

# Sequence Variants of *Toll-Like Receptor 4* Are Associated with Prostate Cancer Risk: Results from the CAncer Prostate in Sweden Study

S. Lilly Zheng,<sup>1</sup> Katarina Augustsson-Bälter,<sup>2</sup> Baoli Chang,<sup>1</sup> Maria Hedelin,<sup>2</sup> Liwu Li,<sup>1</sup> Hans-Olov Adami,<sup>2</sup> Jeanette Bensen,<sup>1</sup> Ge Li,<sup>1</sup> Jan-Erik Johnsson,<sup>4</sup> Aubrey R. Turner,<sup>1</sup> Tamara S. Adams,<sup>1</sup> Deborah A. Meyers,<sup>1</sup> William B. Isaacs,<sup>5</sup> Jianfeng Xu,<sup>1</sup> and Henrik Grönberg<sup>3</sup>

<sup>1</sup>Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina; <sup>2</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; <sup>3</sup>Department of Radiation Sciences, Oncology, University of Umeå, Umeå, Sweden; <sup>4</sup>Department of Urology and Clinical Medicine, Örebro University Hospital, Sweden, and Regional Oncological Center, University Hospital, Uppsala, Sweden; and <sup>5</sup>Department of Urology, Johns Hopkins Medical Institutions, Baltimore, Maryland

## ABSTRACT

Inflammation has been implicated as an etiological factor in several human cancers. Growing evidence suggests that chronic inflammation may also play a role in the etiology of prostate cancer. Considering that genetic susceptibility is a major risk factor for this disease, we hypothesize that sequence variants in genes that regulate inflammation may modify individual susceptibility to prostate cancer. The lipopolysaccharide receptor Toll-like receptor 4 (*TLR4*) is a central player in the signaling pathways of the innate immune response to infection by Gram-negative bacteria and is an important candidate inflammatory gene. We performed a systematic genetic analysis of *TLR4* sequence variants by evaluating eight single-nucleotide polymorphisms that span the entire gene among 1383 newly diagnosed prostate cancer patients and 780 age- and residence-matched controls in Sweden. We found an association between a sequence variant (11381G/C) in the 3'-untranslated region of the *TLR4* gene and prostate cancer risk. The frequency of the variant genotypes (CG or CC) was significantly higher in the patients (24.1%) than in the controls (19.7%;  $P = 0.02$ ). The frequency of risk genotypes among patients diagnosed before the age of 65 years was even higher (26.3%). Compared with men who had the wild-type genotype of this single-nucleotide polymorphism (GG), those with GC or CC genotypes had a 26% increased risk for prostate cancer (odds ratio, 1.26; 95% confidence interval, 1.01–1.57) and 39% increased risk for early onset prostate cancer (before age 65 years; odds ratio, 1.39; 95% confidence interval, 1.02–1.91). The risk attributable to this variant for prostate cancer in Sweden was estimated to be 4.9%. Although the biological mechanism of the observed association remains to be elucidated, our finding supports a role for a bacteria-associated response pathway, possibly acting via inflammation, in the development of prostate cancer.

## INTRODUCTION

Chronic or recurrent inflammation is known to play a causative role in the development of many human cancers, including cancers of the liver, esophagus, stomach, large intestine, and urinary bladder (1). Inflammatory changes have long been recognized in prostate tissues, leading to speculation that chronic inflammation might contribute to prostate cancer development (2). The role of prostate inflammation in prostate cancer was strongly implicated in the recently proposed theory that proliferative inflammatory atrophy serves as a precursor to prostatic intraepithelial neoplasia and to prostate cancer (3, 4). The identification of two candidate prostate cancer susceptibility genes (*RNASEL* and *MSRI*) that encode proteins with critical functions in

host responses to infections provides additional support for a role of inflammation in prostate cancer development (5, 6).

Chronic infection and inflammatory processes that may lead to tumorigenesis are mediated in part through recognition of various stimuli by Toll-like receptors (TLRs; Ref. 7). Among these TLRs, *TLR4* recognizes Gram-negative bacterial products, including lipopolysaccharide (8), the antitumor compound taxol in the mouse but not in humans (9), and human heat shock protein 60 (10). Differential activation of *TLR4* by this array of naturally occurring or synthetic ligands subsequently induces distinct downstream processes, including the expression of inflammatory genes as well as regulation of cell growth/apoptosis. Conceivably, improper regulation or compromised function of *TLR4* may contribute to various inflammatory diseases, including cancer.

On the basis of the potential importance of inflammation and inflammatory genes in prostate cancer development, we hypothesized that sequence variants of *TLR4* are associated with prostate cancer susceptibility. To test this hypothesis, we performed a systematic genetic analysis in a large population-based prostate cancer case-control study in Sweden.

## MATERIALS AND METHODS

**Study Population.** The cases studied came from the large-scale, population-based case-control study CAncer Prostate in Sweden (CAPS). The case participants were recruited from four of the six regional cancer registries that cover the entire population of Sweden. Each of these four registries serves one health care region (Northern, Central, Stockholm, and South Eastern); the four registries altogether encompass ~6 millions inhabitants (67% of Sweden's population). Reporting of newly diagnosed cancer cases to the registries is required by law for both the attending physician and pathologist; therefore, the registries include almost 100% of all cancers diagnosed in Sweden.

In the CAPS study, the source-person-time comprises men living in the area of Örebro and the northern part of Sweden (Västernorrland, Jämtland, Västerbotten, and Norrbotten) from January 1, 2001, as well as men living in the areas of Västmanland, Södermanland, Gävleborg, Dalarna, Värmland, and Uppland from July 1, 2001, until September 2002 (except for Jämtland and the county of Lycksele in Västerbotten, where the source-person-time ended March 1, 2002). The source-person-time was divided into two age-specific study bases. The first study base included men age 35–65 years of age living in all of the regions mentioned above. The second study base included men 66–79 years of age living only in the areas of Örebro, Västmanland, Södermanland, and the northern part of Sweden.

The inclusion criterion for cases in CAPS was pathologically or cytologically verified adenocarcinoma of the prostate (International Classification of Disease for Oncology = C61). After receiving notification of a new case, the administrator at the regional cancer registry mailed a letter to the treating physician informing him or her about the study. The physician was asked to indicate whether the patient was able to participate in the study. If so, the physician mailed a letter to the patient to introduce the study and asked him to send a reply letter to the administrator at the cancer registry. After approval from both the physician and the patient, the study secretaries sent a questionnaire and a kit with tubes for blood sampling to the eligible case. The self-administered questionnaire included questions concerning such items as

Received 10/18/03; revised 2/6/04; accepted 2/9/04.

**Grant support:** Grants from the Swedish Cancer Foundation and Spear Grant from the Umeå University Hospital (Umeå, Sweden). Also partially funded by the Center for Human Genomics at Wake Forest University School of Medicine.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Requests for reprints:** Jianfeng Xu, Wake Forest University School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157. Phone: (336) 713-7500; Fax: (336) 713-7566; E-mail: jxu@wfubmc.edu.

diet (validated food questionnaire), family history, smoking, and physical activity.

In total, 1961 prostate cancer cases were invited to participate; of these, 1444 (73.6%) agreed to participate by donating a blood sample and/or answering the questionnaire. DNA was available for 1383 (95.8%) of the cases that participated. Clinical data that were not included in the Cancer Registry were obtained from the National Prostate Cancer Registry.<sup>6</sup> The cases were linked to the National Prostate Cancer Registry, and clinical information such as tumor-node-metastasis (TNM) stage, Gleason sum, prostate-specific antigen (PSA) level at the time of diagnosis, means of diagnosis, and primary treatment were obtained for 95.3% of the cases. The cases were thereafter classified as either localized ( $T_{1-2}$  and  $N_0/N_X$  and  $M_0/M_X$ ; grade I-II; Gleason sum, 2–7; and PSA <100 ng/ml) or advanced (prone to progressive disease;  $T_{3/4}$  or  $N+$  or  $M+$ , or grade III, or Gleason sum of 8–10, or PSA >100 ng/ml).

Control subjects were randomly selected from the updated Swedish Population Registry based on frequency matching to the expected age distribution (within 5 years) and geographic origin of the cases. After the controls were identified, a letter of introduction to the study was mailed to each control. Three to 4 weeks later, the same questionnaire and blood sampling kit that were used for the cases were mailed to the controls. Of the 1697 randomly selected controls invited, 866 (52.0%) agreed to participate by donating a blood sample and/or completing a questionnaire. DNA was available for 780 (90.9%) of the controls that participated. Eight potential control subjects were excluded after linkage to the National Cancer Registry revealed that they had a diagnosis of prostate cancer before inclusion.

To improve the response rate, cases and controls were recontacted three times; after 1–2 weeks with a follow-up letter, after 6–8 weeks with a new questionnaire and blood draw kit, and after ~12 weeks with a phone call. The clinical characteristics of the study subjects are presented in Table 1. Mean age (age at diagnosis for cases and age at inclusion for controls) for the cases and controls were 66.60 and 67.90 years, respectively.

This study was approved by the Ethical Committees at the Karolinska Institutet and at Umeå University. Written informed consent was obtained from each subject.

**Genotyping Methods.** All participants in this study were instructed to donate blood (4 × 10 ml) at the nearest health clinic or hospital. Samples were thereafter kept at room temperature and sent by overnight mail to the Medical Biobank at Umeå University. After arrival at the Biobank, leukocytes, erythrocytes, plasma, and serum were separated into different tubes. Samples were stored at –70°C until time for analysis. DNA samples were extracted from whole blood by standard methods and were shipped from Umeå, Sweden to the genotyping laboratory in the Center for Human Genomics at Wake Forest University. Each DNA plate contained 2 Centre d'Etude du Polymorphisme Humain controls, a water blank, and blinded internal replicates. Researchers at Wake Forest University were blinded to case status. Genotyping was performed with the MassARRAY system (SEQUENOM). SpectraDesign software was used to generate the primers. Primer sequences and PCR conditions for these sequence variants are available at the authors' website.<sup>7</sup>

**Selection of *TLR4* single-nucleotide polymorphisms (SNPs).** The genomic structure of the *TLR4* gene has not been completely elucidated. The gene is ~11.5 kb and is composed of four exons. There are four known *TLR4* transcript isoforms (A, B, C, and D) that result from alternative splicing of the four exons and different translation start sites. The poly(A) tail has been annotated only for isoform D. Our goal in this study was to evaluate common haplotypes of *TLR4* sequence variants with use of a limited number of SNPs. To achieve this goal, we first selected a subset of reported SNPs from the National Center for Biotechnology Information (NCBI) dbSNP and IIPGA databases, using the criteria of minor allele frequencies ≥5% and a density of 1 SNP/kb across the *TLR4* targeted genomic region, including 2 kb of the promoter, all exons, introns, and the predicted 3'-untranslated region (UTR). We also selected a subset of functional and coding SNPs regardless of the frequency and density. A total of 18 SNPs were selected (Fig. 1), including 4 reported nonsynonymous changes (D299G, V310G, E474K, and Q510H, defined based on isoform A). The previously described nonsynonymous change T399I was not selected because it has been reported to be in strong linkage disequilibrium (LD) with the nonsynonymous change D299G (11). We

Table 1 Characteristics of study subjects in the Cancer Prostate in Sweden 1 study

	Age at diagnosis			
	45–65 years		66–80 years	
	n	%	n	%
T stage				
T <sub>0</sub>	3	0.45	3	0.46
T <sub>1</sub>	259	39.24	200	30.96
T <sub>2</sub>	234	35.45	208	32.20
T <sub>3</sub>	144	21.82	207	32.04
T <sub>4</sub>	20	3.03	28	4.33
Missing	44		33	
N stage				
N <sub>0</sub>	157	85.79	66	77.65
N <sub>1</sub>	26	14.21	19	22.35
Missing	521		594	
M stage				
M <sub>0</sub>	328	83.04	250	79.62
M <sub>1</sub>	67	16.96	64	20.38
Missing	309		365	
Gleason score				
≤4	29	4.67	24	4.21
5	79	12.72	63	11.05
6	236	38.00	185	32.46
7	177	28.50	188	32.98
8	53	8.53	70	12.28
9	40	6.44	37	6.49
10	7	1.13	3	0.53
Missing	83		109	
Differential				
GI	35	13.51	34	13.23
GII	154	59.46	151	58.75
GIII	70	27.03	72	28.01
Missing	445		422	
PSA <sup>a</sup> levels (ng/ml)				
<4	49	7.46	26	4.07
4–9.99	264	40.18	169	26.45
10–19.99	139	21.16	152	23.79
20–49.99	92	14.00	128	20.03
50–99.99	49	7.46	77	12.05
≥100	64	9.74	87	13.62
Missing	47		40	

<sup>a</sup> PSA, prostate-specific antigen.

then genotyped these SNPs in a subset of 96 control subjects. Nine of these SNPs, including E474K and Q510H, were monomorphic in these 96 subjects. The nonsynonymous change V310G was observed only once. The remaining eight SNPs were observed multiple times and thus were selected for genotyping among all case and control subjects whose DNA samples were available at the time of this study.

***TLR4* RNA Analyses.** The 3'-UTR boundary of *TLR4* mRNA was determined by 3' rapid amplification of cDNA end (RACE) on a prostate poly(A)+ RNA template. We used the SMART RACE cDNA amplification kit (Clontech) according to the manufacturer's recommendations. The gene-specific primer used in the RACE reaction was 5'-ggatccctccctgtaccctctcact-gccaggag-3'.

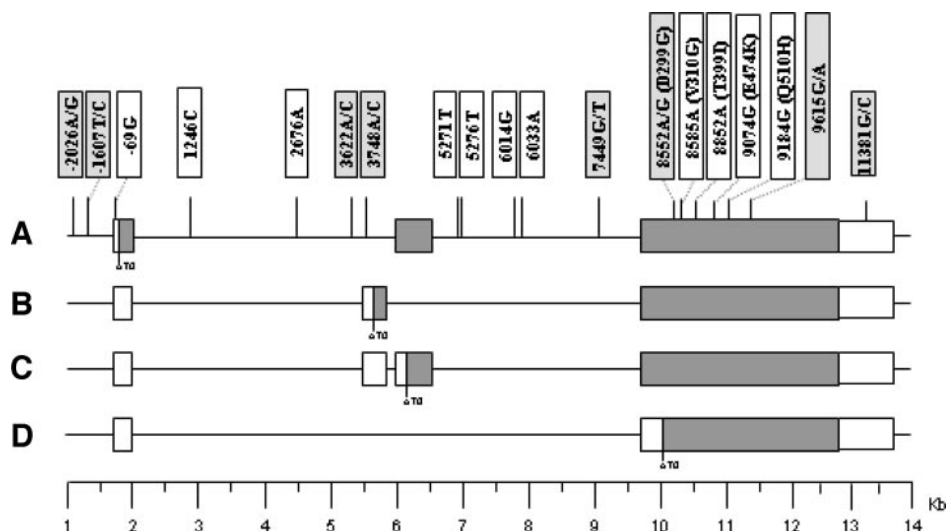
**Statistical Analysis.** Hardy-Weinberg equilibrium tests for each of the eight sequence variants and pair-wise LD tests for all pairs of the eight sequence variants were performed with the Fisher probability test statistic, as described by Weir (12). For each test, 10,000 permutations were performed, and the test statistic of each replicate was calculated. Empirical *P*s for each test were estimated as the proportion of replicates found to be as probable as or less probable than the observed data, as implemented in the software package Genetic Data Analysis. The estimates of pair-wise LD (*D'*) were calculated with the SAS/Genetics computer software package.

Allele frequency differences between the two groups were tested for each SNP by the  $\chi^2$  test with 1 degree of freedom (*df*). Genotype frequency differences were also tested by the  $\chi^2$  test with 2 degrees of freedom. Both tests were performed with the SAS/Genetics computer program. Odds ratios (ORs) of prostate cancer for the variant-allele carriers (homozygous and heterozygous) versus homozygous wild-type allele carriers were estimated by unconditional logistic regression and adjusted for age and geographic region. Attributable risk was estimated using the formula:  $100\% \times p(\text{OR} - 1) / [p(\text{OR} - 1) + 1]$ , where *p* is the prevalence of risk genotypes in a population (13).

<sup>6</sup> <http://www.roc.se>.

<sup>7</sup> [www.wfubmc.edu/genomics](http://www.wfubmc.edu/genomics).

Fig. 1. Toll-like receptor 4 (*TLR4*) genomic structure, with four known transcript isoforms (A, B, C, and D). Sequence variants are indicated at the top of the figure. All variants shown were screened in 96 controls. Shaded variants were genotyped in all study subjects.



Haplotype frequency was estimated with the new statistical method proposed by Stephens *et al.* (14), as implemented in the computer program PHASE.<sup>8</sup> When several runs were performed with different values for the seed of the random number generator, the goodness-of-fit values were similar among the different runs. Association between the haplotypes and prostate cancer risk was performed with a score test developed by Schaid *et al.* (15), as implemented in the computer program HAPLO.SCORE.<sup>9</sup>

## RESULTS

Eight SNPs, including one nonsynonymous change (D299G), were genotyped in all study subjects, including 1383 cases and 780 controls. Because the SNP 7449G/T (rs2149356) significantly deviated from Hardy–Weinberg equilibrium in both cases ( $P < 0.0001$ ) and controls ( $P = 0.0006$ ), it was removed from further statistical analysis. The remaining seven SNPs were in Hardy–Weinberg equilibrium among cases and controls, respectively (all  $P > 0.05$ ). These SNPs were in strong LD: most of the pair-wise  $D'$  estimates were 1.0 with the lowest one being 0.63.

Testing for allele frequency differences between cases and controls revealed one SNP (11381G/C) with a marginally significant difference between cases and controls (Table 2). The frequency of the C allele for the SNP 11381G/C was 0.13 and 0.11 in cases and controls, respectively ( $P = 0.05$ ). The genotype frequencies of CC or CG for this SNP were 24.1% and 19.7% in cases and controls, respectively. The difference in genotype frequencies between cases and controls was larger ( $\chi^2 = 7.77$ ;  $df, 2$ ;  $P = 0.02$ ). In logistic regression analyses (Table 3), men who had the genotypes CC or CG had a 26% increased risk (OR = 1.26, 95% confidence interval, 1.01–1.57) for prostate cancer compared with men with the genotype GG for this SNP (adjusted for age and geographic regions). Furthermore, men who had the genotypes CC or CG were at a 39% increased risk (OR = 1.39; 95% confidence interval, 1.02–1.91) for an early diagnosis of prostate cancer. The high-risk genotypes of SNP 11381G/C were present in 19.7% of control individuals, 24.1% of all cases, and 26.3% of cases diagnosed at age <65 years. The overall proportion of prostate cancer risk in this population that was attributable to the risk genotypes was 4.9%.

There was no statistically significant association between the other six SNPs and prostate cancer risk. The frequency of the variant allele

G of the nonsynonymous change D299G was similar in cases (0.051) and in controls (0.057;  $P = 0.31$ ). Among the prostate cancer patients only, no association was found between any of the SNPs and clinical characteristics such as age of diagnosis (<65 versus  $\geq 65$  years), Gleason score (<8 and  $\geq 8$ ), or PSA levels (data not shown). The difference in the genotype frequency of 11381G/C between cases with a Gleason score <8 (24.2%) and  $\geq 8$  (23.8%) was not statistically significant. The age-adjusted means for total PSA levels were 16.0, 99.1, and 83.7 among the cases with the CC, CG, and GG genotypes of this SNP, respectively ( $P = 0.18$ ).

Haplotype analysis did not provide additional support for an association between *TLR4* haplotypes and prostate cancer risk (data not shown). There were seven inferred haplotypes in this population based on these seven SNPs. The overall frequencies of these haplotypes were not significantly different ( $\chi^2 = 6.52$ ;  $df, 6$ ;  $P = 0.37$ ). The haplotype containing the C allele of SNP S61381 (haplotype A-T-A-A-A-G-C of SNPs -2026A/G, -1607T/C, 3622A/C, 3748A/C, D299G, 9615G/A, and 11381G/C) was present at a higher frequency in cases (12.43%) than in controls (10.35%;  $P = 0.04$ ).

Because there has been some controversy regarding the boundary of the 3'-UTR for the *TLR4* gene, we performed a 3' RACE experiment to further define the 3' end of *TLR4* mRNA expressed in prostate. Our 3' RACE results indicated that the 3'-UTR of *TLR4* is 1691 bp longer than the 1127-bp NCBI-annotated 3'-UTR, with a total length of 2818 bp. These results indicate that the last SNP in our association study, 11381G/C, is actually located in the middle of the *TLR4* 3'-UTR.

## DISCUSSION

There is growing evidence that chronic inflammation may play a role in the development of cancer within several organs, including the prostate. We are in the process of systematically evaluating possible associations between sequence variants of numerous inflammatory genes and prostate cancer risk. *TLR4* is a central player in the signaling pathways that control the innate immune response and was selected to be among the first group of genes to be evaluated. We hypothesized that variations in *TLR4* expression and function may be associated with individual susceptibility to cancer. To test this hypothesis, we performed a systematic genetic analysis of *TLR4* sequence variants in ~2100 men with or without prostate cancer in Sweden. We found a significant association between a sequence variant in the 3'-UTR of the *TLR4* gene. This variant conferred a 26%

<sup>8</sup> <http://www.stats.ox.ac.uk/mathgen/software.html>.

<sup>9</sup> <http://www.mayo.edu/statgen> for the S-PLUS programming language; <http://www.wfubmc.edu/genomics> for the R programming language.

Table 2 Genotype frequency of Toll-like receptor 4 (*TLR4*) single-nucleotide polymorphisms

SNP <sup>a</sup> (dbSNP)	Location	Amino acid	Genotype	Subjects, n (%)		P <sup>c</sup>
				Cases	Controls	
-2026A/G	Promoter		AA	625 (45.45)	341 (43.94)	0.8
			AG	596 (43.35)	354 (45.62)	
			GG	154 (11.20)	81 (10.44)	
-1607T/C	Promoter		TT	991 (72.07)	571 (73.39)	0.34
			TC	350 (25.45)	194 (24.94)	
			CC	34 (2.47)	13 (1.67)	
3622A/C	Intron		AA	1228 (90.56)	680 (88.66)	0.07
			AC	128 (9.44)	82 (10.69)	
			CC	0 (0.00)	5 (0.65)	
3748A/C	Intron		AA	1327 (96.02)	747 (96.02)	0.87
			AC	53 (3.84)	31 (3.98)	
			CC	2 (0.14)	0 (0.00)	
7449G/T (rs2149356)	Intron		GG	603 (51.89)	331 (52.62)	0.83
			TG	423 (36.40)	224 (35.61)	
			TT	136 (11.70)	74 (11.76)	
8552A/G (rs4986790)	Coding	D299G	AA	1241 (90.06)	693 (89.19)	0.31
			AG	136 (9.87)	79 (10.17)	
			GG	1 (0.07)	5 (0.64)	
9615G/A (rs5030721)	Coding		GG	1355 (98.12)	766 (98.21)	0.79
			AG	25 (1.81)	14 (1.79)	
			AA	1 (0.07)	0 (0.00)	
11381G/C	3'-UTR		GG	1047 (75.87)	625 (80.33)	0.05
			CG	318 (23.04)	141 (18.12)	
			CC	15 (1.09)	12 (1.54)	

<sup>a</sup> SNP, single-nucleotide polymorphism; UTR, untranslated region.

<sup>b</sup> From ATG of Isoform A (NM\_138554.1)

<sup>c</sup> Test for allele frequencies.

increased risk of prostate cancer in the study population overall and a 39% increased risk of being diagnosed with prostate cancer before age 65. The risk of prostate cancer in Sweden attributable to this variant was estimated to be 4.9%. Our study represents the first comprehensive evaluation of association between sequence variants of the *TLR4* gene and cancer susceptibility.

Although the observed ~1.3–1.4-fold increase in risk is modest, this is probably consistent with the magnitude of risk that we expect to observe for such a heterogeneous disease. Because genes in multiple pathways (such as androgen metabolism, growth factor, phase I and II detoxification, DNA repair, and inflammation) may alter the risk for prostate cancer, each individual gene is likely to contribute only a modest risk. This phenomenon is observed in other complex diseases, as reported in the most recent meta-analysis of genetic association studies in complex diseases (16). After evaluating 301 published studies that attempted to replicate reported disease associations for 25 different genes, the authors of that pooled meta-analysis confirmed the disease associations for 8 of those genes. Interestingly, seven of these eight genes were associated with modest estimated genetic effects (OR between 1.07 and 1.76) in the pooled analyses. They concluded that there are probably many common variants in the human genome with modest but real effects on common disease risk and that studies using large samples are needed to convincingly identify such variants.

When considering the likelihood that this finding represents a true association between the SNP 11381G/C and the disease, it is important to examine the possibility of spurious effects due to chance (multiple comparisons), as well as confounding due to population stratification. Although there were seven significance tests for the primary hypothesis, these tests were not independent because these SNPs are in strong LD. It is unclear how we might make an appropriate adjustment; however, the degree of inflated type I error should be minimal. Similarly, although the observed differences in allele and genotype frequencies could be due to differences in the genetic background of the two groups, rather than disease status (*i.e.*, population stratification), this is unlikely in this study. Genetic heterogeneity is less of a concern in Sweden than in the United States. In

addition, this was a carefully designed population-based study. Almost all of the patients that met the inclusion criterion enrolled as participants in this study. Control subjects were frequency matched to cases based on residence area and age. Furthermore, the higher frequency of the risk genotypes (CC or CG) among cases diagnosed before the age of 65 years (26.3%) compared with either cases diagnosed at age 65 years or older (22.9%) or controls (19.7%) provided additional support for this association. Finally, the large number of study subjects decreases the possibility for statistical fluctuation and significantly increases our confidence in interpreting the results.

Although SNP 11381G/C is outside the NCBI-annotated 3'-UTR of *TLR4*, previous studies have not clearly defined this 3'-UTR. In fact, several expressed sequence tag clones were then identified by our BLAST search against the human expressed sequence tag database, in which we used ~2 kb of genomic sequence beyond the NCBI-annotated 3'-UTR. This computational analysis result indicated that the 3'-UTR boundary could be further downstream of the NCBI-annotated 3' end of *TLR4*. Further support was provided by our

Table 3 Odds ratios for prostate cancer among Toll-like receptor 4 (*TLR4*) single-nucleotide polymorphisms

SNP <sup>a</sup> and position	Genotypes	Odd ratio <sup>b</sup> (95% CI)
-2026A/G	AA	1.00
	AG/GG	0.94 (0.78–1.12)
-1607T/C	CC	1.00
	TC/TT	1.08 (0.88–1.32)
3622A/C	AA	1.00
	AC/CC	0.81 (0.62–1.10)
3748A/C	AA	1.00
	AC/CC	0.96 (0.61–1.52)
7449C/A (rs2149356)	CC	1.00
	AC/AA	1.02 (0.83–1.24)
8552A/G (rs4986790), D299G	AA	1.00
	AG/GG	0.93 (0.70–1.25)
9615G/A (rs5030721)	GG	1.00
	AG/AA	1.01 (0.52–1.96)
11381G/C	GG	1.00
	CG/CC	1.26 (1.01–1.57)

<sup>a</sup> SNP, single-nucleotide polymorphism; CI, confidence interval.

<sup>b</sup> Adjusted for age and geographic regions.

experimental 3' RACE data, which showed that the 3'-UTR of *TLR4* is 1692 bp longer than the NCBI annotation. Although SNP 11381G/C is located 78 bp beyond the 1127-bp NCBI-annotated 3'-UTR, our computational and laboratory studies show that this SNP is actually located in the center of the 2818-bp *TLR4* 3'-UTR.

The observed association of SNP 11381G/C may be due to biological impact associated with this variant, or it may indirectly reflect another unobserved causal variant of *TLR4* that is in LD with this SNP. For example, SNP 11381G/C may itself influence the stability of the mRNA species. Alternatively, other SNPs in this region that are in strong LD with SNP 11381G/C may alter an AU-rich element motif and affect mRNA stability. There are at least 20 other known SNPs in this 2818-bp 3'-UTR. A large-scale evaluation of these SNPs and functional assessments are needed to address this question. On the other hand, it is unlikely that the observed association reflects the effects of other genes in this region, as the closest known gene [deleted in bladder cancer chromosomal region (*DBCCR1*)] is 2.5 Mb 3' from *TLR4*.

Several reported nonsynonymous changes were either not observed (E474K and Q510H) or observed only once (V310G) among our sample of 96 controls. We therefore could not assess their association with prostate cancer in our study population. Another nonsynonymous change, D299G, was relatively frequent in our population (~11% of control subjects had the variant genotypes) but was not associated with prostate cancer risk. The power calculation based on the prevalence of this SNP and the size of our study population suggested 80% power to detect an association at 5% significance level (two-sided test) if the SNP conferred at least a 1.45-fold increased risk. Multiple studies on the association between D299G and various phenotypes and diseases have been published, including the ability to recognize lipopolysaccharides, susceptibility to Gram-negative infections, premature birth, atherosclerosis, and asthma. The overall results were variable, with some reports providing positive findings (17–20) whereas others did not observe an association (21–23).

In summary, our study provided evidence of an association between a *TLR4* sequence variant and prostate cancer risk. More studies are needed to confirm or refute this finding in independent populations and to understand the mechanism by which *TLR4* sequence variants affect the expression and function of TLR4 in the signaling pathways that control innate immune response. Hopefully, our finding will also encourage additional research interest on the possible role of bacterial infection and inflammation in the development of prostate cancer.

## ACKNOWLEDGMENTS

We thank all study participants in the CAPS1 study. We thank Ulrika Lund for coordinating the study at Karolinska Institute as well as all of the urologists who recruited their patients to this study and provided clinical data to the National Prostate Cancer Registry. We also thank Karin Andersson, Susan

Okhravi-Lindh, Gabriella Thorén-Berglund, and Margareta Åswärd at the Regional Cancer Registries in Umeå, Uppsala, Stockholm-Gotland, and Lindköping. In addition, we thank Sören Holmgren and the personnel at the Medical Biobank in Umeå for skillfully handling the blood samples.

## REFERENCES

1. Coussens LM, Werb Z. Inflammation and cancer. *Nature (Lond)* 2002;420:860–7.
2. Gardner WA Jr, Bennett BD. The prostate—overview: recent insights and speculations *Monogr Pathol* 1992;34:129–48.
3. De Marzo AM, Marchi VL, Epstein JI, Nelson WG. Proliferative inflammatory atrophy of the prostate: implications for prostatic carcinogenesis. *Am J Pathol* 1999;155:1985–92.
4. Shah R, Mucci NR, Amin A, Macoska JA, Rubin MA. Postatrophic hyperplasia of the prostate gland. Neoplastic precursor of innocent bystander? *Am J Pathol* 2001;158:1767–73.
5. Carpten J, Nupponen N, Isaacs S, et al. Germline mutations in the ribonuclease L gene in families showing linkage with HPC1. *Nat Genet* 2002;30:181–4.
6. Xu J, Zheng SL, Komiya A, et al. Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with Prostate cancer risk. *Nat Genet* 2002;32:321–5.
7. Li L. Regulation of innate immunity signaling and its connection with human diseases. *Curr Drug Targets Inflamm Allergy*. In press 2004.
8. Poltorak A, He X, Smirnova I, et al. Defective LPS Signaling in C3H/HeJ and C57BL/10ScCr Mice: Mutations in *Tlr4* Gene. *Science (Wash DC)* 1998;282:2085–8.
9. Vogel SN, Perera PY, Detore GR, et al. CD14 dependent and independent signaling pathways in murine macrophages from normal and CD14 “knockout” (CD14KO) mice stimulated with LPS or taxol. *Prog Clin Biol Res* 1998;397:137–46.
10. Ohashi K, Burkart V, Flohe S, Kolb H. Cutting edge: heat shock protein 60 is a putative endogenous ligand of the toll-like receptor-4 complex. *J Immunol* 2000;164:558–61.
11. Arbour NC, Lorenz E, Schutte BC, et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat Genet* 2000;25:187–91.
12. Weir BS. Genetic data analysis II: methods for discrete population genetic data. Sunderland, MA: Sinauer Association, 1996.
13. Lillienfeld AM, Lillienfeld DE. Foundations of epidemiology, edition 2. New York: Oxford University Press, 1980.
14. Stephens M, Smith NJ, Donnelly PA. New statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001;68:978–89.
15. 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;70:425–34.
16. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;33:177–82.
17. Schmitt C, Humeny A, Becker CM, Brune K, Pahl A. Polymorphisms of TLR4: rapid genotyping and reduced response to lipopolysaccharide of TLR4 mutant alleles. *Clin Chem* 2002;48:1661–7.
18. Lorenz E, Hallman M, Marttila R, Haataja R, Schwartz DA. Association between the Asp299Gly polymorphisms in the Toll-like receptor 4 and premature births in the Finnish population. *Pediatr Res* 2002;52:373–6.
19. Kiechl S, Lorenz E, Reindl M, et al. Toll-like receptor 4 polymorphisms and atherogenesis. *N Engl J Med* 2002;347:185–92.
20. Agnese DM, Calvano JE, Hahn SJ, et al. Human toll-like receptor 4 mutations but not CD14 polymorphisms are associated with an increased risk of gram-negative infections. *J Infect Dis* 2002;186:1522–5.
21. Read RC, Pullin J, Gregory S, Bet al. A functional polymorphism of toll-like receptor 4 is not associated with likelihood or severity of meningococcal disease. *J Infect Dis* 2001;184:640–2.
22. Erridge C, Stewart J, Poxton IR. Monocytes heterozygous for the Asp299Gly and Thr399Ile mutations in the Toll-like receptor 4 gene show no deficit in lipopolysaccharide signalling. *J Exp Med* 2003;197:1787–91.
23. Raby BA, Klimecki WT, Laprise C, et al. Polymorphisms in toll-like receptor 4 are not associated with asthma or atopy-related phenotypes. *Am J Respir Crit Care Med* 2002;166:1449–56.

# Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

## Sequence Variants of *Toll-Like Receptor 4* Are Associated with Prostate Cancer Risk: Results from the CAnCER Prostate in Sweden Study

S. Lilly Zheng, Katarina Augustsson-Bälter, Baoli Chang, et al.

*Cancer Res* 2004;64:2918-2922.

**Updated version** Access the most recent version of this article at:  
<http://cancerres.aacrjournals.org/content/64/8/2918>

**Cited articles** This article cites 18 articles, 4 of which you can access for free at:  
<http://cancerres.aacrjournals.org/content/64/8/2918.full#ref-list-1>

**Citing articles** This article has been cited by 19 HighWire-hosted articles. Access the articles at:  
<http://cancerres.aacrjournals.org/content/64/8/2918.full#related-urls>

**E-mail alerts** [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, use this link  
<http://cancerres.aacrjournals.org/content/64/8/2918>.  
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.