ORIGINAL ARTICLE

GGN repeat length and GGN/CAG haplotype variations in the androgen receptor gene and prostate cancer risk in south Indian men

Krishnaswamy Vijayalakshmi · Kumarasamy Thangaraj · Singh Rajender · Venkatesan Vettriselvi · Perumal Venkatesan · Sunil Shroff · K. N. Vishwanathan · Solomon F. D. Paul

Abstract The ethnic variation in the GGN and CAG microsatellites of the androgen receptor (AR) gene suggests their role in the substantial racial difference in prostate cancer risk. Hence, we performed a case-control study to assess whether GGN repeats independently or in combination with CAG repeats were associated with prostate cancer risk in South Indian men. The repeat lengths of the *AR* gene determined by Gene scan analysis, revealed that men with GGN repeats ≤21 had no significant risk compared to those with >21 repeats (OR 0.91 at 95% CI-0.52–1.58). However, when CAG repeats of our earlier study was combined with the GGN repeat data, the cases exhibited significantly higher frequency of the haplotypes CAG ≤19/GGN ≤21 (OR-5.2 at 95% CI-2.17–12.48, $P < 0.001$) and CAG ≤19/GGN > 21(OR-6.9 at 95%CI-2.85–17.01, $P < 0.001$) compared to the controls. No significant association was observed between GGN repeats and prostate-specific antigen levels and the age at diagnosis. Although a trend of short GGN repeats length in high-grade was observed, it was not significant ($P = 0.09$). Overall, our data reveals that specific GGN/CAG haplotypes (CAG ≤19/GGN ≤21 and CAG ≤19/GGN > 21) of AR gene increase the risk of prostate cancer and thus could serve as susceptibility marker for prostate cancer in South Indian men.

Keywords Androgen receptor · GGN repeats · CAG repeats · Haplotypes · Prostate cancer

K. Vijayalakshmi · V. Vettriselvi · S. F. D. Paul (✉)
Department of Human Genetics, Sri Ramachandra Medical College and Research Institute (Deemed University), Porur, Chennai 600116 Tamil Nadu, India
e-mail: wise_soly@yahoo.com

K. Thangaraj · S. Rajender
Centre for Cellular and Molecular Biology, Hyderabad, India

S. Shroff
Department of Urology and Renal Transplantation, Sri Ramachandra Medical College and Research Institute (Deemed University), Chennai, India

P. Venkatesan
Tuberculosis Research Center (Indian Council of Medical Research), Chennai, India

K. N. Vishwanathan
Department Of General Medicine, Sri Ramachandra Medical College and Research Institute (Deemed University), Chennai, India

Introduction

The incidence of clinical prostate cancer differs substantially between ethnic groups, where African Americans exhibit a 10-to 40-fold higher incidence than Asians (Gronberg 2003). Despite an expanding body of epidemiological data, the etiology of prostate cancer remains poorly understood. However, evidences support the involvement of both genetic and environmental factors, which might also attribute to the ethnic differences in incidence rates.

The growth and development of the prostate gland, together with the maintenance of its physiological integrity, are dependent on the circulating androgens and intact intracellular steroid signaling pathways (Cunha et al. 1987). The effects of androgens are mediated through the androgen receptor (AR), a ligand-activated

nuclear transcription factor encoded by AR gene, located on the X chromosome (Xq11–12). The AR gene comprised of 8 exons, spanning more than 90 kb of the genomic DNA, encodes the AR protein with four functional domains including an amino-terminal transcription activation domain (TAD), the DNA binding domain (DBD), a hinge region and the carboxyl-terminal ligand-binding domain (LBD) (Janne et al.1993). Androgens, particularly dihydrotestosterone, bind to the AR with high affinity and stimulate the transcription of a cascade of androgen-responsive genes. In addition to stimulating the expression of genes associated with the differentiated phenotype of the prostate, such as prostate-specific antigen (PSA), it has been reported that AR may regulate genes involved in cell-cycle control, e.g., cyclin dependent kinases like CDK2, CDK4 and p16 (Lu et al. 1997). Thus, the AR transactivation plays an important role in the normal growth and function of the prostate gland. Exon 1 of the gene contains the polymorphic CAG and GGN repeat motifs, which are approximately 1.1 kb apart encoding polyglutamine and a polyglycine tract, respectively (Edwards et al. 1992). GGN repeat is complex in nature and composed of $(GGT)_3(GGG)_1(GGT)_2(GGC)_n$ (Platz et al. 1998). The number of CAG repeats ranges from 8 to 35 with an average of 20–23 repeats and the GGN tract varies from about 10–35 repeats. In vitro investigations suggest that CAG repeat length correlate inversely with AR transactivation (Chamberlain et al. 1994). The longer CAG repeat length found to be associated with male infertility (Patrizio et al. 2001), while the short CAG repeat length of the AR has been reported to predispose to prostate cancer (Coetzee et al. 1994; Mishra et al. 2005; Krishnaswamy et al. 2006). However, the effect of the variation in the length of the GGN tract on AR activity is unclear. Results from earlier studies on transient transfection of reporter constructs have shown that deletion of the GGN tract resulted in either no alteration or increased or decreased AR transcriptional activity (Jenster et al. 1994, Gao et al. 1996). Jenster et al. (1994) found that complete deletion of the (GGC) n sequence had no substantial effect on AR activity, whereas Gao et al. (1996) found that the same mutation resulted in a diminished capacity (30% reduction) to activate the luciferase gene.

Epidemiological investigations on the association between the number of GGN repeats and prostate cancer risk have produced inconsistent results. Because of the ethnic variation in CAG and GGN repeat lengths of the AR gene and the role of AR in androgenic activity, it has been suggested that the polymorphism may help explain part of the large ethnic difference in prostate cancer risk.

In India, a significant association has been identified between short CAG repeats and prostate cancer risk in North Indian men (Mishra et al. 2005), as well as our earlier study on the South Indian men (Krishnaswamy et al. 2006). This raised curiosity to determine the role of GGN repeat polymorphisms and its linkage with CAG repeats in prostate cancer risk in South Indian men. Therefore, we have analyzed the relation between AR-GGN microsatellite and prostate cancer risk and investigated whether this relation varies with tumor grade, PSA levels and age at diagnosis. Moreover, we assessed whether specific combination of CAG and GGN microsatellite alleles shows significant association with prostate cancer risk. In addition, we also tested for a possible linkage between the repeats among all the individuals studied, irrespective of the disease status.

Materials and methods

Subjects

The present case-control study comprised of 86 histologically confirmed prostate cancer patients and 119 male control subjects from the southern part of India. The controls comprised of 79 healthy, age-matched and unrelated individuals with normal serum PSA levels (≤4 ng/ml), digital rectal examination showing no abnormality and with no history of cancer and 40 subjects with benign prostatic hyperplasia (BPH). Both patients and control individuals were from the same ethnic background. Relevant clinical and pathological data were collected for all the patients. The age of prostate cancer patients ranged from 44 to 98 years with mean age of 67.5 years, in BPH patients it was 55–77 years with mean of 65.5 years and in normal healthy controls the age ranged between 50 and 81 years with mean of 66.5 years. Pathological grading of the tumors by Gleason scores (GS) were obtained and the patients were stratified as low grade if their Gleason scores were less than 7 and high grade if their Gleason scores were greater than or equal to 7. The Gleason score was less than 7 in 47 patients and greater than or equal to 7 in 39 patients. The study was approved by the Institutional Medical and Ethics Committee. Blood samples were collected from both the patients and controls with an informed written consent.

Genotyping of GGN repeat polymorphism

Genomic DNA was isolated from blood leucocytes by standard phenol/chloroform method (Sambrook et al.

1989). Exon 1 of *AR* gene was genotyped using the primers flanking the GGN repeat motif: 5′FAM-CCGCTTCCTCATCCTGGCACAC 3′ (forward primer) and 5′ GCCGCCAGGGTACCACACATC 3′ (reverse primer). Each PCR was carried out in a 10 µl reaction mixture containing 20 ng DNA, 1 µl of 10 × PCR buffer, 5 pM of each primer, 200 µM dNTPs (deoxynucleotide triphosphates), 0.4 µl of 100% DMSO, 0.6 µl of 100% Glycerol and 0.5 U of AmpliTaq Gold (Perkin–Elmer). PCR conditions consisted of initial denaturation of 96°C for 12 min, followed by 30 cycles each consisting of 1 min 30 s at 96°C, 1 min at 60°C and 3 min at 72°C followed by a final extension at 72°C for 5 min. For GeneScan analysis, 3.0 µl of the PCR product was mixed with 0.2 µl of LIZ500™ and 6.8 µl of formamide. Upon, denaturation for 5 min at 95°C and cooling for 5 min on ice, the samples were run on ABI 3730 DNA analyzer (Applied Biosystems, USA). The raw data were further analyzed using GeneMapper software to determine the number of repeats. The PCR and the genotyping were repeated for all the samples to confirm the number of repeats.

Genotyping of CAG repeat polymorphism was performed by PCR followed by GeneScan analysis on ABI3730 Genetic Analyzer as described earlier (Krishnaswamy et al. 2006)

Statistical analysis

The descriptive measures like mean, median and standard deviation of various characteristics such as age, PSA levels, tumor grade, CAG and GGN repeats of the subjects were calculated. Comparison of the mean GGN repeat length among the cases and controls were carried out using unpaired *t* test. The mean GGN repeat was used to categorize the subjects into two groups and the relative risk associated with GGN repeats was determined by calculating odds ratio (OR). The difference in proportion of specific CAG and GGN alleles between the cases and controls were evaluated by calculating OR. Using the χ^2 test, we assessed whether the distribution of GGN repeats varied by the level of CAG repeats and also determined the linkage disequilibrium between the repeats separately among the cases and controls. To further test whether GGN microsatellite contributes to the risk in combination with CAG microsatellite, logistic regression analysis was carried out with CAG and GGN as binary covariates.

Mean GGN repeat in different groups of Gleason score, age and PSA were compared by *t* test. GGN repeat distributions within each prognostic factor (age, grade and PSA values) were also calculated. One-way analysis of variance (ANOVA) was used to compare more than two mean values. All the tests were two-sided and the level of significance is taken as 5%. All the statistical analysis of data were performed using the statistical software SPSS (version 13).

Results

The mean age of prostate cancer patients, BPH and healthy controls were 67.5, 65.5 and 66.5 years, respectively. Mean serum PSA level measured at the time of diagnosis was 44.4 ng/ml in prostate cancer patients. The PSA values were in the normal range (≤4 ng/ml) in both the BPH patients and the healthy controls. Selected characteristics of prostate cancer cases and controls are presented in Table 1.

The number of GGN repeats among cases and healthy controls ranged from 15 to 26 with a mean of 21, whereas in BPH the number of GGN repeat was between 15–23 with mean of 21. Thus, no significant difference was seen in the mean GGN repeat among the cases and controls; also, within the control group the mean GGN repeats length was 21 in both BPH and healthy controls. The patient and control allelic

Table 1 Principle characteristics of study subjects

Characteristics	Prostate cancer patients (n = 86)	Control subjects (n = 119)	
		Healthy controls (n = 79)	BPH patients (n = 40)
Age (years)			
Mean ± SD	67.5 ± 8.8	66.5 ± 6.7	65.5 ± 6.6
Range	44.0–98.0	50.0–81.0	55.0–77.0
PSA (ng/ml)			
Mean ± SD	44.4 ± 31.5	–	–
Gleason score			
2	3.4%	–	–
3	8.0%	–	–
4	21.8%	–	–
5	12.6%	–	–
6	9.2%	–	–
7	24.1%	–	–
8	14.9%	–	–
9	4.6%	–	–
10	1.4%	–	–
CAG repeats			
Mean ± SD	17.0 ± 3.3	20.7 ± 3.9	21.3 ± 3.3
Median	17	21	21
Range	9–25	11–29	13–25
GGN repeats			
Mean ± SD	21.2 ± 1.8	21.03 ± 2.1	21.18 ± 1.6
Median	21.45	21.3	21.42
Range	15–26	15–26	15–23

distributions for the (GGN) n polymorphism are shown in Fig. 1.

For the polyglycine tract (GGT)$_3$(GGG)$_1$(GGT)$_2$(GGC) n, there was no variation in the number of GGT and GGG trinucleotides in all the samples analyzed, although the number of GGC repeats was highly variable. The pattern was always three GGT, one GGG, and two GGT, followed by a variable number of GGC repeats. Hence a 22 GGN repeat refers to 6 repeats corresponding to the consensus sequence and 16 GGC repeats. In the (GGN) n system, the 2 alleles, 21 and 22 repeats were predominant and together they accounted for 78% in patients and 75% in control populations of all 12 alleles genotyped.

In order to assess the risk associated with GGN repeats, the study subjects were dichotomized based on the mean GGN repeat. Men with GGN repeats ≤21 had no significant risk of prostate cancer compared to those with >21 repeats (OR 0.91 at 95% CI = 0.52–1.58) (Table 2).

The CAG repeat polymorphism analyzed in our previous study revealed a significant difference in the mean CAG repeats between prostate cancer patients and controls (17.0 vs. 20.7; $P < 0.001$) and, men with CAG repeat length ≤19 had a significantly increased risk for cancer than those with >19 CAG repeats (OR-5.90 at 95% CI 3.2–11.2; $P < 0.001$) (Krishnaswamy et al. 2006). In order to determine whether specific combination of CAG and GGN alleles differed significantly between cases and controls, we combined the CAG repeat data with the GGN repeats, observed in the present study. There was prevalence of CAG ≤19/GGN ≤21 and CAG ≤19/GGN >21 haplotypes in cases compared to controls. Thus, men with CAG ≤19/GGN
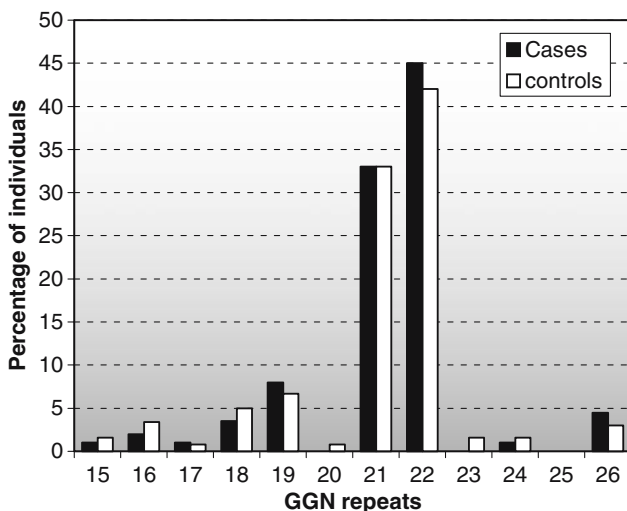
Table 2 Risk of prostate cancer in relation to the number of GGN repeats in exon 1 of AR gene

No of GGN repeats	Prostate cancer patients (n = 86) (%)	Controls (n = 119) (%)	OR	95% CI	P
>21	44 (51)	58 (49)	1.0		
≤21	42 (49)	61 (51)	0.91	0.52–1.58	NS

OR Odds ratio, *CI* confidence interval, *NS* non significant

≤21 (OR-5.2 at 95% CI-2.17–12.48, $P < 0.001$) and CAG ≤19/GGN >21(OR-6.9 at 95% CI–2.85–17.01, $P < 0.001$) had an increased risk compared to men with CAG >19/GGN >21 whereas individuals with CAG >19 and GGN≤21 were not at an increased risk for cancer (OR-1.1; 95% CI-0.14–2.83) (Table 3). Logistic regression analysis with CAG and GGN as binary covariates also showed a significant association ($P < 0.001$).

We further tested a possible association (linkage) between CAG and GGN microsatellites separately among the cases and controls by cross-classifying them into groups based on mean GGN (≤21 and >21) and mean CAG (≤19 and >19) repeat lengths (Table 4). However, we did not observe a significant linkage between the two microsatellites among the cases as well as among the controls.

We also analyzed GGN repeat polymorphism of the prostate cancer patients by categorizing them into different groups based on age, Gleason score and PSA levels (Table 5). With respect to tumor grade, patients with well and moderately differentiated tumor were classified as low grade (GS < 7) and those with poorly differentiated tumor as high grade (GS ≥ 7). Although a trend towards short mean GGN repeat length with high grade was observed, it was non-significant (P–0.09).

With respect to age at diagnosis, subjects were stratified into four groups based on quartiles (≤62, 63–66, 67–72, >72), and into two groups with regard to PSA levels, with mean PSA value as the cut-off. The mean GGN repeat within each of the age groups and PSA groups revealed no significant difference. Moreover, stratified analysis of GGN repeats distribution based on the age of onset, tumor grade and PSA levels revealed no significant association with any of the variables (Table 5).

Discussion

Studies on the association of AR-GGN repeat length and prostate cancer risk have produced conflicting

Fig. 1 Distribution of GGN repeats in Androgen receptor gene among prostate cancer patients and controls

Table 3 Risk of prostate cancer in relation to the combined distribution of number of CAG and GGN repeats in exon 1 of AR gene

Repeats	Patients (%)	Controls (%)	OR	95% CI	P
CAG > 19/GGN > 21	10 (11.6)	39 (32.7)	1.0		
CAG ≤ 19/GGN ≤ 21	32 (37)	24 (20)	5.2	2.17–12.48	<0.001
CAG ≤ 19/GGN > 21	34 (39.5)	19 (15)	6.9	2.85–17.01	<0.001
CAG > 19/GGN ≤ 21	10 (11.6)	37 (31)	1.1	0.14–2.83	NS

OR Odds ratio, CI confidence interval, NS non significant

Table 4 Frequency distribution of AR gene CAG and GGN repeat lengths among prostate cancer cases and controls

Group	No. of GGN repeats	No. of CAG repeats		Total (%)
		≤19 (%)	>19 (%)	
Cases[a] (n = 86)	≤21	32 (48.5)	10 (50)	42 (48.8)
	>21	34 (51.5)	10 (50)	44 (51.2)
Total		66 (100)	20 (100)	86 (100)
Controls[b] (n = 119)	≤21	24 (55.8)	37 (48.7)	61 (51.3)
	>21	19 (44.2)	39 (51.3)	58 (48.7)
Total		43 (100)	76 (100)	119 (100)

[a] χ^2 test of independence: two sided P = 0.905

[b] χ^2 test of independence: two sided P = 0.455

Table 5 Comparison of AR-GGN repeats of prostate cancer patients as a variable with age, grade and serum PSA levels at diagnosis

Variable	No. of cases (n)	GGN repeats Mean ± SD	P value	Distribution of GGN repeats within each prognostic factor		P value
				No. of cases with ≤21 GGN repeats (%)	No. of cases with >21 GGN repeats (%)	
Grade						
GS < 7 (Low grade)	47	21.53 ± 1.63	0.09	21 (44.7)	26 (55.3)	NS
GS ≥ 7 (High grade)	39	20.85 ± 2.11		21 (53.8)	18 (46.2)	
Age						
≤62	22	21.14 ± 1.86	NS	14 (64)	8 (36)	NS
63–66	24	21.17 ± 2.09		11 (46)	13 (54)	
67–72	24	21.04 ± 1.37		11 (46	13 (54)	
>72	16	21.69 ± 2.30		6 (37)	10 (63)	
PSA level						
≤48	53	21.15 ± 2.30	0.665	26 (49.1)	27 (50.9)	NS
>48	33	21.33 ± 0.89		16 (48.5)	17 (51.5)	

GS Gleason score, PSA prostate specific antigen, NS non significant, SD standard deviation

results. Our study reveals that singly the GGN repeats are not associated with prostate cancer risk, but when combined with CAG repeats, men with CAG ≤19/GGN ≤21 and CAG ≤19/GGN >21 have an increased risk compared to men with CAG >19/GGN >21. However, we did not observe any statistically significant association between GGN repeats length and age of diagnosis, Gleason score and PSA levels.

The distribution of GGN microsatellites has been reported to differ significantly among different ethnic groups. High-risk African–Americans were found to possess the lowest frequency (20%) for GGN allele, with 22 repeats; whereas the comparable values for intermediate-risk whites and low-risk Asians were 57 and 70%, respectively (Irvine et al. 1995). Among the Western (Platz et al. 1998, Chen et al. 2002) as well as Chinese men (Hsing et al. 2000) GGN repeat length of 23 was predominant. Hence, they suggested that 23 GGN repeats might represent the coding sequence for optimal AR protein conformation and activity. However, in our study only 1.7% of the subjects had 23 GGN repeats and the repeats clustered around 22; where, 42% of the subjects had 22 repeats and 33% had 21 repeats, thus revealing the polymorphic nature

and distinct ethnic variation in GGN repeat tract length.

Similar to our results, no significant genotype-specific prostate cancer risk was found with GGN repeat polymorphism among Caucasians in Britain (Edwards et al. 1999), French–German men (Correa-Cerro et al. 1999) and Caucasians in America (Chen et al. 2002). Moreover, men with short GGN repeats were not at increased risk in familial prostate cancer cases (Miller et al. 2001; Cicek et al. 2004). In addition, a recent study on early onset prostate cancer in British men also reported lack of association between GGN repeats and prostate cancer risk (Forrest et al. 2005).

In contrast to our results, men with GGC repeats ≤16 have been reported to exhibit higher risk estimates than men with >16 repeats (Stanford et al. 1997). Moreover, a significantly increased frequency of the GGC repeat ≤16 has been reported in hereditary as well as sporadic prostate cancer in a study predominant of Caucasians (Chang et al. 2002). In addition, among Chinese men those with <23 GGN repeats had 12% increased risk of prostate cancer compared to those with ≥23 GGN repeats (Hsing et al. 2000). Thus, our results on GGN repeats correlate with and deviate from some of the previous studies revealing the ethnic differences in AR GGN polymorphism and association with prostate cancer (Table 6).

When we combined both the CAG and GGN repeats, we observed men with the haplotypes CAG ≤ 19/GGN ≤ 21 and CAG ≤ 19/GGN > 21 to exhibit increased risk compared to men with CAG > 19/GGN > 21 haplotype. Table 7 reveals the combined distribution of the repeats observed in different studies. Our results are thus consistent with earlier studies, where the subgroups with two short repeats (CAG < 22; GGC ≤ 16) had a twofold increased risk relative to those with long repeats (CAG ≥ 22; GGC > 16) (Stanford et al. 1997). Platz

et al. (1998) reported an increased risk for those with a GGN = 23 and CAG < 21 compared to a GGN other than 23 and a CAG > 23. In contrast, men with CAG < 22 and GGN ≤ 23 repeats were not at increased risk of prostate cancer (Chen et al. 2002).

Since the AR gene is located on the X-chromosome, the two microsatellites, which are in close proximity at this locus, can be associated with each other or in other words one would expect to find linkage disequilibrium between them. Irvine et al. (1995) have observed significant linkage disequilibrium between the CAG and GGN repeats only among cases and not among controls. However, Platz et al. (1998) reported linkage disequilibrium among cases and controls. In contrast, our results revealed no significant linkage between the repeats among cases as well as controls. The absence of linkage between the repeats might indicate that either one or both the repeats mutate at a relatively high rate and independent of each other.

Studies on the association of GGN repeats and the age of diagnosis have revealed contrasting findings. Stanford et al. (1997) observed men with GGC ≤ 16 to be at increased risk regardless of the age at diagnosis. But Miller et al. (2001) reported a reduced risk among men diagnosed at the age ≤ 66 years and an increased risk among men diagnosed at the age ≥ 66 years. Moreover, a reduced risk was observed with ≤17 repeats in men aged 70 years or older and no evidence of any association in men <70 years (Chen et al. 2002). However, our study did not find any association between age and the prostate cancer risk.

Although a trend of short mean GGN repeat length with high grade was observed in our study the association was not significant. Hakimi et al. (1997) have reported short GGN repeats to identify a sub-population of patients with clinically localized disease. However, Edwards et al. (1999) have observed long GGN alleles at higher frequency in advanced stages and

Table 6 Studies on AR-GGN repeat polymorphism and prostate cancer risk

Study	Study population	Results	
		GGN repeats	OR (%CI)
Present study*	Asians/ South Indians	≤21 vs. >21	0.91 (0.52–1.58)
Irvine et al. 1995	Non Hispanic whites	Others vs. 16	1.18 (data not available)
Stanford et al. 1997	Caucasians/United States	≤16 vs. >16	1.60 (1.07–2.41)
Platz et al. 1998	Physicians' Health Study/USA	23 vs. not 23	1.2 (0.97–1.49)
Edwards et al. 1999	Caucasians/United Kingdom	≤16 vs. >16	1.06 (0.57–1.96)
Hsing et al. 2000	Asian/China	<23 vs. ≥23	1.12 (0.71–1.78)
Miller et al. 2001	Caucasians	≤16 vs. >16	0.98 (0.46–2.06)
Chu chen et al. 2002	Caucasians/America	≤17 vs. >17	0.8 (0.57–1.12)
Chang et al. 2002	Caucasians/United States	≤16 vs. ≥17	1.51 (0.99–2.32)
Cicek et al. 2004	Caucasians and African Americans/United States	≤16 vs. >16	1.04 (0.71–1.52)
Forrest et al. 2005	British men/United Kingdom	≤16 vs. >16	1.06 (0.78–1.44)

Table 7 Studies on combined distribution of AR-CAG and GGN repeat polymorphism in prostate cancer

Study	Population	Results	
		CAG and GGN/GGC	OR (%CI)
Present study	Asians/ South Indians	≤19, ≤21 vs. >19, >21	5.2 (2.17–12.48)
		≤19, >21 vs. >19, >21	6.9 (2.85–17.01)
		>19, ≤21 vs. >19, >21	1.1 (0.139–2.83)
Irvine et al. 1995	Non Hispanic whites	<22, not 16 vs. ≥22, 16	2.1 (NA)
Stanford et al. 1997	Caucasians/ United States	≥22, ≤16 vs. ≥22, >16	1.15 (0.56–2.35)
		<22, >16 vs. ≥22, >16	1.54 (0.83–2.86)
		<22, ≤16 vs. ≥22, >16	2.05 (1.09–3.84)
Platz et al. 1998	Physicians' Health Study /USA	>23,23 vs. >23 not 23	1.17 (0.77–1.77)
		21–23, not 23 vs.>23 not 23	1.39 (0.93–2.06)
		21–23,23 vs. >23 not 23	1.22 (0.82–1.83)
		<21, not 23 vs.>23 not 23	1.49 (1.02–2.15)
		<21,23 vs. >23 not 23	1.62 (1.07–2.44)
Hsing et al. 2000	Asian/China	≥23, <23 vs. ≥23, ≥23	1.48 (0.76–2.88)
		<23, ≥23 vs. ≥ 23, ≥23	1.85 (1.21–2.82)
		<23, <23 vs. ≥23, ≥23	1.75 (0.9–3.41)
Miller et al. 2001	Caucasians	≥22, ≤16 vs. ≥22, >16	0.63 (0.18–2.2)
		<22, >16 vs. ≥22, >16	0.69 (0.17–2.74)
		<22, ≤16 vs. ≥22, >16	1.06 (0.25–4.46)
Chu chen et al. 2002	Caucasians/ America	≥22, ≤17 vs. ≥22, >17	0.54 (0.32–0.91)
		<22, >17 vs. ≥22, >17	0.55 (0.31–0.98)
		<22, ≤17 vs. ≥22, >17	0.56 (0.32–0.98)
Chang et al. 2002	Caucasians/ United States	≥22, ≤16 vs. ≥22, ≥17	1.27 (0.68–2.39
		≤21, ≥17vs. ≥22, ≥17	0.65 (0.32–1.30)
		≤21, ≤16 vs. ≥22, ≥17	1.14 (0.62–2.12)

grades. They also reported long GGN alleles to be associated with shorter time to relapse and worse overall survival. Thus we propose that assessing the role of CAG and GGN repeats with relapse and overall survival of the patients will enable prediction of the growth behavior of early-stage tumors and thus validate the prognostic significance of the repeats in prostate cancer.

Since the AR transactivation results in PSA secretion, we assessed the association of PSA levels with AR GGN genotype but did not observe any significant association. As the GGN repeats are in the transactivation domain, it is possible that a single–amino acid difference disrupts the binding affinity of the domain enough to up- or down-regulate the formation of a critical regulatory complex. However, the two functional studies reported so far are contradictory, with one revealing no substantial effect of the deletion of GGN repeats on AR activity (Jenster et al. 1994), whereas another reported diminished AR activity (Gao et al.1996). Thus the functional significance of the GGN repeats needs to be further evaluated.

The differences in the results reported for each study population may be partially explained by the gene–environment interaction. Differences in the study design and reference CAG and GGN lengths may also contribute to the divergent results in the epidemiological studies. It has also been proposed that the polymorphic CAG and GGN repeats function as low penetrance prostate cancer alleles that may require additional genetic or environmental factors to result in increased cancer risk (Nwosu et al. 2001). Moreover, they might be in linkage disequilibrium with other disease causing mutations in the AR gene or with other unknown adjacent genes that affect prostate cancer risk. Although numbers of repeats have a quantitative feature, frequency of 21 and 22 GGN repeats is very high and others are almost negligible. Therefore the positive association could be due to difference between the allele with 21 repeats and the allele with 22 repeats, rather than due to quantitative effect of number of repeats.

To the best of our knowledge, this is the first study to investigate the association between GGN repeat polymorphism and the relative risk of prostate cancer in Indian men. Our results suggest that specific haplotypes of AR attribute to risk of prostate cancer. Because of the significance of AR in prostate cancer, investigation of factors that interact with the polyglutamine and polyglycine region of the AR gene to alter AR function and modulate prostate cancer risk is an important area for future research.

References

Chamberlain N, Driver E, Miesfeld R (1994) The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acid Res 22:3181–3186

Chang BL, Zheng SL, Hawkins GA, Isaacs SD, Wiley KE, Turner A, Carpten JD Bleecker ER, Walsh PC, Trent JM, Meyers DA, Isaacs WB, Xu J (2002) Polymorphic GGC repeats in the androgen receptor gene are associated with hereditary and sporadic prostate cancer risk. Hum Genet 110:122–129

Chen C, Lamharzi N, Weiss NS, Etzioni R, Dightman DA, Barnett M, Di Tommaso D, Goodman G (2002) Androgen receptor polymorphisms and the incidence of prostate cancer. Cancer Epidemiol Biomarkers Prev 11:1033–1040

Cicek MS, Conti DV, Curran A, Phillippa JN, Paris PL, Graham C, John SW (2004) Association of prostate cancer risk and aggressiveness to androgen pathway genes: SRD5A2, CYP17, and the AR. Prostate 59:69–76

Coetzee GA, Ross RK (1994) Re: prostate cancer and the androgen receptor. J Nat Cancer Inst 86:872–873

Correa-Cerro L, Wohr G, Haussler J, Berthon P, Drelon E, Mangin P, Fournier G, Cussenot O, Kraus P, Just W, Paiss T, Cantu JM, Vogel W (1999) (CAG) n and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French–German population. Eur J Hum Genet 7:357–362

Cunha GR, Donjacour AA, Cooke PS, Mee S, Bigsby RM, Higgins SJ, Sugimura Y (1987) The endocrinology and developmental biology of the prostate. Endocr Rev 8:338–362

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241–253

Edwards SM, Badzioch MD, Minter R, Hamoudi R, Collins N, Ardern-Jones A, Dowe A, Osborne S, Kelly J, Shearer R, Easton DF, Saunders GF, Dearnaley DP, Eeles RA (1999) Androgen receptor polymorphisms: association with prostate cancer risk, relapse and overall survival. Int J Cancer 84:458–465

Forrest MS, Edwards SM, Houlston R, Kote Jarai Z, Key T, Allen N, Knowles MA, Turner F, Ardern Jones A, Murkin A, Williams S, Oram R, CR-UK/BPG UK prostate cancer study collaborators et al (2005) Association between hormonal genetic polymorphisms and early onset prostate cancer. Prostate Cancer Prostatic Dis 8:95–102

Gao T, Marcelli M, McPhaul MJ (1996) Transcriptional activation and transient expression of the human androgen receptor. J Steroid Biochem Mol Biol 59:9–20

Gronberg H (2003) Prostate cancer epidemiology. Lancet 361:859–864

Hakimi JM, Schoenberg MP, Rondinelli RH, Piantadosi S, Barrack ER (1997). Androgen receptor variants with short

glutamine or glycine repeats may identify unique sub populations of men with prostate cancer. Clin Cancer Res 3:1599–1608

Hsing AW, Gao YT, Wu G, Wang X, Deng J, Chen YL, Sesterhenn IA, Mostofi FK, Benichou J, Chang C (2000) Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case-control study in China. Cancer Res 60:5111–5116

Irvine RA, Yu MC, Ross RK, Coetzee GA (1995) The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. Cancer Res 55:1937–1940

Janne OA, Palvimo JJ, Kallio P, Mehto M (1993) Androgen receptor and mechanism of androgen action. Ann Med 25:83–89

Jenster G, Ruiter PE, Van der Korput HA, Kuiper GG, Trapman J, Brinkmann AO (1994) Changes in the abundance of androgen receptor isotypes: effects of ligand treatment, glutamine-stretch variation, and mutation of putative phosphorylation sites. Biochemistry 33:14064–14072

Krishnaswamy V, Kumarasamy T, Venkatesan V, Shroff S, Jayanth VR, Paul SFD (2006) South Indian men with reduced CAG repeat length in the androgen receptor gene have an increased risk of prostate cancer. J Hum Genet 51:254–257

Lu S, Tsai SY, Tsai MJ (1997) Regulation of androgen-dependent prostatic cancer cell growth: androgen regulation of CDK2, CDK4, and CKI P16 genes. Cancer Res 57:4511–4516

Miller EA, Stanford JL, Hsu L, Noonan E, Ostrander EA (2001) Polymorphic repeats in the androgen receptor gene in high-risk sibships. Prostate 48:200–205

Mishra DK, Thangaraj K, Mandhani A, Kumar A, Mittal RD (2005) Is reduced CAG repeat length in androgen receptor gene associated with risk of prostate cancer in Indian population? Clin Genet 68:55–60

Nwosu V, Carpten J, Trent JM, Sheriden R (2001) Heterogeneity of genetic alterations in prostate cancer: evidence of the complex nature of the disease. Hum Mol Genet 10:2313–2318

Patrizio P, Leonard DGB, Chen K-L, Hernandez-Ayup S, Trounson AO (2001) Larger trinucleotide repeat size in the androgen receptor gene of infertile men with extremely severe oligozoospermia. J Androl 22:444–448

Platz EA, Giovannucci E, Dahl DM, Krithivas K, Hennekens CH, Brown M, Stampfer MJ, Kantoff PW (1998) The androgen receptor gene GGN microsatellite and prostate cancer risk. Cancer Epidemiol Biomark Prev 7:379–384

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd edn. Cold Spring Harbor Laboratory Press, New York 9.17–9.19

Stanford JL, Just JJ, Gibbs M, Wicklund KG, Neal CL, Blumenstein BA, Ostrander EA (1997) Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. Cancer Res 57:1194–1198