

Trinucleotide repeat length in the first exon of the androgen receptor gene may be associated with prostate carcinogenesis and facilitate prediction of prostate cancer aggressiveness

Artur Lemiński¹, Agnieszka Bińczak-Kuleta², Mariusz Kaczmarczyk²,
Karolina Skonieczna-Żydecka³, Andrzej Ciechanowicz², Marcin Słojewski¹

¹ Department of Urology and Urological Oncology, Pomeranian Medical University, Szczecin, Poland

² Department of Clinical and Molecular Biochemistry, Pomeranian Medical University, Szczecin, Poland

³ Department of Gerontobiology, Pomeranian Medical University, Szczecin, Poland

Introduction Prostate cancer (PC) is a major health problem among men in Western countries. The incidence of this tumor has been rising in most industrialized countries, including Poland, although in the United States, the number of new PC cases has declined slightly in recent years,^{1,2} along with a similar trend in mortality.² Despite these promising data, which are the result of widespread prostate-specific antigen (PSA) screening and an aggressive therapeutic approach, it remains a challenge to predict the aggressiveness of PC in an individual patient. The decline in PC mortality needs to be balanced with the overdiagnosis and overtreatment inherent to existing screening protocols.³ Therefore, there is an urgent need for new, more specific markers that would enable clinicians to predict clinical significance of PC and facilitate the choice of treatment.⁴

PC is an androgen-dependent disease. Androgens exert their action through the androgen receptor (AR), a ligand-activated transcription regulator located in the cell cytoplasm. The gene encoding the AR is located on the X chromosome (Xq11-12) and contains 8 exons, totaling 3600 base pairs (Supplementary material online, *Figure S1*). The variability of sequences of the AR gene is caused by regions with a variable number of trinucleotide repeats present within the first exon: (CAG)_n – rs4045402, (GGN)_n – rs3138869, and (CCN)_n, encoding polyglutamine, polyglycine, and polyproline, respectively. The (CCN)_n region is well conserved, while the repeat number of the (CAG)_n and (GGN)_n regions varies from 8 to 62 and from 10 to 31, respectively. Numerous authors have reported an inverse correlation between the

number of CAG and GGN repeats and AR transactivation activity.^{5,6} This phenomenon is believed to predefine the sensitivity of peripheral tissue to circulating androgen, possibly influencing the susceptibility to PC along with clinical and pathological characteristics of the tumor.⁷⁻⁹

Patients and methods We designed a prospective study, which was approved by the local ethics committee. A total of 120 consecutive men with PC, who consented to undergo a genetic analysis of blood, were enrolled into the study. We obtained 5 ml of blood from each patient to containers with EDTA-K (Sarstedt, Nümbrecht, Germany). Clinical data acquired for analysis included age at diagnosis, PSA, and the Gleason score (GS). We used a pathological stage after radical prostatectomy (RP) and clinical staging (digital rectal examination, transrectal ultrasound, and bone scan) to classify patients into 2 groups: organ-confined PC when pT≤2c after RP or locally advanced/metastatic PC when cT3/4, pT≥3a after RP, pN1, or M1 in a bone scan.

The median age of patients was 67 years (range, 48–87 years), and median PSA was 12.8 ng/ml (range, 3.44–2500 ng/ml). There were 72 patients (60.5%) with a GS of 6 or lower and 47 patients (39.5%) with a GS of 7 or higher; 39 patients were assigned to the organ-confined group, whereas 44 patients were considered as locally advanced / metastatic. The data on staging were unavailable in 37 patients. Our control group comprised umbilical cord blood samples obtained from 120 male newborns, whose parents agreed to an investigational use of samples.

Correspondence to:

Artur Lemiński, MD, PhD,
Klinika Urologii i Onkologii Urologicznej,
Pomorski Uniwersytet Medyczny,
ul. Powstańców Wielkopolskich 72,
70-111 Szczecin, Poland,
phone: +48 91 466 11 00,
e-mail: artur.leminski@pum.edu.pl
Received: November 8, 2016.

Revision accepted:

January 24, 2016.

Published online: February 19, 2016.

Conflict of interests: none declared.

Pol Arch Med Wewn. 2016;

126 (3): 201-203

doi: 10.20452/pamw.3297

Copyright by Medycyna Praktyczna,

Kraków 2016

TABLE 1 Frequencies of shorter (CAG)_n ≤ 18 repeat and longer (CAG)_n > 18 repeat alleles of the androgen receptor gene in patients with low-grade (Gleason score ≤ 6) and high-grade (Gleason score ≥ 7) prostate cancer

Gleason score	CAG n ≤ 18	CAG n > 18	P value
≤ 6, n (%)	14 (19.44)	58 (80.56)	0.017
≥ 7, n (%)	2 (4.26)	45 (95.74)	

Genomic DNA was isolated from blood leukocytes using the QIAapm® DNA Mini Kit (Qia-gen, Hilden, Germany); all samples were sufficient for amplification. Repeat length polymorphisms (RLPs) of microsatellite regions were identified with polymerase chain reaction (PCR), using specific pairs of primers: [(CAG)_n 5'-TCCAGAGC-GTGC GCGAAGTGAT-3' (sense) and 5'-CCT-GTGGGGCCTCTACGATGG-3' (antisense) – T_h 65°C (36 cycles); as well as (CCN)_n 5'-TCGCGAC-TACTACAACCTTCCACT-3' (sense) and 5'-AG-GTCCCCATAGCGGCACTG-3' (antisense) – T_h 54°C (33 cycles); as well as (GGN)_n 5'-GCAGT-GCCGCTATGGGGACCT-3' (sense) and 5'-CCT-GAGGGGGCCGAGTGATAGC-3' (antisense) – T_h 59°C (39 cycles)]. One batch of each sense primer was labeled with a fluorescence dye for identification. The length of the PCR products was evaluated with capillary electrophoresis, using an ABI PRISM® 3100-Avant genetic analyzer (Applied Biosystems, Foster City, California, United States) with a GeneScan™ 500 LIZ™ size standard (Supplementary material online, Figure S2). Random samples underwent sequencing to validate the identification of polymorphisms with capillary electrophoresis. Sequencing was performed with the ABI PRISM® 3100-Avant analyzer, and data were analyzed with the Sequencing Analysis v5.1 software (Applied Biosystems).

Statistical analyses were performed with STATISTICA 8 (StatSoft, Tulsa, Oklahoma, United States). Qualitative variables (AR alleles, discrete clinical data) were analyzed with the χ^2 test. Normality of distribution was evaluated with the Shapiro–Wilk test, and distributions of alleles were compared using the Mann–Whitney test. A P value of less than 0.05 was considered significant.

Results RLPs were identified in (CAG)_n and (GGN)_n regions (range (CAG)_n: patients [14–28], controls [15–39]; range (GGN)_n: patients [16–27], controls [17–30]). The (CCN)_n region had a constant length of 8 repeats. We aimed to evaluate whether the RLPs of the AR gene influenced earlier PC onset (age at diagnosis, <55 years vs ≥55 years) and PC prognostic group assignment, including the level of PSA (<10 ng/ml vs ≥10 ng/ml) and pathological characteristics (GS ≤6 vs ≥7; stage: organ confined vs locally advanced / metastatic).

The shorter variants of the (CAG)_n region (≤18 repeats) were nearly 5-fold more frequent among patients with a GS of 6 or lower: (CAG)_n ≤18, 19.44%, compared with patients with a GS of 7 or higher: (CAG)_n ≤18, 4.26%; $P = 0.017$; power, 0.66 (TABLE 1). There was no significant correlation

between (CAG)_n and (GGN)_n repeats and age at diagnosis, PSA levels, or cancer stage.

Furthermore, to evaluate the association between AR gene RLPs and propensity to PC, allele frequency distributions were compared between patients and controls. The distributions of (CAG)_n and (GGN)_n alleles were comparable in both groups (Supplementary material online, Figures S3 and S4). However, there were significantly more alleles of the (CAG)_n region of 18 repeats or fewer (13.3% vs 5.8%; $P = 0.048$; power, 0.51) and more alleles of the (GGN)_n region of 19 repeats or fewer (10.8% vs 4.2%; $P = 0.049$; power, 0.50) in patients than in controls. To simulate the combined influence of polymorphisms in both regions on the susceptibility to PC, a total repeat count (TRC) was calculated for each participant. The distribution of TRCs was similar in both groups; however, there were more participants with a TRC of 42 or lower among patients than among controls (25% vs 13.3%; $P = 0.021$; power, 0.63).

Discussion We report the first study of (CAG)_n and (GGN)_n RLPs of the AR gene in the context of PC risk and tumor characteristics, conducted among the Caucasian men of Polish origin diagnosed with PC. The unscreened control group included male newborns, from whom umbilical cord blood samples were secured for genetic studies. This study design has not been used so far in the above setting. The influence of the RLPs within the first exon of the AR gene on prostate carcinogenesis remains unclear. We found smaller number of repeats (≤18) within the (CAG)_n region predisposing to PC. This is consistent with the most comprehensive meta-analysis on this subject to date, which involved 4274 patients and 5275 controls, showing that shorter (CAG)_n alleles are correlated with the risk of PC.¹⁰ However, the absolute difference between cases and controls in this study was less than 1 repeat, precluding the use of the CAG repeat number as a marker of PC risk in a clinical setting.¹⁰ Another confounding factor is a broad diversity of cut-off values of repeat number considered “optimal” for the prediction of PC in series published to date.^{7,8,11} We chose the cut-off values identical with those in the study of Mononen et al.⁹ Furthermore, our study recognized a GGN repeat count of 19 or lower as predisposing to PC. This relationship was also confirmed by Zeegers et al.¹⁰; however, as with the (CAG)_n region, there is no consensus on the number of GGN repeats most predictive for PC. The combined influence of both polymorphisms on the risk of PC has been evaluated according to the method of

Silva-Neto et al,⁷ with a total repeat count calculated for each subject. We found a TRC of 42 or lower to be predisposing to PC, while Silva-Neto reported a TRC of 37 or lower. This discrepancy stems not only from different populations studied and diverse cut-off thresholds used for each polymorphism, but also from the lack of association between the GGN repeat number and the PC risk found in the Brazilian study. The combined effect of short alleles of the (CAG)_n and (GGN)_n regions on the intensity of signal transduction through the AR has been recently studied by Rodríguez-González et al,¹¹ who found that patients with (CAG)_n repeats of 21 or fewer and (GGN)_n repeats of less than 23 had the strongest PC tissue staining for AR and PSA proteins, along with a higher GS. In contrast to the above papers, our study revealed that (CAG)_n alleles of 18 repeats or fewer were nearly 5-fold more frequent in patients with a GS of 6 or lower than in patients with high-grade PC. Our findings are consistent with those reported by Madjunkova et al⁸ among the Macedonian patients with PC.

The unscreened control groups, such as newborn cohorts, are frequently used in case-control genetic association studies.¹² In our study, patients with PC and newborn controls were ethnically and geographically matched, and regardless of inherent differences in some demographic characteristics between the groups, this study design is not likely a major limitation. Because of a preliminary character of the study, we did not use multiple testing corrections; therefore, the results should be interpreted with caution.

The (CAG)_n and (GGN)_n RLPs of the AR gene are likely to play a role in the pathogenesis of PC, with shorter variants of (CAG)_n and (GGN)_n microsatellites being more prevalent in patients with PC than in unscreened controls. Despite conflicting data from the literature, our research in Polish men supports the hypothesis that shorter (CAG)_n microsatellite length indicates a favorable tumor grade. Although not sufficiently reliable as a stand-alone molecular marker, the (CAG)_n repeat count could be incorporated into molecular PC risk-stratification instruments along with other validated molecular markers. This will help to reliably allocate patients to prognostic groups and to personalize the approach in counseling the patients on treatment modalities.

Acknowledgments Prof. Paul D. Abel, Imperial College, London, United Kingdom, for his inspiration and expert advice.

Supplementary material online Supplementary material is available with the online version of the article at www.pamw.pl.

REFERENCES

- 1 Dobruch J, Borówka A, Antoniewicz AA, Chłosta P. [Epidemiology of prostate cancer: changes observed in Poland in 1991–2000]. *Urol Pol*. 2005; 58: 26–30. Polish.

- 2 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin*. 2015; 65: 5–29.
- 3 Hugosson J, Carlsson S. Overdetection in screening for prostate cancer. *Curr Opin Urol*. 2014; 24: 256–263.
- 4 Boström PJ, Bjartell AS, Catto JW, et al. Genomic predictors of outcome in prostate cancer. *Eur Urol*. 2015; 68: 1033–1044.
- 5 Ding D, Xu L, Menon M, et al. Effect of GGC (glycine) repeat length polymorphism in the human androgen receptor on androgen action. *Prostate*. 2005; 62: 133–139.
- 6 Ding D, Xu L, Menon M, et al. Effect of a short CAG (glutamine) repeat on human androgen receptor function. *Prostate*. 2004; 58: 23–32.
- 7 Silva Neto B, Koff WJ, Biolchi V, et al. Polymorphic CAG and GGC repeat lengths in the androgen receptor gene and prostate cancer risk: analysis of a Brazilian population. *Cancer Invest*. 2008; 26: 74–80.
- 8 Madjunkova S, Eftimov A, Georgiev V, et al. CAG Repeat Number in the Androgen Receptor Gene and Prostate Cancer. *Balkan J Med Genet*. 2012; 15: 31–36.
- 9 Mononen N, Ikonen T, Autio V, et al. Androgen receptor CAG polymorphism and prostate cancer risk. *Hum Genet*. 2002; 111: 166–171.
- 10 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: 1765–1771.
- 11 Rodríguez-González G, Cabrera S, Ramírez-Moreno R, et al. Short alleles of both GGN and CAG repeats at the exon-1 of the androgen receptor gene are associated to increased PSA staining and a higher Gleason score in human prostatic cancer. *J Steroid Biochem Mol Biol*. 2009; 113: 85–91.
- 12 Moskvina V, Holmans P, Schmidt KM, Craddock N. Design of case-controls studies with unscreened controls. *Ann Hum Genet*. 2005; 69: 566–576.