

ORIGINAL ARTICLE

Polymorphisms in the androgen receptor CAG repeat sequence are related to tumour stage but not to ERG or androgen receptor expression in Malaysian men with prostate cancer

Jia Shin Jessica TAN¹, Kien Chai ONG², Diana Bee-Lan ONG¹, Yuan Seng WU⁴, Azad RAZACK³, Shanggar KUPPUSAMY³, Jasmine LIM³, and Anthony RHODES^{1,4}

¹Departments of Pathology, ²Biomedical Science and Surgery ³Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; ⁴School of Medicine, Faculty of Health and Life Sciences, Taylor's University, Selangor, Malaysia.

Abstract

Introduction: Polymorphic expression of a CAG repeat sequence in the androgen receptor (AR) gene may influence the activity of the AR and the occurrence of prostate cancer and the TMPRSS2-ERG fusion event. Furthermore, this polymorphism may be responsible for the ethnic variation observed in prostate cancer occurrence and expression of the ERG oncogene. We investigate the expression of AR and ERG in the biopsies of Malaysian men with prostate cancer and in the same patients relate this to the length of the CAG repeat sequence in their AR gene. **Materials and Methods:** From a PSA screening initiative, 161 men were shown to have elevated PSA levels in their blood and underwent prostatic tissue biopsy. DNA was extracted from the blood, and exon 1 of the AR gene amplified by PCR and sequenced. The number of CAG repeat sequences were counted and compared to the immunohistochemical expression of ERG and AR in the matched tumour biopsies. **Results:** Of men with elevated PSA, 89 were diagnosed with prostate cancer, and 72 with benign prostatic hyperplasia (BPH). There was no significant difference in the length of the CAG repeat in men with prostate cancer and BPH. The CAG repeat length was not associated with; age, PSA or tumour grade, though a longer CAG repeat was associated with tumour stage. ERG and AR were expressed in 36% and 86% of the cancers, respectively. There was no significant association between CAG repeat length and ERG or AR expression. However, there was a significant inverse relationship between ERG and AR expression. In addition, a significantly great proportion of Indian men had ERG positive tumours, compared to men of Malay or Chinese descent. **Conclusions:** CAG repeat length is not associated with prostate cancer or expression of ERG or AR. However, ERG appears to be more common in the prostate cancers of Malaysian Indian men than in the prostate cancers of other Malaysian ethnicities and its expression in this study was inversely related to AR expression.

Keywords: Prostate cancer, CAG repeat, ERG, ethnicity

INTRODUCTION

Prostate cancer is one of the most common cancers occurring in men. However, a large proportion of screen-detected prostate cancers are relatively indolent and unlikely to result in mortality.¹ One of the main aims of cancer research is to identify a biomarker capable of discriminating between indolent cancers which require no further treatment and aggressive cancers which require surgical or radiological intervention.

In this respect, a common oncogenic event occurring in up to sixty-percent of prostate cancers in western cohorts is the fusion of the androgen-regulated promoter sequence of the Transmembrane Serine Protease 2 (TMPRSS2) gene to the ETS transcription factor family member ETS-Related Gene (ERG).^{2,3} Androgen binding to the promoter region of the TMPRSS2-ERG fusion then results in increased ERG expression and is considered to be oncogenic as ERG acts as a transcription factor for the

downstream regulation of a number of genes, known to be important in tumour proliferation and invasion.⁴ The TMPRSS2 and ERG genes are located less than 3 Mb apart on chromosome 21² and experimental models have demonstrated that the fusion of these two genes is a result of androgen signalling and gamma irradiation.⁵ Consequently, one of the proposed mechanisms for the formation of the TMPRSS2-ERG gene fusion is that high levels of androgen activity induces TMPRSS2 and ERG gene proximity, and therefore increases the probability of the fusion event occurring.^{3,5,6} Consequently, it is hypothesised that increased AR signalling promotes the formation of TMPRSS2-ERG fusion gene and hence the expression of ERG oncoprotein.³ However, the underlying cause of this increased AR activity, responsible for the fusion, remains unanswered.

Androgen, with its receptor, influences the growth, development and differentiation of the prostate gland.⁷ The binding of androgens to AR along with the presence of transcriptional coactivators leads to the transcription of target genes.⁷ The AR gene is located on the X chromosome (Xq11-q12) with 4 domains: a transactivating amino-terminal domain, a DNA binding domain, a hinge region and a carboxyl-terminal ligand (steroid) binding domain.⁸ Exon 1 of the AR gene encodes for the large N-terminal transactivating domain, exons 2 and 3 encode the DNA binding domain while the remaining exons code for the ligand-binding domain.⁸ In exon 1, there is a trinucleotide microsatellite of cytosine, adenine and guanine (CAG) which is polymorphic in humans. The CAG codon codes for the amino acid glutamine in the N-terminal transactivation domain of the AR protein.⁹ In the healthy population, the CAG repeat lengths usually range from 8-35 repeats and it has been shown that there is an inverse linear relation between CAG repeat length and AR transactivation function.¹⁰ Expansion of CAG trinucleotide repeats of above the normal length (>40), results in an extended poly-glutamine tract and is associated with human genetic disease, such as X-linked spinal and bulbar muscular atrophy (Kennedy's disease); a rare neuromuscular disease where patients also experience androgen insensitivity, decreased virilisation, testicular atrophy, reduced sperm count and infertility.⁸ In functional studies, the progressive expansion of the poly-glutamine tract in the human AR is associated with a linear decrease in transactivation function.⁹

Whilst shorter CAG repeat lengths (and hence a shorter poly-glutamine tract) is hypothesised to be responsible for greater sensitivity towards androgen stimulation and pose an advantage for malignant prostate growth.¹¹ Studies have demonstrated an association of shorter CAG repeat lengths with increased prostate cancer risk¹², earlier age of cancer onset in Caucasian men^{13,14}, and aggressive early-stage prostate cancer.¹²

Intriguingly, the CAG repeat length and prostate cancer incidences have been reported to differ between ethnicities¹⁵ with the average CAG repeat length to be shortest in African Americans, followed by Caucasians, Chinese and longest in Hispanic men.¹⁵ These CAG repeat lengths corresponded to high, intermediate and low incidence and mortality rate of prostate cancer in these populations.¹⁵ However, other studies investigating the association of CAG repeats with prostate cancer risk have been inconclusive.^{16,17}

Despite being detected in more than half of the prostate cancer patients in the Western population, Asian samples have demonstrated much lower frequency of TMPRSS2-ERG fusions and ERG expression.^{18,19} We hypothesise that the difference in CAG repeat length and its effect on AR activity may explain the difference seen in the occurrence of the fusion event and ERG expression in men with prostate cancer of different ethnicities. In support of this, a recent study reported that CAG repeats to be lower in TMPRSS2:ERG positive tumours compared with TMPRSS2:ERG negative prostate cancer.²⁰ In the current study, we investigate whether the occurrence of ERG expression in the prostate cancers of Malaysian men is related to the length of the CAG repeat sequence in the AR gene. If men with a shorter CAG repeat sequence in their DNA are more likely to have ERG positive tumours, it would support the findings of previous experimental *in vitro* studies, suggesting that the activity of the AR plays a major role in the occurrence of the TMPRSS2-ERG fusion gene.

MATERIALS AND METHODS

Prostatic biopsies with matched blood samples

From a prospective study involving the collection of blood from a PSA screening initiative conducted by urologists from the Department of Surgery, University of Malaya, 161 men were shown to have elevated PSA levels and subsequently underwent prostatic tissue biopsy.

Clinico-pathological data were collected for each patient from the University of Malaya Medical Centre online database or the patient's medical record folders (ethical approval #NMRR-10-1400-7968, Malaysian Ministry of Health). Blood samples were collected in BD Vacutainer Serum Separator Tubes. Approximately 30 minutes after blood collection, serum was separated by centrifugation at 1000xg for 15 minutes and aliquoted into cryovials for storage at -80°C before analysis. All the tissue biopsies were fixed in neutral buffered formalin for 24 hours.

DNA extraction

For genomic DNA extraction, the Maxwell® 16 Blood Purification kit (Promega, Madison, USA) was used as specified in the manufacturer's instructions. DNA yields and purity were determined using a Quantus™ Fluorometer (Promega, USA).

Polymerase Chain Reaction (PCR)

The PCR reaction was performed using a Veriti™ Thermal Cycler (Applied Biosystems, CA, USA), utilising a forward primer; (5'-TCCCGCAAGTTTCCTTCTCT-3') and a reverse primer (5'-CCCACTTTCCTGGCTTAA-3') spanning exon 1 of the androgen receptor (AR) gene.

Gel electrophoresis and DNA sequencing

The DNA products amplified were separated using gel electrophoresis. The separated DNA products were then visualised using a Gel Doc XR+ documentation system (Bio-Rad). Targeted DNA product was verified based on molecular size. The targeted DNA products of the AR gene were then cut out carefully before sending for purification and sequencing. The forward and reverse primers used for PCR were also used for DNA sequencing. Chromas Lite software was then used to count the number of CAG triplets in the polyglutamine tract (CAG nCAA) for each sample.

Immunohistochemistry

Sections were cut at 4 micrometers and mounted onto Superfrost Plus Slides (Thermo Scientific, USA). The sections were de-paraffinised in xylene and rehydrated in a series of graded alcohols before antigen retrieval in Tris EDTA (pH 9) for 30 mins at 100°C. Incubation with primary ERG antibody (Clone EP111, Dako, Denmark) was carried out for one hour at room temperature, whilst staining for AR

Clone AR441 (Dako, Denmark), diluted in TBS-Tween 20, was carried out overnight at 4°C. The primary antibody was detected using DAKO REAL EnVision Detection System utilising a horseradish peroxidase label (Dako, Denmark). The reaction was visualised using 3,3-diaminobenzidine tetrahydrochloride (DAB) chromogen. Nuclei were counterstained with Harris's haematoxylin (Leica, Germany). Experimental runs contained negative controls in which the primary antibody was omitted. Staining of endothelial cells of small vessels functioned as the internal positive control for ERG; whilst cases of benign prostatic hyperplasia were used as positive controls for AR. Both ERG and AR displayed the expected nuclear immune-reactivity. ERG was recorded as positive when either part or the whole of the tumour compartments showed positive staining. AR staining was scored using a semi-quantitative method²¹. With this system the percentage and intensity of staining is recorded as follows: Percentage: score 0 (negative), 1 (less than 20%), 2 (20-70% positivity) and 3 (more than 70% positivity); Intensity 0 (negative), 1 (weak), 2 (moderate) and 3 (strong). Based on the product of the two scores, AR was then categorised as either negative (product score of 0 or 1) or positive (product score >1). The histopathological evaluation was performed by an experienced clinical pathologist (DBLO).

Statistical analysis

Tests for normality of linear variables (AR, CAG repeat number, PSA, and age at diagnosis) were performed by the Kolmogorov–Smirnov statistic with a Lilliefors significance level. Tests for an association between linear variables were performed by Pearson correlation coefficient or Spearman rank correlation (when data did not meet the assumption of normality). Student's T-test and logistic regression were performed on CAG repeat length distribution to investigate its association with prostate cancer, BPH and between ethnic groups. The median CAG repeat length was used when comparing CAG repeats to ERG and AR expression and clinical and pathological parameters. Comparison of AR and ERG immunostaining and with clinical and pathological data were analysed using the chi-square test. The statistical tests were performed using SPSS software (SPSS Inc., Chicago, IL) with $p < 0.05$ considered statistically significant.

RESULTS

Of the 161 men with elevated PSA levels, 89 were identified as having invasive prostate cancer whilst the remaining 72 men had evidence of benign prostatic hyperplasia (BPH). The prostate cancer cases comprised of 27 Malay, 31 Chinese and 31 Indians patients. Just over 40% of the patients had Gleason’s scores of 8 and above. With respect to the stage; nearly half of the tumours (47%) were of late stage. The control group comprised 72 cases of BPH of which 28 were of Malay, 29 Chinese and 15 Indian ethnicities. The allele distributions for the CAG polymorphism for prostate cancer and BPH cases are shown in Figure 1. The most

frequent allele was the 22 CAG repeat length for both the prostate cancer and BPH groups. There was no significant difference between the mean CAG repeat length for the prostate cancer group (mean 22.8, SD 2.4, range 17-29) and the BPH group (mean 22.1, SD 3.6, range 13-30), unpaired t-test ($p=0.21$). To compare our results with those already published; we dichotomised our cases and controls based on the median repeat length of 22 in the prostate cancer group (≤ 22 and > 22). The proportion of cases above and below this cut point was similar for the BPH and prostate cancer cases (OR of 0.78, 95% CI 0.398-1.532 ($p=0.5$)) (Figure 2, Table 2).

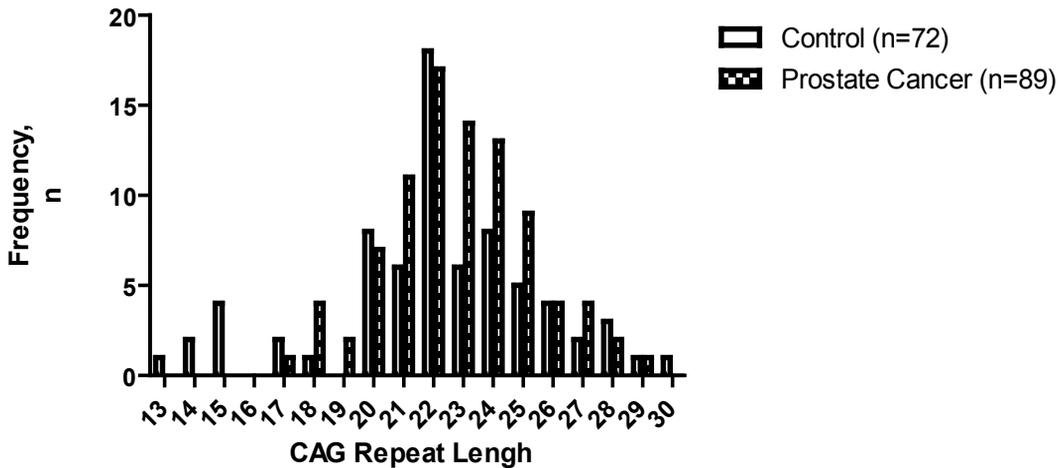


FIG. 1: Distributions of CAG polymorphic repeats in patients with prostate cancer and benign prostatic hyperplasia (Control group)

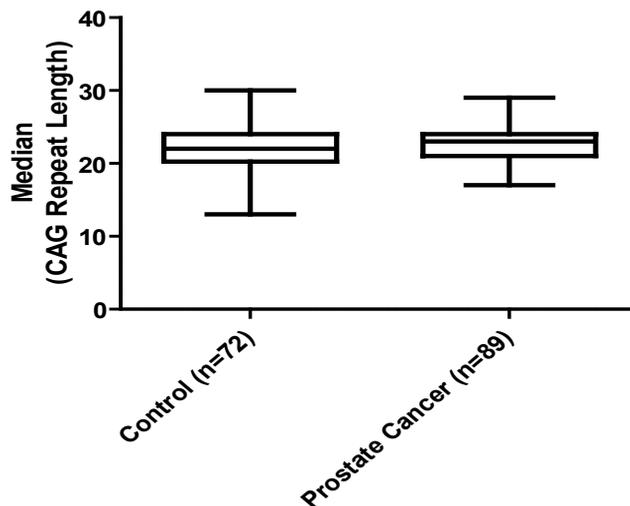


FIG. 2: Comparison of median CAG repeat in men with prostate cancer and men with benign prostatic hyperplasia (Control group).

TABLE 2: Odds Ratio (O) and 95% CI comparing long versus short CAG repeat alleles in cases and Benign Prostatic Hyperplasia (BPH) controls

Group	CAG Repeats		Total
	≤22	>22	
Prostate cancer	25 (28.1%)	64 (71.9%)	89
BPH (controls)	24 (33.3%)	48 (66.7%)	72

OR= 0.781 (0.398-1.532)

Next, we compared the CAG repeat length between different ethnicities in the prostate cancer and BPH groups. Whilst the median CAG repeat lengths were shorter in prostate cancer cases of Chinese (CAG=22) ethnicity compared to Malay and Indian men (both CAG=23), these differences were not significant ($p=0.578$) (Table 3 and Figure 3).

The relation between the polymorphism of repetitive CAG repeats and clinical and pathological parameters

In men with prostate cancer, no significant association was found between CAG repeat length and; age ($p = 0.986$), iPSA ($p = 0.355$) or Gleason score ($p = 0.980$). However, CAG repeat length was positively correlated with TNM staging (Spearman's rank $r = 0.229$, $p=0.038$), D'Amico risk classification for prostate

TABLE 3: The median CAG repeat lengths in the Androgen Receptor of Malaysian men with prostate cancer and benign prostatic hyperplasia (BPH)

Ethnicity	Median CAG repeat length			
	<i>n</i>	BPH	<i>n</i>	Prostate cancer
Malay	28	22.5	27	23
Chinese	29	22	31	22
Indian	15	22	31	23

Comparison of Median CAG repeat length in prostate cancer cases and BPH across ethnic groups

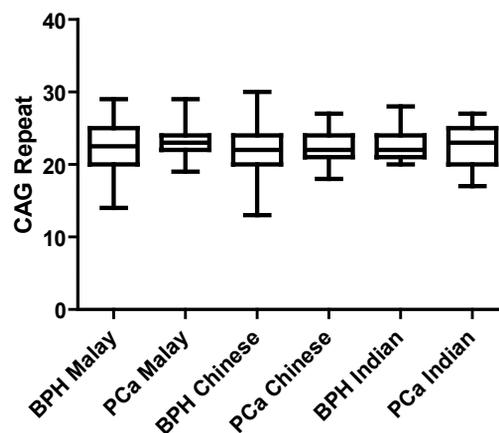


FIG. 3: Comparisons of Median CAG polymorphic repeats in the prostate cancer group and BPH group across ethnicity.

cancer (Spearman’s rank $r = 0.218$, $p = 0.036$) and UICC staging of prostate cancer (Spearman’s rank $r = 0.230$, $p = 0.038$).

Expression of ERG and AR in men with prostate cancer and comparison to CAG repeat length

Of the 89 cases of prostate cancer with matched blood samples, 58 cases with adequate invasive cancer in the biopsy samples were tested for expression of ERG and AR, which were expressed in 21/58 (36.2%) and 50/58 (86.2%) of cases, respectively. No significant associations were found between ERG or AR expression and patient age, PSA level at diagnosis, Gleason score, TNM staging and metastasis. However, a significantly great proportion of the Indian men with prostate cancer had ERG positive tumours

(11/18, 61%) compared to men of Malay ethnicity (8/24, 33%) and Chinese ethnicity (2/16, 13%) ($p=0.034$). (Table 4). We further compared the CAG repeat length in ERG positive and negative cases (Figure 4). The average CAG repeat length in the ERG negative group (22.84) was not significantly different to that of the ERG positive group (22.64) ($p=0.796$). In addition, there was also no significant difference between the CAG repeat length of AR-positive (22.8) and AR-negative (23.29) group ($p= 0.913$). The low numbers of cases did not allow for a meaningful comparison of ethnicity, CAG repeat length and ERG expression. When comparing ERG to AR expression in the tissue biopsies, they were found to have an inverse relationship ($p<0.01$).

TABLE 4: Correlation of immunohistochemical ERG and AR expression with clinicopathological parameters on needle biopsy

Parameters	ERG Expression		p	AR expression		p
	-ve (n=37)	+ve (n=21)		-ve (n=8)	+ve (n=50)	
Mean age (range)	68.65 (52-82)	68.67 (53-81)	0.929	67.5 (53-82)	68.84 (52-81)	0.65
PSA (ng/ml) (mean; range)	236 (2-2603)	142 (4-899)	0.728	80 (9-358)	221 (2- 2603)	0.505
Ethnicity			0.013			0.704
Malay	16 (43.2)	8 (38.1)		4 (50.0)	20 (40.0)	
Chinese	14 (37.8)	2 (9.5)		1 (12.5)	15 (30.0)	
Indian	7 (18.9)	11 (52.4)		3 (37.5)	15 (30.0)	
Gleason Score			0.383			0.361
<7	6 (16.2%)	4 (19.0%)		1 (12.5%)	9 (18.0%)	
7	10 (27.0%)	9 (42.9%)		1 (12.5%)	18 (36.0%)	
≥8	21 (56.8%)	8 (38.1%)		6 (75%)	23 (46.0%)	
TNM Staging			0.610			0.168
T1/T2	12 (32.4%)	9 (42.9%)		5 (62.5%)	16 (32%)	
T3	6 (16.2%)	2 (9.5%)		1 (12.5%)	7 (14%)	
T4, N1/Mx, M1	19 (51.4%)	10 (47.6%)		2 (25%)	27 (54%)	
D’amico Risk			0.640			0.459
1	1 (2.7%)	1 (4.8%)		1 (12.5)	1 (2.0%)	
2	3 (8.1%)	2 (9.5%)		0	5 (10.0%)	
3	33 (89.2%)	17 (81.0%)		7 (87.5%)	43 (86.0%)	
4	0	1 (4.8%)		0	1 (2.0%)	

Significant values are in bold.

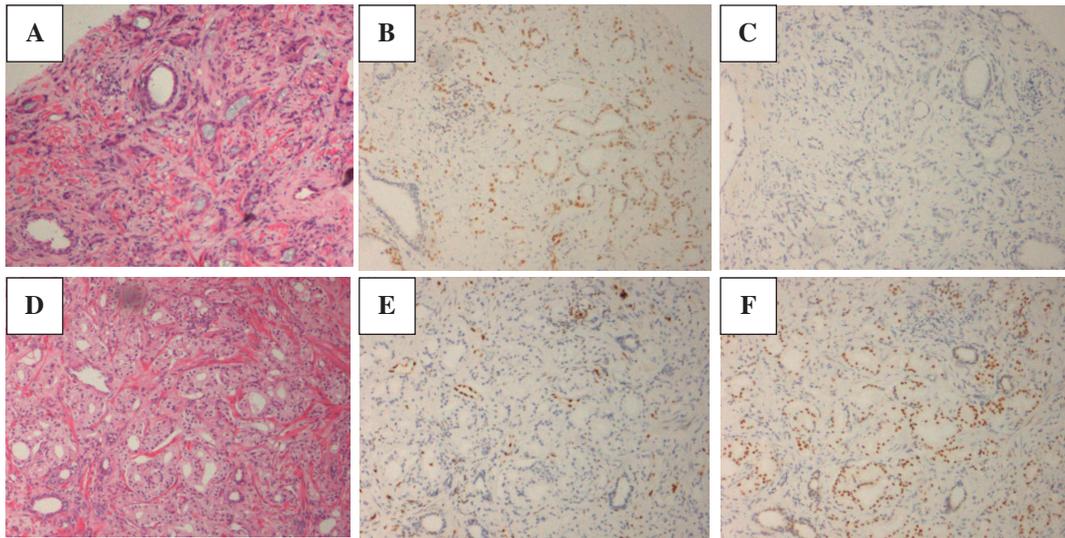


FIG. 4: Prostatic adenocarcinomas stained for both AR and ERG. Immunohistochemical expression of ERG and androgen receptor (AR) in two different prostatic adenocarcinomas; Case 1: A) Haematoxylin and Eosin (H&E) B) ERG positive, C) AR negative, Case 2: D) H&E, E) ERG negative, F) AR positive. NB: Normal endothelial cells also stain positive for ERG and serve as a useful internal positive control (e). Magnification x200.

DISCUSSION

The association between the CAG repeat length polymorphism and prostate cancer remains elusive as the results from different studies are conflicting.^{16,17} Whilst a meta-analysis revealed an overall shorter CAG repeat lengths in cancer patients compared to controls, the difference was less than one repeat.²²

This is the first study from Malaysia that has compared the CAG repeat length in the prostatic tissue biopsies of patients with prostate cancer and with BPH. Both shorter and longer CAG repeat length has been reported in cancer cases compared to benign prostate lesions.^{22,23} However, in our study the mean CAG repeat length for both prostate cancer and BPH groups was very similar.

Other than age and family history, ethnicity has been suggested to be one of the most important risk factors for prostate cancer.¹⁵ Previous studies reveal a difference in the mean number of CAG repeats across different ethnic groups with the shortest sequence of repeats seen in African Americans, intermediate sequences in Caucasians, and largest sequences in Asians.¹⁵ However, when we stratified the results of the prostate cancer cases based on ethnicity and CAG repeat length, we failed to detect any association.

Some studies have demonstrated a significant association between CAG repeat length and the age of diagnosis.^{13,14} In our study, when the CAG

repeat lengths were stratified for age and PSA at diagnosis, no significant associations were found.

A shorter CAG repeat length in prostate cancer patients with high grade (Gleason score > 7), late stage disease, or distant metastasis have been reported.^{13,24} However, many studies including our own, do not support the association between shorter CAG repeat length and Gleason score or later stage.^{14,25} On the contrary, we found a significant association between longer CAG repeat length, the D'amico risk classification and the UICC staging system.

Studies from Asia show differing frequencies of the TMPRSS2:ERG fusion or ERG expression compared to Western studies.^{18,19} It was suggested that these differences may be due to the relatively lower androgen levels and longer CAG repeats in Asian men in comparison to Western men.²⁶ In the present study, in which we investigated the expression of ERG in prostate biopsies, ERG tumour expression occurred in 36% of prostate cancer cases, which is consistent with studies reporting a lower frequency in Asian men with prostate cancer.²⁷ However, we found no association between the CAG repeat length and the expression of ERG or AR in the matched samples of patients' serum and tissue biopsy. This concurs with similar results reported by Mao *et al.* (2014).²⁸

There have been contrasting results with respect to tumour expression of ERG and tumour

grade. Whilst some studies report no association with Gleason scores^{29,30} others have reported an association with lower Gleason score and longer survival.³¹ In our study, we failed to detect any statistically significant association between the tumour ERG expression and preoperative serum PSA, age of diagnosis, UICC staging, and Gleason score.

When we stratified the ERG expression by ethnicity, we noticed a significant difference in the frequency of ERG expression. The Malaysian Indian patients demonstrated the highest positivity, followed by Malay and Chinese Malaysians. This is in concordance to a previously reported study involving a different cohort of patients following radical prostatectomy in which there was considerable tumour heterogeneity in ERG expression when all tissue blocks were examined. This in turn was significantly associated with patient ethnicity; with the highest and lowest levels of ERG heterogeneity observed in the tumours of Chinese and Indian patients, respectively.³² Intriguingly, the current study along with the two other studies conducted in the region, on completely different cohorts of men all showed this significant difference in ERG expression in the tumours of Malaysian Indian and Malaysian Chinese men with prostate cancer. This suggests that ERG expression may be associated with one of the main driving mutations for prostate cancer in Indian patients from Malaysia, but not in Malaysian Chinese patients whose tumours are predominantly ERG negative.^{27,32}

In the current study, there was a significant inverse relationship between tumour expression of AR and ERG. *In vitro* studies by Yu *et al.* (2010) previously demonstrated that overexpression of ERG significantly decreased AR transcripts as well as AR protein expression in multiple cell lines, whereas knock down of the ERG gene in VCaP cells resulted in AR upregulation.³³ Mapping of the genomic landscape surrounding AR and ERG, and the use of *in vitro* studies, shows that ERG is capable of binding to AR and a majority of AR target genes to disrupt AR signalling and prostate specific differentiation and can potentiate a stem cell-like de-differentiation program in the cell lines. Consequently, it is hypothesised that in prostate cancer ERG may act as an early molecular switch to revert cells to a more primitive state.³³ Such repression of AR by ERG may be reflected in the current study where we observed a significant inverse relationship between ERG and AR expression

in clinical biopsies. Previous studies have failed to consistently delineate a prognostic role for ERG in prostate cancer. We suggest that the true prognostic value of ERG in clinical cases may only become apparent when it is investigated in relation to its effects on AR expression and prostate cell differentiation in larger clinical studies.

In summary, we found no differences in the CAG repeat length in men with prostate cancer and men with BPH. Similarly, we found no association between ethnicity and mean CAG repeat length, either in patients with prostate cancer or BPH. However, we report a positive association between the mean CAG repeat length and tumour stage. There was no association between CAG repeat length and ERG or AR expression in the tumours of prostate cancer patients and ERG and AR expression were found to be inversely related. Lastly, as previously reported on a different Malaysian cohort³², the expression of ERG was significantly more common in the prostate cancers of Malaysian men of Indian ethnicity, than those of either Malay or Chinese descent.

Acknowledgement

We acknowledge financial assistance from the following University of Malaya grants; J-20012-73870, BKP grant- BKS053-2017, PPP grant- PG180-2015B, UM.C/625/1/HIR/MOE/MED/12, UM.C/625/1/HIR/224.

REFERENCES

1. Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate Cancer. *J Clin Oncol* 2010;28: 1117-23.
2. Perner S, Demichelis F, Beroukhim R, *et al.* TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res* 2006; 66: 66.
3. Haffner MC, Aryee MJ, Toubaji A, *et al.* Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nat Genet.* 2010; 42(8): 668-75.
4. Hagen RM, Adamo P, Karamat S, *et al.* Quantitative analysis of ERG expression and its splice isoforms in formalin-fixed, paraffin-embedded prostate cancer samples: association with seminal vesicle invasion and biochemical recurrence. *Am J Clin Pathol* 2014; 142: 533-40.
5. Mani RS, Tomlins SA, Callahan K, *et al.* Induced chromosomal proximity and gene fusions in prostate cancer. *Science* 2009; 326: 1230.
6. Lin C, Yang L, Tanasa B, *et al.* Nuclear receptor-induced chromosomal proximity and DNA breaks

- underlie specific translocations in cancer. *Cell* 2009; 139: 1069-83.
7. Heinlein CA, Chang C. Androgen Receptor in Prostate Cancer. *Endocr Rev* 2004; 25: 276-308.
 8. Li J, Al-Azzawi F. Mechanism of androgen receptor action. *Maturitas* 2009; 63: 142-48.
 9. Ding D, Xu L, Menon M, Reddy GP, Barrack ER. Effect of a short CAG (glutamine) repeat on human androgen receptor function. *Prostate* 2004, 58:23-32.
 10. Chamberlain NL, Driver ED, Miesfeld RL. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Res* 1994; 22: 3181-6.
 11. Hakimi JM, Rondinelli RH, Schoenberg MP, Barrack ER. Androgen-receptor gene structure and function in prostate cancer. *World J Urol* 1996; 14: 329-37.
 12. Hakimi JM, Schoenberg MP, Rondinelli RH, Piantadosi S, Barrack ER. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin Cancer Res* 1997; 3: 1599-608.
 13. Bratt O, Borg A, Kristoffersson U, Lundgren R, Zhang QX, Olsson H. CAG repeat length in the androgen receptor gene is related to age at diagnosis of prostate cancer and response to endocrine therapy, but not to prostate cancer risk. *Br J Cancer* 1999;81: 672-6.
 14. Beilin J, Harewood L, Frydenberg M, *et al.* A case-control study of the androgen receptor gene CAG repeat polymorphism in Australian prostate carcinoma subjects. *Cancer* 2001; 92: 941-9.
 15. Irvine RA, Yu MC, Ross RK, Coetzee GA. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res* 1995; 55: 1937-40.
 16. Lindström S, Ma J, Altshuler D, *et al.* A large study of androgen receptor germline variants and their relation to sex hormone levels and prostate cancer risk: Results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium. *Mol Endocrinol* 2010; 24: 1498-98.
 17. Price DK, Chau CH, Till C, *et al.* Androgen receptor CAG repeat length and association with prostate cancer risk: Results from the prostate cancer prevention trial. *J Urol* 2010; 184: 2297-302.
 18. Mao X, Yu Y, Boyd LK, *et al.* Distinct genomic alterations in prostate cancers in Chinese and Western populations suggest alternative pathways of prostate carcinogenesis. *Cancer Res.* 2010; 70(13): 5207-12.
 19. Lee K, Chae JY, Kwak C, Ku JH, Moon KC. TMPRSS2-ERG gene fusion and clinicopathologic characteristics of Korean prostate cancer patients. *Urology* 2010; 76: 1268.e7-13.
 20. Bastus NC, Boyd LK, Mao X, *et al.* Androgen-induced TMPRSS2:ERG fusion in non-malignant prostate epithelial cells. *Cancer Res* 2010; 70: 9544-48.
 21. McCarty KS, Jr., Szabo E, Flowers JL, *et al.* Use of a monoclonal anti-estrogen receptor antibody in the immunohistochemical evaluation of human tumors. *Cancer Res* 1986; 46: 4244-48.
 22. Zeegers MP, Kiemeny LA, Nieder AM, Ostrer H. How strong is the association between CAG and GGN repeat length polymorphisms in the androgen receptor gene and prostate cancer risk? *Cancer Epidemiol Biomarkers Prev* 2004; 13: 1765-71.
 23. Zhai XL, Qu XW, Guo L, Ha QH. Correlation study between the polymorphism of repetitive sequence in gene CAG of androgen receptor and the occurrence and progression of prostate cancer. *Asian Pac J Trop Med* 2014; 7: 301-4.
 24. Giovannucci E, Stampfer MJ, Krithivas K, *et al.* The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci U S A* 1997; 94: 3320-3.
 25. Hsing AW, Gao YT, Wu G, *et al.* 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* 2000; 60: 5111-6.
 26. Tirabassi G, Cignarelli A, Perrini S, *et al.* Influence of CAG Repeat Polymorphism on the Targets of Testosterone Action. *International Journal of Endocrinology* 2015; 2015: 298107.
 27. Kelly GM, Kong YH, Dobi A, *et al.* ERG oncoprotein expression in prostate carcinoma patients of different ethnicities. *Molecular and Clinical Oncology* 2015; 3: 23-30.
 28. Mao X, Li J, Xu X, *et al.* Involvement of different mechanisms for the association of CAG repeat length polymorphism in androgen receptor gene with prostate cancer. *Am J Cancer Res* 2014;4: 886-96.
 29. Perner S, Mosquera JM, Demichelis F, *et al.* TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. *Am J Surg Pathol* 2007; 31: 882-8.
 30. Lapointe J, Kim YH, Miller MA, *et al.* A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. *Mod Pathol* 2007; 20: 467-73.
 31. Winnes M, Lissbrant E, Damber JE, Stenman G. Molecular genetic analyses of the TMPRSS2-ERG and TMPRSS2-ETV1 gene fusions in 50 cases of prostate cancer. *Oncol Rep* 2007; 17: 1033-6.
 32. Tan JSJ, Ong KC, Ong DBL, *et al.* Heterogenous expression of ERG oncoprotein in Malaysian men with adenocarcinoma of the prostate. *Malays J Pathol* 2018; 40: 103-10.
 33. Yu J, Yu J, Mani RS, *et al.* An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* 2010; 17: 443-54.