*	*Survival analysis*	
rs11264341	*rs4971100*	*rs2066981*	*rs2990245*	*rs9628662*	*Controls*	*Cases*	P	*HR (95% CI)*	P
C	G	C	C	T	45.0	44.8	0.93	1.00 (Reference allele)	
T	A	T	T	T	21.8	21.3	0.67	1.04 (0.86–1.25)	0.69
T	A	T	T	G	20.2	20.3	0.98	1.03 (0.85–1.24)	0.79
C	G	T	T	G	7.8	8.0	0.72	0.95 (0.71–1.26)	0.71

determined whether levels of total MUC1 or specific isoforms are important in cancer progression.

The MUC1 protein isoforms produced by SNP rs4072037 have previously been associated with various cancers.[9,11] Although rs4072037 failed in our analysis, we genotyped the adjacent rs2066981 marker in strong LD with rs4072037 ($r^2 = 0.96$), thus providing basically the same genetic information. We, however, observed no correlation between rs2066981 and prostate cancer risk or survival.

Previous studies on MUC1 variants in cancer have analysed mRNA expression patterns in tissue from the relevant organ. The variants (determined by rs4072037) observed in ovarian cancer[9] and breast cancer cell lines[11] were predominantly associated with the G-allele. In contrast, genomic DNA from blood samples was analysed in this study. The only other report on *MUC1* variants stemming from this SNP in prostate cancer[25] demonstrated the loss of heterozygosity between blood and tumour DNA, with the G-allele being lost. Thus, it may be that the regulation of rs4072037 variants occurs in a tissue- and disease-specific manner.

The study population used has a power of 85% to detect the association of an SNP with minor allele frequency of 0.2 and an odds ratio of 1.3 (assuming an additive inheritance model). Our large sample size makes CAPS a well-powered study with a high probability of detecting a true casual allele through the association. Therefore, we believe that this is a true negative finding.

In summary, the genetic variation in 40 Kb surrounding *MUC1* does not influence the risk of prostate cancer, disease severity or prostate cancer-specific survival.

Acknowledgements

The authors thank all the study subjects who participated in the CAPS study and the urologists who included their patients. We also thank the Regional Cancer Registries and the CAPS steering committee including Drs Jan Adolfsson, Jan-Erik Johansson and Eberhart Varenhorst. We are thankful to the personnel at the Medical Biobank in Umeå for skilfull handling of the blood samples and the personnel at Mutation Analysis Facility (MAF) at Karolinska Insititutet for excellent genotyping. The authors declare that no conflicts of interest exist. This work was financially supported by the Swedish Cancer Foundation, Karolinska Institutet, Karolinska University Hospital and the Cancer Society in Stockholm, the Radiumhemmets forskningsfonder, Cancer- and Allergifonden and the Erling–Persson Foundation.

References

1 Lapointe J, Li C, Higgins JP *et al*: Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci USA* 2004; **101**: 811–816.
2 Arai T, Fujita K, Fujime M, Irimura T: Expression of sialylated MUC1 in prostate cancer: relationship to clinical stage and prognosis. *Int J Urol* 2005; **12**: 654–661.
3 Cozzi PJ, Wang J, Delprado W *et al*: MUC1, MUC2, MUC4, MUC5AC and MUC6 expression in the progression of prostate cancer. *Clin Exp Metastasis* 2005; **22**: 565–573.
4 Gendler SJ: MUC1, the renaissance molecule. *J Mammary Gland Biol Neoplasia* 2001; **6**: 339–353.
5 Stepensky D, Tzehoval E, Vadai E, Eisenbach L: O-glycosylated versus non-glycosylated MUC1-derived peptides as potential targets for cytotoxic immunotherapy of carcinoma. *Clin Exp Immunol* 2006; **143**: 139–149.
6 Wesseling J, van der Valk SW, Hilkens J: A mechanism for inhibition of E-cadherin-mediated cell-cell adhesion by the membrane-associated mucin episialin/MUC1. *Mol Biol Cell* 1996; **7**: 565–577.
7 Ligtenberg MJ, Gennissen AM, Vos HL, Hilkens J: A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. *Nucleic Acids Res* 1991; **19**: 297–301.
8 Fowler JC, Teixeira AS, Vinall LE, Swallow DM: Hypervariability of the membrane-associated mucin and cancer marker MUC1. *Hum Genet* 2003; **113**: 473–479.
9 Obermair A, Schmid BC, Packer LM *et al*: Expression of MUC1 splice variants in benign and malignant ovarian tumours. *Int J Cancer* 2002; **100**: 166–171.
10 Carvalho F, Seruca R, David L *et al*: MUC1 gene polymorphism and gastric cancer an epidemiological study. *Glycoconj J* 1997; **14**: 107–111.
11 Schmid BC, Buluwela L, Liu Q *et al*: Expression of MUCI splice variants correlates with invasive growth of breast cancer cell lines. *Breast Cancer Res Treat* 2002; **76**: 211–219.
12 Zheng SL, Sun J, Wiklund F *et al*: Cumulative association of five genetic variants with prostate cancer. *N Engl J Med* 2008; **358**: 910–919.
13 Lindstrom S, Adami HO, Balter KA *et al*: Inherited variation in hormone-regulating genes and prostate cancer survival. *Clin Cancer Res* 2007; **13**: 5156–5161.
14 Frazer KA, Ballinger DG, Cox DR *et al*: A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851–861.
15 Gabriel SB, Schaffner SF, Nguyen H *et al*: The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
16 Stram DO, Haiman CA, Hirschhorn JN *et al*: Choosing haplotype-tagging SNPS based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. *Hum Hered* 2003; **55**: 27–36.
17 Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
18 Lindstrom S, Wiklund F, Adami HO, Balter KA, Adolfsson J, Gronberg H: Germ-line genetic variation in the key androgen-regulating genes androgen receptor, cytochrome P450, and steroid-5-alpha-reductase type 2 is important for prostate cancer development. *Cancer Res* 2006; **66**: 11077–11083.
19 Jurinke C, van den Boom D, Cantor CR, Koster H: Automated genotyping using the DNA MassArray technology. *Methods Mol Biol* 2001; **170**: 103–116.
20 http://lib.stat.cmu.edu/R/CRAN/index.html.
21 Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA: Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002; **70**: 425–434.
22 Tregouet DA, Tiret L: Cox proportional hazards survival regression in haplotype-based association analysis using the Stochastic-EM algorithm. *Eur J Hum Genet* 2004; **12**: 971–974.
23 Parry S, Hanisch FG, Leir SH *et al*: N-Glycosylation of the MUC1 mucin in epithelial cells and secretions. *Glycobiology* 2006; **16**: 623–634.
24 Andren O, Fall K, Andersson SO *et al*: MUC-1 gene is associated with prostate cancer death: a 20-year follow-up of a population-based study in Sweden. *Br J Cancer* 2007; **97**: 730–734.
25 Strawbridge RJ, Nistér M, Brismar K, Grönberg H, Li C: MUC1 as a putative prognostic marker for prostate cancer. *Biomarker Insights* 2008; **3**: 303–315.