ANDROGEN RECEPTOR POLYMORPHISMS: ASSOCIATION WITH PROSTATE CANCER RISK, RELAPSE AND OVERALL SURVIVAL

Stephen M. Edwards¹, Mike D. Badzioch², Ralph Minter³, Rifat Hamoudi⁴, Nadine Collins⁵, Audrey Arden-Jones³, Anna Dowe⁵, Simon Osborne⁵, Jo Kelly⁴, Robert Shearer³, Doug F. Easton⁶, Grady F. Saunders⁵, David P. Dearnaley⁶ and Rosalind A. Eeles¹,³,☆

¹CRC Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, UK
²Unit of Genetic Epidemiology, International Agency for Research on Cancer, Lyon, France
³Royal Marsden NHS Trust, Sutton, Surrey, UK
⁴CRC Section of Genetic Epidemiology, Strangeways Laboratories, Cambridge, UK
⁵University of Texas, M.D. Anderson Cancer Center, Houston, Texas, USA
⁶Academic Radiotherapy Unit, Royal Marsden NHS Trust, and Institute of Cancer Research, Sutton, Surrey, UK

Several reports have suggested that one or both of the trinucleotide repeat polymorphisms in the human androgen receptor (hAR) gene, (CAG)n, coding for polyglutamine and (GGC)n, coding for polyglycine, may be associated with prostate cancer risk; but no study has investigated their association with disease progression. We present here a study of both hAR trinucleotide repeat polymorphisms not only as they relate to the initial diagnosis but also as they are associated with disease progression after therapy.

Lymphocyte DNA samples from 178 British Caucasian prostate cancer patients and 195 control individuals were genotyped by PCR for the (CAG), and (GGC), polymorphisms in hAR. Univariate Cox proportional hazard analysis indicated that stage, grade and GGC repeat length were individually significantly associated with disease-free survival (DFS) and overall survival (OS). The relative risk (RR) of relapse for men with more than 16 GGC repeats was 1.74 (95% CI 1.08–2.79) and increasing from any cause, 1.98 (1.13–3.45). Adjusting for stage and grade, GGC effects remained but were not significant (RRDFS = 1.60, p = 0.052; RROS = 1.65, p = 0.088). The greatest effects were in stage T1–T2 (RRDFS = 3.56, 95% CI 1.13–11.21) and grade 1 (RRDFS = 6.47, 95% CI 0.57–72.8) tumours. No differences between patient and control allele distributions were found by odds-ratio analysis, nor were trends with stage or grade evident in the proportion of short CAG alleles. Non-significant trends with stage and grade were found in the proportion of short GGC alleles. The (GGC)n polymorphism in this population is a significant predictor of disease outcome. Since the (GGC)n effect is strongest in early-stage tumours, this marker may help forecast aggressive behaviour and could be used to identify those patients meritting more radical treatment. Int. J. Cancer (Pred. Oncol.) 84:458–465, 1999.

© 1999 Wiley-Liss, Inc.

Prostate cancer has been described as a disease that will affect every man if he lives long enough. Many studies have shown that age is the most important risk factor for prostate cancer in virtually all populations studied. It is also believed that underlying the wide variation among populations in the age-adjusted rates of clinical disease there is a universally high rate of development of histological or latent prostate cancer that also increases with age (Carter et al., 1991). It has been reported that AR may regulate genes involved in cell-cycle control, e.g., CDK2, CDK4 and p16 (Luo et al., 1997). The 8 exons of hAR define 3 major protein functional domains: a C-terminal ligand-binding domain, a central DNA-recognition and-binding domain and an N-terminal transcription-activation domain. Within the latter domain are 2 single–amino acid repeat length polymorphisms: the first, a polyglutamine tract, is coded by a repeated CAG (Sledsens et al., 1992); the second, a variable number of glycine residues, is coded by a tandem array of GGC sequences (Sledsens et al., 1993).

When case-control studies have found associations between prostate cancer risk and one or both of the hAR trinucleotide repeat polymorphisms, the risks have tended to be strongest for men with shorter (CAG)n alleles and higher clinical stage and tumour grade, i.e., more aggressive tumours (Irving et al., 1995; Giovannucci et al., 1997; Hardy et al., 1996; Ingles et al., 1997; Stanford et al., 1997). Hakimi et al. (1997) observed that the frequency of shorter (CAG)n alleles was substantially higher in patients and that a significant proportion of these patients, who were candidates for prostatectomy, had unsuspected metastatic disease identified as lymph node-positive prostate cancer. Hardy et al. (1996) extended

Grant sponsors: National Cancer Institute; Prostate Cancer Charitable Trust; Prostate Research Campaign UK; Neil MacTaggart Fund; Cancer Research Campaign UK.

S.M. Edwards and M.D. Badzioch are considered equal first authors.

M.D. Badzioch’s former address is: University of Texas, M.D. Anderson Cancer Center, Houston, TX, USA.

*Correspondence to: CRC Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG UK. Fax: +44 (0)181 770 1489. E-mail: ros@icr.ac.uk

Received 14 December 1998; Revised 27 April 1999
the analysis to include possible effects on the clinical course of disease (PSA-related relapse), but found no association with (CAG) length. Differences in average (CAG) repeat length have been found to exist among ethnic groups and have correlated with the prostate cancer risk among those populations (Irvine et al., 1995). In vitro investigations have suggested that variation in (CAG) length affects the AR capacity to stimulate transcription at target gene promoters (Kazemi-Esfarjani et al., 1995; Chamberlain et al., 1994). These functional results indicate that fewer CAG triplets result in a more active receptor in controlled conditions and are correlated with ethnic variations in prostate cancer risk.

Length variation of the (GGC)n sequence has been less widely studied for several reasons, resulting in a possibly biased assessment of risk. For instance, the sequence of GGC triplets is inherently more difficult to efficiently amplify using PCR, is less polymorphic and has not been shown to have significant ethnic variation. The results from (GGC)n analysis are also conflicting. Whereas Irvine et al. (1995) found long (GGC)n alleles to be associated with increased risk, others (Stanford et al., 1997; Hakimi et al., 1997) found that men with short alleles were at increased risk.

Overall, these results implicate the involvement of normal AR polymorphisms in the natural history of prostate cancer, but the mechanism or nature of that involvement remains poorly defined. For example, which of them is involved in prostate cancer pathogenesis, or are they both involved? Irvine et al. (1995) have provided evidence that the (CAG)n and (GGC)n sequences may be in linkage disequilibrium in prostate cancer patients and not in control individuals. Also, to date, all studies of AR polymorphisms and prostate cancer have been undertaken in North America. Studies of migrating populations have shown that prostate cancer risks may be significantly influenced by place of residence (Swerdlow et al., 1995; Shimizu et al., 1991). Dietary fat intake has been suggested to be the basis for some of the environmentally induced differences in prostate cancer rates between populations (Whitemore et al., 1995). Overall, the extent to which the selection of hAR polymorphisms in prostate cancer may be influenced by environmental factors remains to be addressed.

We hypothesised that if an association between hAR and prostate cancer risk exists, it would also be seen, and may even be stronger, in relation to the post-diagnosis clinical course of the disease. It was our belief that disease aggressiveness and clinical behaviour would be more closely monitored and less susceptible to confounding by extraneous factors such as socio-economic status than would be disease status at diagnosis. Additionally, although studies of the conservative management of localised disease have determined overall progression rates (Chodak et al., 1994), it remains difficult to predict the time course of progression within each component of the localised stages, and further determinants and markers of disease advancement need to be found. The present study of both hAR trinucleotide polymorphisms emphasises disease outcome, rather than characteristics at diagnosis, in a European population. We have genotyped 178 British prostate cancer patients to assess risk by comparison with population-based controls, to correlate their genotypes with several clinical characteristics of the disease and to determine if any differences exist in hAR-associated risk patterns between previously published reports and these British patients.

MATERIAL AND METHODS

Patients

Prior to August 1993, 201 referred cases of prostate cancer were systematically ascertained at the Royal Marsden NHS Trust (RMH), London, UK, to study the effects of inherited mutations and polymorphisms on prostate cancer development. Of these, 183 were Caucasian, 13 were black (native Africans and Caribbeans) and 5 were Asian. Since the numbers of black and Asian men were not sufficient to study ethnicity-specific effects, they are not included in this analysis. On review, 5 of the Caucasian patients were found to not have primary cancer of the prostate, i.e., stage T0, and so were excluded. The average age at presentation of the 178 Caucasian patients was 68.1 years (SD = 6.82 years), with a range of 48 to 80 years, and the median date of admission to RMH was March 1992. There was no association of age with year of presentation. After receiving informed consent through interview, demographic data, detailed family history and a 10-ml peripheral blood sample were obtained. Tumour specimens were also collected for histological review. Tumours were classified as grade 3 if any poorly differentiated cells/foci were present (Gleason score 8 or greater) and grade 2 if only moderate differentiation was seen (Gleason score 4 to 7) and grade 1 if all cells/foci were well-differentiated (Gleason score <4). We were unable to obtain pathology specimens or to verify tumour grade in 16 cases, and these were not considered for further analysis except where indicated. Disease staging was performed by clinical examination, abdomino-pelvic computerised tomography and bone scans and serum PSA measurement. Patients were categorised into 1 of 6 groups based on the TNM classification system: T1, non-palpable tumour; T2, palpable but organ-confined; T3 and T4, penetrating margins but no metastasis identified; N+, pelvic nodal involvement; and M+, widespread metastasis. Patients were treated with hormone therapy alone, radiotherapy or a combination of these. Those presenting with higher stage disease were generally given hormone treatment alone. After therapy, patients were routinely assessed for disease progression at 6-month intervals until April 1997, the end of the study. At this time, 80 (49.4%) were alive and well at last routine clinic visit, 54 (33.3%) had died and 28 (17.3%) had not been seen by the clinic for over 6 months.

Controls

Control genotypes were obtained from the DNA of Caucasian females who were age- and location-matched to breast cancer patients for a previous breast cancer case-control study conducted by the UK National Case-Control Study Group (1989). Up to 195 women were available for analysis. These DNA samples provided the source of data for allele frequency comparisons, and as samples were from females, data are presented on up to 390 X chromosomess. All controls were cancer-free, randomly selected via general practice surgeries and aged up to 45 years.

DNA extraction

DNA was extracted from blood samples (10 ml) as previously described (Edwards et al., 1997) with the inclusion of a second protease K digestion at 50°C for 2 to 3 hr. DNA was dissolved in 0.2 to 0.4 ml of water (BDH, Poole, UK) and stored at –20°C until required.

Radioactive end-labelling

The appropriate amplification primer (see below) was end-labelled in a 30 µl reaction volume containing 1× reaction buffer (Stratagene, Cambridge, UK), 0.33 units T4-PNK enzyme (Stratagene) and 0.17 to 0.34 mCi 32P-γ-dATP (ICN, High Wycombe, UK; 7,000 Ci/mmol).

(CAG)n amplification

PCR was conducted with primers 5′-tccccgaagttatggcaagc and 5′-ctggtggagacactctca. Reactions (25 µl) were conducted with Perkin-Elmer (Beaconsfield, UK) reaction buffer [50 mM KCl, 10 mM Tris-HCl (pH 8.3)], 0.8 mM MgCl2 (Perkin-Elmer), 0.2 mM each dNTP (0.8 mM total, Stratagene), primers at 0.1 OD/ml (0.5 µM final), 0.75 units AmpliTag Gold (Perkin-Elmer) and 100 ng DNA. Thermocycling consisted of 95°C for 9 min, 4 cycles of 1 min at 94°C, 30 sec at 64°C and 1 min at 72°C, followed by a ’touchdown’ reduction of annealing temperature by 2°C every 4 cycles until 54°C (20 cycles on ‘touchdown’), then another 29 cycles at 54°C annealing, with a final polymerase extension for 10 min at 72°C.
(GGC)_n amplification

PCR was conducted with primers 5'-tcctggcacactcttcac and 5'-gcagggctacaactcaggt. Reaction conditions were as described above with the following modifications: 1.3 mM MgCl₂, 0.25 mM each dNTP [1.00 mM total with 3:1 mixture of 7-deaza-dGTP: dGTP (Pharmacia, Milton Keynes, UK)], 2.5% v/v DMSO (Sigma, Poole, UK; molecular grade), primers at 0.15 OD/ml (0.7µM final concentration) and 1.0 unit AmpliTaq Gold. Thermocycling consisted of 12 min at 95°C, 21 cycles of 1 min at 96°C, 1 min at an initial annealing temperature and 1 min at 72°C, using a “computed touchdown” algorithm starting with an annealing temperature of 64°C and ending at 54°C. This was equivalent to reducing the annealing temperature by 0.5°C per cycle. A further 25 cycles were conducted at the annealing temperature of 54°C, with a final polymerase extension for 10 min at 72°C.

Polyacrylamide gels

PCR products were denatured at 95°C for 6 min with 23 µl of formamide (containing bromophenol blue and xylene CFF), placed on ice for 3 min, then loaded onto a 6% polyacrylamide gel (40%; 19:1 bis-acrylamide; National Diagnostics, Hull, UK) containing urea (50% w/v). Gels were loaded with 3 µl of sample and run at 80 W for 3 to 4 hr in 1× GTB buffer (glycerol tolerant buffer, Tris-base, taurine, sodium-EDTA). Gels were moved to Whatman 3MM paper, dried and exposed to autoradiography film (Amer sham, Aylesbury, UK).

Repeat length determination

Representative samples were chosen for repeat length determination using ABI 377 and 310 automatic fluorescent DNA sequencers (ABI, Perkin-Elmer) in “Sequencing” and “Genescan” modes of the ABI software. Aliquots of these samples were then run routinely on the gels to aid in the identification of sample genotypes. For (CAG)ₙ genotyping, a ladder was constructed from samples with 15, 17, 19, 21, 23, 25, 27, 28 and 32 repeats and was run in every 8th lane. For (GGC)ₙ genotyping, a sample with 16 repeats was run at regular intervals.

Statistical analysis

Odds ratios (ORs) were calculated without age or other adjustments, and all χ² calculations were done with correction for continuity. To test for trends in ORs, a χ² with 1 degree of freedom was generated according to equation 4.31 in Breslow and Day (1980). Cox proportional hazard analysis, t-tests and Kaplan-Meier survival estimates were generated by Stata (College Station, TX) software, version 5.0, which also compared survival curves by log-rank tests.

RESULTS

The patient and control allele distributions for the (CAG)ₙ polymorphism are shown in Figure 1a and those for the (GGC)ₙ polymorphism in Figure 1b. The allele frequencies for controls were obtained from Caucasian females (see “Material and Methods”). Data are presented as number of chromosomes analysed (390 and 284, respectively). For each region, the same number of repeated triplets was most common in both patients and controls: 21 CAG repeats and 16 GGC repeats. Overall, there was no statistically significant difference between patient and control distributions for either polymorphism. However, there were distinct differences between the 2 polymorphisms. In the (CAG)ₙ system, 5 of the 19 observed alleles occurred at a frequency of over 10%. In the (GGC)ₙ system, only 2 alleles, those with 16 and 17 repeats, accounted for over 80% of all 13 alleles genotyped in patient and control populations. A single additional (GGC)ₙ allele, of 12 repeats, occurred at a frequency of greater than 5%.

The ORs for the risk of prostate cancer by type and length of repeat are shown in Table I. Calculations were performed using the entire series of patients (n = 178) and separately using only those patients for whom complete histological data were available (n = 162). The results indicate that there is no statistical association between prostate cancer risk and the number of repeated sequences in either the (CAG)ₙ or (GGC)ₙ polymorphism of hAR. Allele distributions were subdivided at other repeat lengths with no significant association found (data not shown). Table II shows the numbers of patients of each clinical stage who had long and short alleles at each polymorphism. Although the ratios of (CAG)ₙ repeat length (short ≥ 21 repeats vs. long > 21 repeats) for stages T3–T4 did not differ from those for T1–T2, there is evidence of a trend for an increasing proportion of longer (CAG)ₙ repeat lengths in association with metastatic disease; however, this was not significant (p = 0.07). A similar non-significant trend (p = 0.10) was observed within (GGC)ₙ repeat length stage-specific proportions. Analysis by histological grade (Table III) indicated that although
neither polymorphism differed significantly, longer alleles may be associated with higher grade.

Table IV presents the results of univariate Cox proportional hazard analysis of factors potentially related to disease free survival (DFS) and overall survival (OS). During the course of the study, 75 (46.3%) of the patients experienced treatment failure (relapse) and 54 (33.3%) died. Each of these possible risk factors was assessed as both a categorical and a continuous variable, and the results did not depend on how the factor was coded. Overall, age at diagnosis was significantly associated with increased risk of relapse and death, compared to men with T1/T2 disease. Disease stage was significantly associated with both DFS and OS. For stages T3–T4, the risk of relapse and death were, respectively, 76% and 87% greater than those for stages T1–T2. The risk of relapse continued to rise but at a slower rate in men with confirmed metastatic spread [N+ or M+; relative risk (RR) = 2.07, 95% CI 1.06–4.03]. This group, however, had nearly a 6-fold increased risk of death compared to men with T1/T2 disease.

Significant trends in DFS and OS were also found in relation to tumour grade. Patients with tumours of moderately differentiated histology (grade 2) were 2.49 times more likely to relapse than those with well-differentiated tumours (grade 1). The risk significantly rose to over 4-fold (RR = 4.07, 95% CI 1.55–10.67) in men with poorly differentiated tumours (grade 3). Since none of the patients with grade 1 tumours died during the study period, these men were combined with those with grade 2 tumours to establish the baseline risk. Compared to grades 1/2, men with poorly differentiated tumours (grade 3) had a significantly worse OS (RR = 2.40, 95% CI 1.40–4.13).

The (CAG)ₙ polymorphism was not a significant factor in either DFS or OS. However, in both cases, longer alleles (>21 repeats) were associated with a mild decrease in disease risk. The (GGC)ₙ trinucleotide repeat polymorphism was significantly related to both DFS and OS when examined in a univariate model. Long (GGC)ₙ alleles (>16 repeats) increased the risk of relapse by a factor of 1.74 (95% CI 1.08–2.79). These alleles were also associated with a 1.98-fold (95% CI 1.13–3.45) increased risk of death.

In our study, 50 men had pre-treatment PSA levels available for analysis. For these, non-significant associations between PSA levels and DFS and OS were found, which, as expected, indicated a worse prognosis with increasing PSA levels.

Since both stage and grade have clinical relevance and were also significant factors in predicting DFS and OS in our patient population, we examined their associated RRs by Kaplan-Meier survival estimates, to verify their effects. Figure 2 presents these survival estimates for DFS. For both risk factors, the expected patterns were observed, i.e., higher stages and grades were associated with decreased DFS.

The results from multivariate Cox proportional hazard analysis are presented in Table V. This analysis included age in years at diagnosis and the factors stage, grade, (CAG)ₙ and (GGC)ₙ, coded as grouped variables, as in Table IV. Log(PSA) was included in a
In our patient population, men with long (CAG)ₙ alleles were at increased risk of a poor prognosis and worse survival. By stage and grade, the association of long alleles with worse survival was progressively more severe. In both genotype groups, the survival probability did not fall below approximately 0.25 during the course of the study. A similar delayed effect of 33 months before long (GGC)ₙ alleles became prominent associated with worse OS was also observed (Fig. 3b).

(GGC)ₙ genotype was significantly associated with DFS and OS when tested as a univariate predictor but became non-significant, p = 0.082 (DFS) and p = 0.088 (OS), when stage and grade were simultaneously included. To explain this finding, we thought that there may be some interaction between (GGC)ₙ and stage and/or (GGC)ₙ and grade in terms of their effects on DFS and OS. This possibility was also suggested by the previous finding of long (GGC)ₙ alleles being mildly correlated with stage (Table II) and grade (Table III). We tested each of the other possible risk factors, age, grade, stage and (CAG)ₙ genotype, in a 2-factor analysis of DFS and OS and found that the inclusion of grade information was most responsible for the decrease in (GGC)ₙ effect, with stage having the second largest effect on (GGC)ₙ (data not shown). We therefore analysed (GGC)ₙ genotype separately for 2 levels each of grade and stage. These results are presented in Table VI. In both measures of disease outcome, DFS and OS, long (GGC)ₙ alleles were associated with a worse prognosis at all levels of stage and grade. (GGC)ₙ genotypes on DFS were associated with both risk factors, with the largest effects in early, T1–T2, stages (RR = 3.56, 95% CI 1.13–11.21) and grade 1 (RR = 6.47, 95% CI 0.57–72.8) tumours. Long (GGC)ₙ alleles were also associated with decreased OS at all levels of stage and grade, though the RRs were smaller and the variation in the effects among the clinical/pathological classes was less than for DFS. In this case, the overall trend by stage was not evident; the highest risk, RR = 2.22 (95% CI 0.76–6.44), occurred in stages T3–T4. The effect of (GGC)ₙ length was significant in the combined group, grades 1 and 2 (RR = 2.21, 95% CI 1.01–4.85). A similar analysis of the (CAG)ₙ sequence showed that this polymorphism was not a significant factor in DFS or OS in any of the stage/grade categories (data not shown).

As Irvine et al. (1995) reported linkage disequilibrium between the hAR polymorphisms in prostate cancer patients, we tested the independence of 2 classifications of (CAG)ₙ alleles with 3 classes of (GGC)ₙ alleles by Pearson’s χ² test for independence of rows and columns, and the results were non-significant (χ²(ad) = 1.31, p = 0.511). We also performed χ²-tests to estimate the significance of the difference in mean (CAG)ₙ repeat number when partitioned by (GGC)ₙ length. The difference of 0.23 repeats was also not significant (t = 0.563, p = 0.574).

DISCUSSION

In the British population, we have found that compared to controls overall prostate cancer risk was not significantly associated with either hAR (CAG)ₙ or (GGC)ₙ variation. This result is in accord with several other studies in which no overall significant genotype-specific prostate cancer risk was found with either the (CAG)ₙ or (GGC)ₙ polymorphism compared with control groups (Irvine et al., 1995; Giovannucci et al., 1997; Stanford et al., 1997). When we looked for stage and grade associations, we found similar effects in the 2 polymorphisms; i.e., long alleles in both polymorphisms appeared at higher frequency in higher stages and higher grades. Although the association was stronger for the (GGC)ₙ polymorphism, the suggested association for the (CAG)ₙ system is opposite to what other investigators have generally found. One concern was that the clinical staging schemes are different in the United Kingdom (TNM system) than in the United States (ABC system). However, since both systems reflect the natural growth and metastatic potential of prostate cancer and score tumour progression on the same continuum, from intracapsular to capsular penetration to documented metastasis, we do not believe this would have affected our results.

We next examined the 2 polymorphisms as well as 4 other potential risk factors for their involvement in DFS and OS. Disease stage and grade were individually significant factors for each of these end-point measures. These correlations were not unexpected since both stage and grade are used as prognostic indicators in the clinic. Also in the univariate analysis, the (CAG)ₙ system behaved in a manner similar to that hypothesised by other workers in that short (CAG)ₙ alleles were associated with more aggressive disease. In our patient population, men with long (CAG)ₙ alleles were at less risk of relapse and death than those with short alleles. The risk differentials were not significant however. (GGC)ₙ variation was significantly related to DFS and OS in that longer alleles (i.e., those with 17 or more repeats) were associated with shorter time to
relapse and worse OS. This (GGC)n correlation is a new finding and with 17 repeats, similar to our (GGC)n higher-risk alleles produced have a functional role, at least in regulating tumour growth, if not in much stronger than, and independent of, DFS and OS association

To our knowledge, only 2 studies have reported a functional role for disease-free survival and overall survival.

These (GGC)n results indicate that the AR polyglycine tract may be associated with the disease aggressiveness, Quetelet index, family history and age. The association with disease aggressiveness was also suggested by the results. Data presented by Hakimi et al. (1997) suggest that short (GGC)n repeats, in addition to increasing the risk of prostate cancer, identify a subpopulation of patients with clinically localised disease. The report suggests that short glycine repeats may identify men who develop prostate cancer but with decreased aggressiveness. This is an intriguing suggestion and supports our observations on OS and DFS.

Irvine et al. (1995) found a significant risk association between short (CAG)n alleles and long (GGC)n alleles in all patients (p = 0.008) as well as when white patients only were examined (p = 0.015). In contrast and in agreement with Stanford et al. (1997) and Lumbroso et al. (1997), we found no evidence for allelic association between the 2 polymorphisms.

To our knowledge, only 2 studies have reported a functional role of the polyglycine tract in vitro: Jenster et al. (1994) found that complete deletion of the (GGC)n sequence had “no substantial effect on the activity of the hormone regulated AR”, whereas Gao et al. (1996) found that this same mutation resulted in a diminished capacity to activate the MMTV luciferase gene. Complex protein interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative interactions involving steroid receptors and positive and negative

### Table V: Multivariate Cox Proportional Hazards Analysis: RR and p Values for Risk Factors on Disease-Free Survival and Overall Survival

<table>
<thead>
<tr>
<th>Factor (n groups)</th>
<th>Disease-free survival</th>
<th>Overall survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR (n = 162 (75)) (^{1})</td>
<td>RR (n = 162 (54)) (^{1})</td>
</tr>
<tr>
<td></td>
<td>RR (n = 50 (21)) (^{1})</td>
<td>RR (n = 50 (15)) (^{1})</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.99 (0.414) 0.91 (0.660) 0.90 (0.660)</td>
<td>1.01 (0.517) 0.98 (1.46) 0.525 (1.46)</td>
</tr>
<tr>
<td>Stage (3)</td>
<td>1.30 (1.03) 0.86 (0.49) 0.39 (0.175)</td>
<td>2.81 (0.001) 4.17 (0.006)</td>
</tr>
<tr>
<td>Grade (3)</td>
<td>1.71 (0.005) 2.56 (0.024) 0.61 (0.097)</td>
<td>4.17 (0.006)</td>
</tr>
<tr>
<td>CAGL (2)</td>
<td>0.65 (0.052) 2.06 (0.173) 1.65 (0.088)</td>
<td>4.33 (0.021)</td>
</tr>
<tr>
<td>GGCL (2)</td>
<td>1.60 (1.70) 0.203</td>
<td>1.65 (0.525)</td>
</tr>
</tbody>
</table>

\(^{1}\)Number at risk (number affected). – CAGL and GGCL, alleles grouped as long (L), see Table IV.

### Table VI: Univariate Cox Proportional Hazards Analysis: RR of GGC Genotype on Disease-Free Survival and Overall Survival by Stage and Grade

<table>
<thead>
<tr>
<th>Factor</th>
<th>Disease-free survival</th>
<th>Overall survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR (95% CI)</td>
<td>RR (95% CI)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1–T2</td>
<td>3.56 (1.13–11.21)</td>
<td>2.16 (0.36–12.9)</td>
</tr>
<tr>
<td>T3–T4</td>
<td>1.76 (0.87–3.54)</td>
<td>1.22 (0.76–6.44)</td>
</tr>
<tr>
<td>N+–M+</td>
<td>1.03 (0.47–2.25)</td>
<td>1.65 (0.80–3.45)</td>
</tr>
</tbody>
</table>

### Figure 3 – Kaplan-Meier survival estimates by (GGC)n length (a) for disease-free survival and (b) for overall survival. Patients categorised by repeat length as indicated. p values indicate log-rank significance tests.
acting co-factors are beginning to be elucidated (Shibata et al., 1997). Since the (GGC)n sequence is in the transactivation domain, it is possible that the single-amino acid difference disrupts the binding affinity of the domain enough to up- or down-regulate the formation of a critical regulatory complex.

A graded relationship or interaction between the effects of (GGC)n genotype and tumour stage and grade in DFS was also found in our study and is important for 2 reasons. First, it suggests that the role of (GGC)n in regulating tumour growth is likely to be of heightened relevance during a limited time period out of the total multiple-step progression required for the transition from normal cell growth regulation to fully hormone-independent metastatic disease. Additionally, since the greatest RRs occurred in the earliest-stage tumours, these effects of (GGC)n genotype are likely to occur relatively early in the pathway. Second, it suggests potential clinical application, and this aspect is discussed later.

We have specifically examined the role of both hAR trinucleotide repeat polymorphisms as they relate to disease outcome. We found no evidence of an association between the (CAG)n polymorphism and prostate cancer risk but a significant association between the (GGC)n polymorphism and disease progression and outcome. We examined hAR genotypes in a population outside the United States. Whether the present findings are unique to British patients is unknown. However, dietary and other environmental factors have been associated with the clinical incidence of prostate cancer (Rose et al., 1986), and it has been suggested that the metabolic link between fat intake and prostate cancer risk may include androgenic pathways (Ross and Henderson, 1994). Also, it has often been found that the fat–prostate cancer risk correlation may be strongest in men with high-grade/advanced cancer (Wenig et al., 1995; Ross and Henderson, 1994). It is therefore possible that gene–environment interactions may affect the relationship between the hAR polymorphisms and prostate cancer progression.

With regard to clinical considerations, treatment option decisions for early-stage tumours, particularly those detected through PSA screening, are becoming frequent and have both economic as well as quality-of-life implications. Although further research is needed to validate and assess our present findings in the clinical setting, use of the (GGC)n genotype may offer relevant information to help predict the growth behaviour of early-stage tumours and determine who should be offered radical treatment.

ACKNOWLEDGEMENTS

We thank the Royal Marsden NHS Trust and M.D. Anderson patients who enrolled and contributed to this study. We thank Drs J. Peto and M. Stratton for providing the control DNA samples and Mrs. S. Seal for kindly databasing and storing the samples. We thank Dr D. Godgar for helpful discussion. Sequencing was conducted in the Jean Rook Gene Cloning Laboratory, which is supported by BREAKTHROUGH Breast Cancer Charity 328323. The study was supported by an NCI Predoctoral Fellowship in Cancer Prevention at the University of Texas M.D. Anderson Cancer Centre. Funds were also generously provided by The Prostate Cancer Charitable Trust, Prostate Research Campaign UK, The Neil MacTaggart Fund and The Cancer Research Campaign UK.

REFERENCES


