Journal of Genetics (2023)102:23
https://doi.org/10.1007/s12041-023-01421-1

RESEARCH ARTICLE

Differential expression of the androgen receptor gene is correlated with CAG polymorphic repeats in patients with prostate cancer

MARTÍN IRIGOYEN ARREDONDO[1,2], JOSÉ ROMERO QUINTANA[1], ELIAKYM ARÁMBULA MERAZ[3], FERNANDO BERGEZ HERNÁNDEZ[2], FRED LUQUE ORTEGA[4], ALEJANDRA MARTÍNEZ CAMBEROS[2], MARCO ÁLVAREZ ARRAZOLA[5], ENRIQUE ROMO MARTÍNEZ[7], DORA CEDANO PRIETO[3], JOSÉ CONTRERAS GUTIÉRREZ[6] and NOEMI GARCÍA MAGALLANES[7]*

[1]*Posgrado en Ciencias Biomédicas, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Sinaloa, Calz de las Americas Nte 2771, Cd Universitaria, Burócrata, 80030 Culiacán Rosales, Sinaloa, México*
[2]*Lic. en Ciencias Biomédicas, Universidad Autónoma de Occidente, Av del Mar 1200, Tellería, 82100 Mazatlán, Sinaloa, México*
[3]*Laboratorio de Genética y Biología Molecular, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Sinaloa, Calz de las Americas Nte 2771, Cd Universitaria, Burócrata, 80030 Culiacán Rosales, Sinaloa, México*
[4]*Laboratorio de Ciencias Básicas, Facultad de Odontología, Universidad Autónoma de Sinaloa, C. Josefa Ortiz de Domínguez s/n, Cd Universitaria, Universitaria, 80013 Culiacán Rosales, Sinaloa, México*
[5]*Álvarez & Arrazola Radiólogos, Avenida Insurgentes, 1390, Lopez Mateos, 82140 Mazatlán, Sinaloa, México*
[6]*Centro de Investigación y Docencia en Ciencias de la Salud, Hospital Civil de Culiacán, Universidad Autónoma de Sinaloa, Prol. Álvaro Obregón 1422, Tierra Blanca, 80030 Culiacán Rosales, Sinaloa, México*
[7]*Laboratorio de Biomedicina y Biología Molecular, Ingeniería en Biotecnología, Universidad Politécnica de Sinaloa, Carretera Municipal Libre Mazatlán-Higueras km 3, col. Genaro Estrada, 82199 Mazatlán, Sinaloa, México*
*For correspondence. E-mail: ngarcia@upsin.edu.mx.

Received 24 October 2022; revised 22 December 2022; accepted 8 January 2023

Abstract. Prostate cancer (PCa) is one of the most common types of cancer in men. The aetiology of the disease is not well established, but it has been related to one of the main pathways of regulation of prostate proliferation, mediated by androgens. The androgen receptor (*AR*) gene encodes the androgen-receptor protein, which functions as a transactivation factor for steroid hormones. It has been proposed that the *AR* gene transcription levels are mediated by short tandem repeats corresponding to the CAG sequence. However, there are conflicting results in this relationship. We evaluated the expression levels of the *AR* gene and identified the number of CAG repeats (CAG_n) in the Mexican population, establishing the relationship between expression levels and increase in the number of CAG repeats. We evaluated the expression levels of *AR* in tissue samples of PCa and benign prostate disease, such as benign prostatic hyperplasia, or prostatitis, to determine the difference in their expression levels. Our results showed a statistically insignificant underexpression of 0.64-fold decrease in *AR* levels of PCa patients compared to benign prostate disease patients ($P = 0.623$) and suggest that the number of CAG_n was correlated with the relative expression of the *AR* gene ($P = 0.009$) and this correlation was positive, moderate, and proportional ($\rho = 0.467$) and no correlation was found between CAG_n with other clinical features.

Keywords. androgen receptor; AR pathway; CAG repetitions; prostate pathology; relative expression.

Introduction

Prostate cancer (PCa) is a disease characterized by massive cell proliferation events, the onset, and progression of PCa have been associated with this cell proliferation and deregulation of the cell cycle, caused when one or more DNA damage repair pathways become inefficient, modifying normal cellular pathways, which favours abnormal cell growth (Ben-Salem and Balaji Venkadakrishnan 2021). Accordingly, one of the most characterized pathways that regulate PCa development and growth is that mediated by androgens, which is a ligand-dependent nuclear transcription factor. The

Published online: 27 February 2023

androgen receptor gene (*AR*) encodes the androgen-receptor protein (ARP) that functions as a transcription factor activated by steroid hormones (Davey and Grossmann 2016). This series of transcriptional factors can be activated through ligand binding, with nuclear receptors being the main ligand-dependent transcriptional factors, which modulate the transcription of specific genes (Mitsis *et al*. 2020).

The ARP is controlled by testosterone and 5α-dihydrotestosterone (5α-DHT), which exert their biological effects by binding to ARP. Upon binding to the hormone ligand, the receptor dissociates from accessory proteins, moves to the nucleus, dimerizes, and then binds to the androgenic response element (ARE), to induce transcription of specific androgen-responsive genes that recruit transcription co-activators and co-suppressors (Crumbaker *et al*. 2017; Aurilio *et al*. 2020). *AR* gene expression is maintained throughout the PCa progression and most cases of hormone-refractory or androgen-independent PCa maintain constant *AR* expression (Heinlein and Chang 2004). Mutations in the *AR* gene, especially those that cause loss of specificity for the ARP ligand, contribute to the progression of PCa (Cai *et al*. 2011).

The most characterized mutations in *AR* correspond to two polymorphic trinucleotide repeat segments encoding polyglutamine and polyglycine tracts in the N-terminal transactivation domain of its protein (located on the first exon). The polyglutamine tract is encoded by a trinucleotide CAG repeat (Hsing *et al*. 2000; Gómez *et al*. 2016). This polymorphic site has been associated with receptor activity, where fewer CAG repeats encode for a more active receptor and higher androgenic activity, even with the same amount of androgens (Beilin *et al*. 2000; Hsing *et al*. 2000; Buchanan 2004; Gómez *et al*. 2016). Some studies have established the number of repeats in patients with PCa that range from 12 to 30 CAG repeats, with 22 repeats being the wild-type allele (Nam *et al*. 2000; Paz-Y-Miño *et al*. 2016). In this sense, previous studies have shown that a short CAG repeat length considerably increases the risk of developing PCa in contrast to longer CAG repeat lengths, which seem to result in the reduction of transcriptional activity of the *AR* gene (Hsing *et al*. 2000; Gu *et al*. 2012; Sun and Lee 2013; Gómez *et al*. 2016). Due to the importance of androgenic activity and the establishment of CAG polymorphic variants concerning the development of PCa, and the lack of information for Hispanic populations, it is suggested that the relationship of this polymorphism could elucidate and establish the association between the relative expression of the *AR* gene to the number of CAG repetitions. Therefore, the objective of this work was to establish this association.

Materials and methods

Study population

The study included a total of 71 patients, of which 41 were diagnosed with PCa and 30 were diagnosed with some benign prostate pathology (BPD). Tissue and peripheral blood samples were obtained from the patients from the private clinic Álvarez & Arrazola Radiólogos in Sinaloa, Mexico, collected between August 2019 and October 2021. To participate in the study, the patients filled out an evaluation questionnaire and signed a letter of informed consent. The study design was evaluated by the ethical committee of the corresponding medical institution.

Histopathological analysis

The tissue samples were obtained through ultrasound-guided biopsies and analysed by a pathologist, where a complete report of the morphological characteristics of the tumour (cellular differentiation) was made. The results were reported according to the Gleason score and were classified by the International Society of Urological Pathology grade (ISUP).

DNA extraction

Peripheral blood samples were obtained by venipuncture and stored in tubes with EDTA anticoagulant from both study groups. From the collected samples, DNA extraction from whole blood was performed following the Gustincich method (Gustincich *et al*. 1991). DNA concentration and integrity were evaluated in a spectrophotometer at a wavelength of 260 nm and 280 nm. DNA from peripheral blood was used to perform fragment analysis.

Fragment analysis

An injection mixture of 1 μL of PCR product, 0.25 μL of standard size GS500 ROX, and 12 μL of Hi-Di formamide (Applied Biosystems, Foster City, USA) was prepared. Following denaturation with cycles of 93°C for 2 min and ice for 2 min, the mixture was injected into an ABI PRISM 310 Genetic Analyzer (from Applied Biosystems) and run under the following conditions: dye primer FAM4, capillary 30 cm, injection voltage at 15 kV, injection time 5 s, run time 25 min, filter set F and POP4 polymer. The size of the wild-type CAG repeat (224 bp) was taken as a reference to discriminate the number of repetitions of all samples.

RNA extraction and reverse transcription

Total RNA extraction from tissue samples was performed using the miRNEasy kit (Qiagen, Hilden, Germany), following the supplier's recommendations. Subsequently, RNA quantification was performed to check the concentration and purity index employing a UV-visible spectrophotometer (Thermo Scientific). cDNA synthesis was performed from the RNA samples using the TaqMan Advanced miRNA cDNA Synthesis kit (Applied Biosystems). The RNA from

the tissue was used to perform the relative quantification by real-time PCR.

Real-time PCR relative expression

Relative expression of the *AR* gene was quantified by real-time PCR (StepOneplus system of Applied Biosystems), using FAM-labeled Taqman® probes (Applied Biosystems, Foster City, USA) and using GAPDH as normalizer. Relative expression data were obtained using the method proposed by Taylor *et al*. (2019), which uses the ΔΔCt method as a basis (Livak and Schmittgen 2001), and quantitative analyses of relative expression were subsequently performed using the $\log^2$ transformed normalized expression.

Statistical analysis

Statistical analysis was performed with The Statistical Package for Social Sciences for Windows (v. 20.0; SPSS). The normality of the data was checked using the Kolmogorov–Smirnov test. For the analysis of the mean differences of various continuous variables between the PCa and BPD groups, the statistical test Mann–Whitney U, Kruskal Wallis, and ANOVA were applied, which were applied when they correspond to the type of distribution presented by the data. To determine the correlations between the variables, the Pearson and Spearman correlation coefficients were used, corresponding to each case. To evaluate the relative impact of the clinical characteristics, a predictive model was carried out based on simple linear regression analysis, using the relative expression of *AR* as the dependent variable.

Results

Clinicopathological characteristics

In this study, all the patients in the PCa group were diagnosed with prostate adenocarcinoma. The average age of PCa patients was 66.55 ± 12.13 years, and the average age of BPD group was 62.45 ± 6.94 years (P = 0.008). According to the evaluation of the body mass index (BMI), it was determined that the average age of the patients in the PCa group was 24.68 ± 9.16 compared to the BPD group with an average of 25.58 ± 3.01 (P = 0.738). The analysis of prostate-specific antigen (PSA) in patients of the PCa group showed a mean of 65.92 ± 203.19 ng/mL in contrast to the BPD group with a mean of 68.29 ± 178.94 ng/mL, a statistically insignificant difference was observed between the groups (P = 0.945). While performing the histopathological analysis in the PCa group, eight different Gleason scores were identified: 6 (3 + 3), 7 (3 + 4), 7 (4 + 3), 8 (4 + 4), 8 (3 + 5), 8 (5 + 3), 9 (4 + 5) and 9 (5 + 4) with a frequency of 4.88%, 31.71%, 31.71%, 2.44%, 4.88%, 2.44%, 7.32% and 14.63%, respectively. According to the ISUP grade, the following frequencies were observed: ISUP 1 as 4.88%, ISUP 2 as 31.71%, ISUP 3 as 31.71%, ISUP 4 as 9.76%, and ISUP 5 as 21.95%. A statistical difference was observed between the ages of the ISUP grades (P = 0.042), likewise, it was determined that there were differences between the weights of the patients with PCa between the different ISUP grades (P = 0.038).

Fragment analysis STR-CAG

Regarding fragment analysis, 11 alleles were identified corresponding to 13, 15, 16, 17, 17, 18, 19, 20, 21, 22, 23 and 25 CAG repeats. The 22 CAG repeat is the wild type and has a size of 224 bp (figure 1). The distribution of CAG repeats (CAG_n) in the PCa and BPD groups is shown in table 1. The CAG_{16}, CAG_{22}, and CAG_{23} are the most common in the PCa group with 15.8% and CAG_{19} at 15.8% in the BPD. According to the wild type (CAG_{22}), the CAG repetitions were classified into two groups ≤22 and >22 repetitions to proceed with the corresponding analyses. The analysis of the distribution of age, weight, relative *AR* expression, PSA, and BMI, demonstrated that there were no differences between the distribution of these variables between the groups ≤22 and >22 CAG repetition (table 2). Likewise, no association was found between the CAG_n in the presence of PCa (P = 0.16) and the ISUP grade for aggressiveness (P = 0.215).

Relative AR expression, variable correlation, and linear regression analysis

It was observed that the *AR* gene is underexpressed in patients of the PCa group, showing a change factor of 0.64 in contrast to the BPD group, and a statistically insignificant difference was observed (P = 0.155). To establish the different relationships between the expression of the study genes concerning the clinicopathological variables, correlation analyses were performed as described above. The aim was to identify the clinicopathological variables that, along with gene expression levels, could be used as a basis for establishing a predictive model. When performing the analysis, we observed that the expression of the *AR* is correlated with the BMI, and PSA levels being statistically significant. In contrast to other critical clinical characteristics such as ISUP grade, age, and weight, no statistical significance was observed. Furthermore, the number of CAG_n repetitions was correlated with the relative expression of the *AR* gene (P = 0.009) and this correlation was positive, moderate, and proportional (ρ = 0.467) and no correlation was found between CAG_n with other clinical features (figure 2).

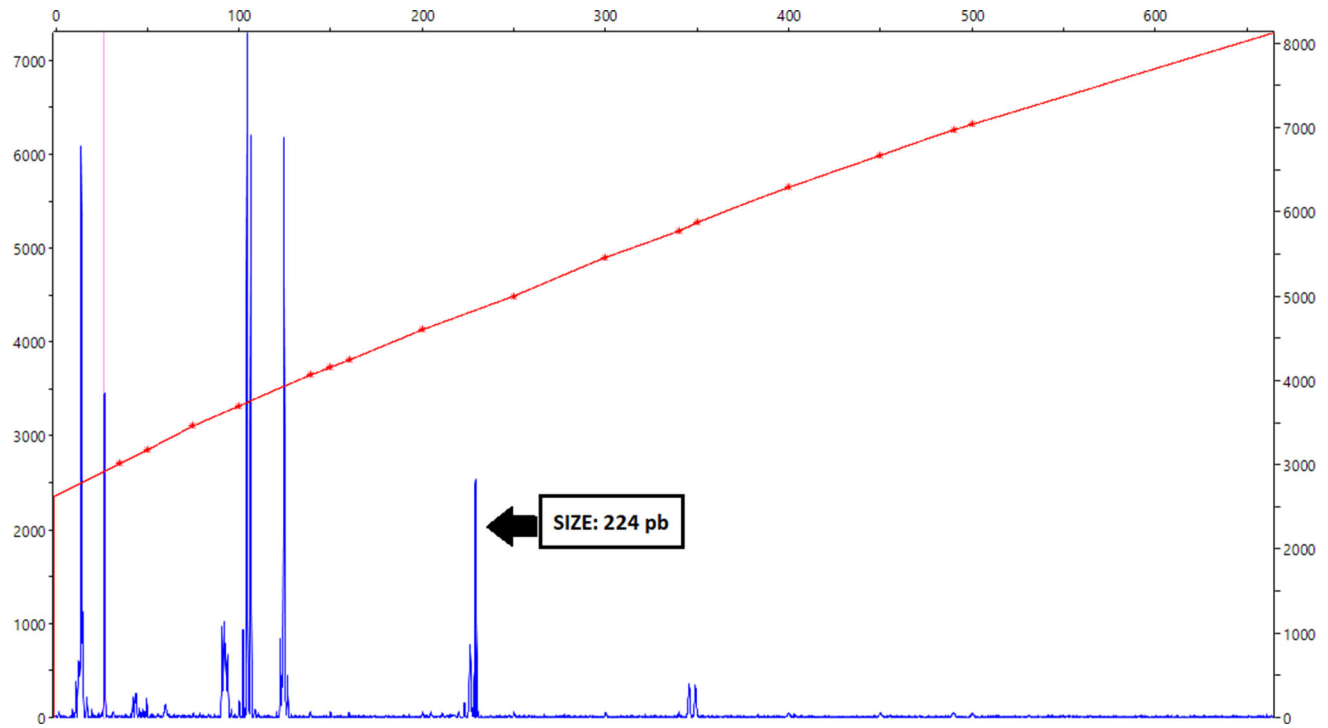

Figure 1. Fragment analysis of $(CAG)_n$. On the *x*-axis, the corresponding fragment size is observed, and on the *y*-axis, the height of each peak is shown. The size of 224 bp corresponds to allele 22 (wild type).

Table 1. Allele distribution of CAG_n in PCa and BPD groups.

	CAG_n distribution	
CAG_n	PCa group (%)	BPD group (%)
CAG_{13}	0.0	11.1
CAG_{15}	5.3	0.0
CAG_{16}	15.8	0.0
CAG_{17}	10.5	11.1
CAG_{18}	10.5	0.0
CAG_{19}	5.3	33.3
CAG_{20}	0.0	22.2
CAG_{21}	10.5	11.1
CAG_{22}	15.8	11.1
CAG_{23}	15.8	0.0
CAG_{25}	10.5	0.0

Table 2. Distribution of means and standard deviation of clinicopathological variables, ISUP grade, and PCa diagnosis frequency in CAG groups.

	CAG groups		
Variables	≤22	>22	*P*-value
PSA	40.51 ± 58.17	43.23 ± 63.41	0.426
Age	68.25 ± 9.27	69.38 ± 11.44	0.599
Weight	74.00 ± 3.92	76.13 ± 12.30	0.448
BMI	25.45 ± 1.10	26.70 ± 2.58	0.776
AR $\log^2$	− 1.17 ± 2.32	− 0.72 ± 1.88	0.623
	CAG distribution		
ISUP grade	≤22	>22	*P*-value[¥]
1	0%	4%	0.165
2	14%	7%	
3	4%	21%	
4	0%	0%	
5	11%	7%	
Study group			
PCa	29%	39%	0.215
BPD	4%	29%	

[*]Variable units, PSA (ng/mL); age (years); weight (kg); BMI (kg/m^2); *AR* $\log^2$ (fold-change).
P-value, Kruskal Wallis test.
P-value, [¥]Pearson's chi-squared test.

According to simple linear regression, values such as the R-square (r^2), intercept, slope, and standard error of estimation (Sy.x) were calculated from the equation of the straight line for the independent variables: PSA, BMI, and CAG_n. The model built for the serum PSA levels and the *AR* relative expression (r^2= 0.003), it was determined that for every 1 ng/mL of serum PSA, the relative expression of the *AR* mRNA would have an increase of 0.003-fold change, with an estimation error of ± 1.85. For the BMI, for every 1 kg/m^2 the *AR* expression decreases 0.50-fold change, with an estimation error of ± 2.00. In the case of CAG_n, for each repetition, the expression of the *AR* increases 0.081-fold change, with an estimation error of ± 1.93 (table 3).

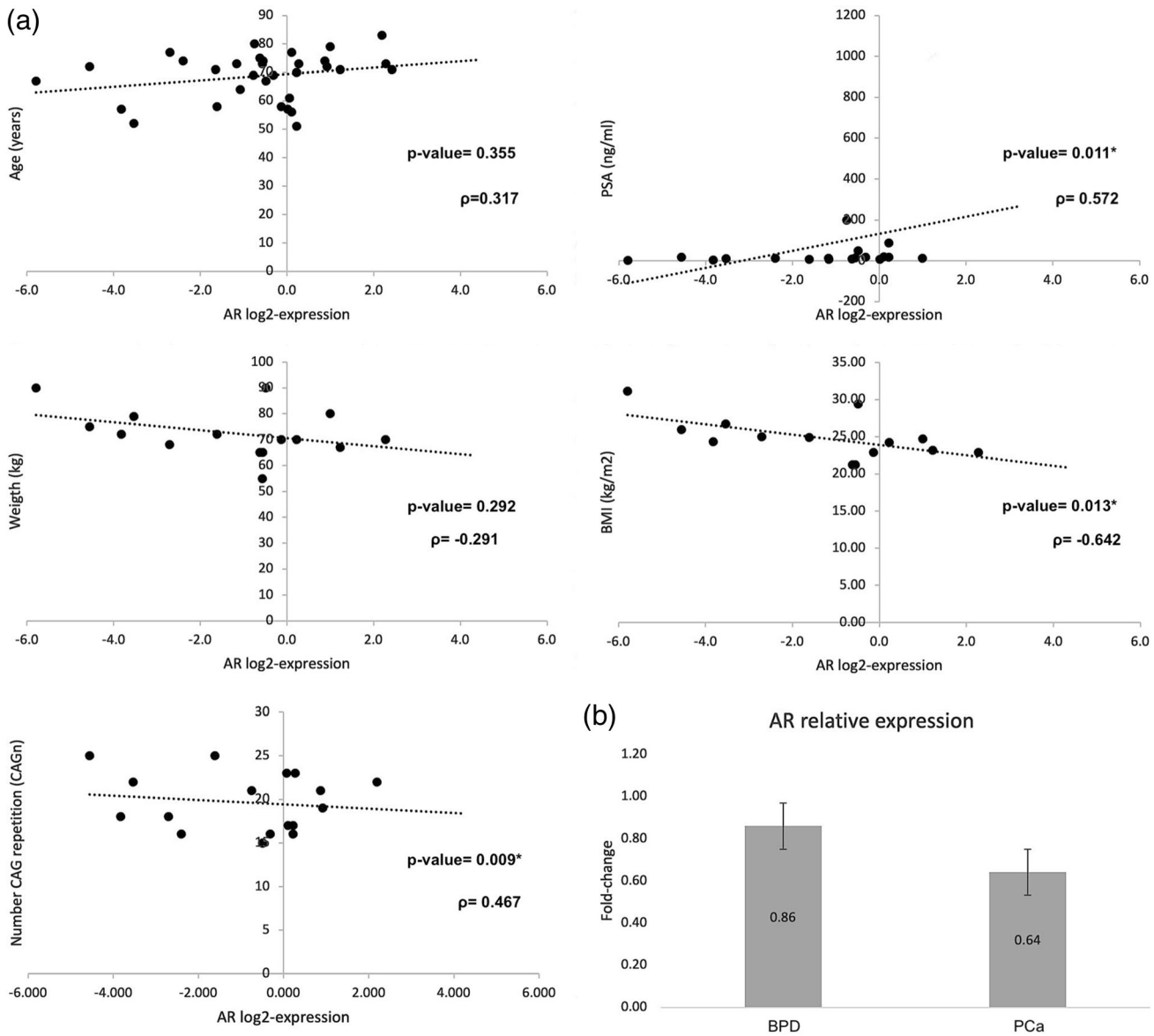

Figure 2. (a) Correlation analysis between the relative expression of the *AR* gene and clinicopathological characteristics. (b) Comparative analysis of the relative expression of the *AR* gene between the study groups.

Discussion

The aetiological processes that lead to the development of PCa are not well established. It is a multifactorial disease but with a high molecular component, important for its development. We focussed our efforts on two molecular components: gene expression, which is an essential complex process for cells, playing a fundamental role in converting the specific sequences of a gene into functional products (Mitsis *et al*. 2020), and trinucleotide repeats which have been highly associated with the development of different human pathologies due to microsatellite instability (Albà and Guigó 2004; Ellegren 2004; Paz-Y-Miño *et al*. 2016). Therefore, the analysis of gene expression and associated mutations is considered pertinent to better understand the aetiology of this pathology. *AR* gene expression and CAG_n polymorphic repeats are proposed to be involved in the prostate growth regulation pathway, such as the androgen-mediated signalling pathway. The variants reported in the CAG_n are highly polymorphic and associated with ethnic factors (Hsing *et al*. 2000), thus, it may be important to determine their association with PCa in different populations. Our results show that in the Mexican population, the distribution of CAG_n genotypes ranged from 13 to 25 repeats, which is consistent with previous studies where the number of repeats was characterized in Mexican, African, Caucasic, and Asian populations (Edwards *et al*. 1999; Hsing *et al*. 2000; Patiño-García *et al*. 2007; Paz-Y-Miño *et al*. 2016). In our study, the relationship between CAG_n and the risk of developing PCa was determined, however, no

Table 3. Simple linear regression analysis between clinicopathological characteristics and relative expression of the *AR* gene.

	Regression statistics			Regression parameter	
Estimation	Correlation coefficient r	R-square (r^2)	Sy.x	Intercept	Slope
PSA-ARlog2	0.57	0.33	1.85	−1.50	0.003
BMI-ARlog2	−0.64	0.41	2.00	11.07	−0.500
CAGn-ARlog2	0.47	0.22	1.93	0.70	−0.080

association was found. This agrees with a study carried out in a French-German population (Correa-Cerro *et al*. 1999), a study carried out by Patiño-Garcia, where their results showed that in Mexican men there was no significant difference in the *AR* polymorphism in patients with PCa compared to controls (Patiño-García *et al*. 2007), and other studies supported these results (Edwards *et al*. 1999; Lindström *et al*. 2010; Price *et al*. 2010). However, there is contradictory evidence, most of the studies showed a positive association between CAG_n and PCa (Hsing *et al*. 2000; Paz-Y-Miño *et al*. 2016). Nevertheless, the large population frequency of short CAG_n suggests that, if causal, variants in CAG lengths could have substantial public health implications (Nelson and Witte 2002).

Two groups of CAG_n repetitions were established, based on studies that define the risk alleles (Edwards *et al*. 1999; Hsing *et al*. 2000; Patiño-García *et al*. 2007; Lindström *et al*. 2010; Nenonen *et al*. 2010; Price *et al*. 2010; Paz-Y-Miño *et al*. 2016), being the group > 22 the one that is reported as a risk allele. Therefore, the following groups were established; > 22 and ≤ 22. An analysis of clinicopathological variables such as age, PSA levels, and weight was performed to establish differences in distribution means between the > 22 and ≤ 22 groups. Our evidence reveals that the distribution of age, PSA levels, and weight was the same among the established categories. The results agree with Xu *et al*. (2002), showing no differences in serum PSA levels between > 22 and ≤ 22. On the other hand, our reports are similar to other studies, such as the one carried out by Gsur *et al*. (2002), where they found that there was no difference between the distribution of PSA levels in three categories of CAG_n (6–20, 21–22, 23–32). In contrast, Xue *et al*. (2001), showed that higher PSA levels in healthy men were accompanied by shorter CAG_n repeats. Nevertheless, these differences in study designs and analyses make these studies difficult to compare. We found no evidence in the literature that assesses age and weight among the established categories.

Regarding ISUP grade and BMI, no association was found with CAG_n. Several studies have found a lack of association between tumour aggressiveness (Gleason score) and CAG_n (Correa-Cerro *et al*. 1999; Edwards *et al*. 1999; Gsur *et al*. 2002; Lindström *et al*. 2010). This contradicts other studies, where the results suggest that shorter CAG repeat length may be involved not only with the development of prostate cancer but also in the potential aggressiveness of the disease (Edwards *et al*. 1999; Nelson and Witte 2002; Paz-Y-Miño *et al*. 2016). Something important to mention is that the association of the repetitions with the aggressiveness of the tumour was not found in the literature using the ISUP grade classification, which allows our comparisons to be blinded by this factor.

In the present work, we observed that the *AR* gene was underexpressed in PCa patients when compared with BPD patients, and statistical insignificance was found. Moreover, 100% of the patients in the PCa group were diagnosed with prostate adenocarcinoma, which occurs in glandular cells, components of the prostatic epithelium. The constitution is made up of various cell types, including stromal and epithelial cells. Stromal cells have been shown to express *AR* and to be involved in prostate development and prostate carcinogenesis. Further, before and during the development of the prostate, the expression of *AR* in the prostatic epithelium is undetectable and at the same time, stromal cells express high levels of *AR*, and *AR* signalling in stromal cells targets different genes compared to epithelial cells (Aurilio *et al*. 2020). According to the literature, the *AR* signalling pathway is required for both development of the normal prostate gland and PCa. The overexpression of *the AR* gene has been highly associated with castration-resistant PCa (CRPC), overexpressed in $\sim$80–90% of all CRPCs and are hypothesized to be hypersensitive to androgens, rather than androgen-independent. Thus, overexpression of *AR* enhances receptor binding to chromatin in the presence of low androgen concentrations (Urbanucci *et al*. 2013).

Although there were no differences between the expression levels of the *AR* gene between the groups, we observed that there was a correlation between the expression levels concerning the PSA and the BMI. Determining a moderate positive and negative correlation, respectively. This association with PSA levels can be explained because AR stimulates the transcription of genes by binding androgen-response elements in promoters of target genes. One of these androgen-regulated target genes is the *PSA* gene. The expression of *PSA* is mainly induced by androgens and regulated by the AR at the transcriptional level (Gsur *et al*. 2002; Kim and Coetzee 2004). Regarding BMI, this correlation can be explained because obesity is accompanied by changes in the plasma levels of certain hormones and changes in their secretion and/or clearance patterns. In men, obesity and overweight are characterized by a progressive

decrease in testosterone levels with increasing body weight (Pasquali 2006; Álvarez-Castro *et al*. 2011).

Moreover, a correlation was found between the expression levels of the *AR* with the CAG_n (treated for this analysis as a quantitative variable), this association is being reported first time using a relative expression, which to our knowledge is the first study in Mexico. We observed a moderate positive correlation between the relative expression levels of the gene and the CAG_n. As mentioned above, there are contradictory results about the relationship between the number of repetitions for the development of the disease, in most studies a negative linear association between *AR* protein function and the CAG repeat numbers is generally assumed (Hsing *et al*. 2000; Nenonen *et al*. 2010; Paz-Y-Miño *et al*. 2016), from the prognostic value of *AR* expression levels determined by immunohistochemistry (IHC), suggesting that short repetitions CAG_n may impose higher transactivation on the receptor and have an increased binding affinity for androgens, due to the inverse relationship between the number of glutamine residues in the polyglutamine tract and transcriptional activity. This may make the prostate more vulnerable to chronic androgen over-stimulation and increased proliferative activity, which, in turn, could increase the rate of somatic mutations among tumour suppressor genes (Nelson and Witte 2002; Nenonen *et al*. 2010). This difference between the established relationships could be explained due to the different types of analysis performed, where in all previously reported studies, the expression analysis is performed at the protein level, compared to our study, where the mRNA levels of the gene were analysed.

Finally, the estimators in the simple linear regression models have a low influence on the expression levels of the *AR*, so it is suggested that other factors could be related. Regression line models may be used to give predicted values; however, care must be taken not to make predictions outside the ranges (minimum and maximum) that were established within the model. On the other hand, correlation and regression analyses do not infer causality, and more rigorous analyses are required if a causal inference is to be made.

Conclusion

The present study in a Mexican population is the first to analyse the expression of the *AR* gene in combination with CAG_n repeats in patients with PCa. The association of the expression levels of the *AR* gene with the number of CAG_n repetitions was revealed, this being directly proportional. In addition, the expression of *AR* was associated with high-value clinical characteristics such as PSA and BMI. We built a simple predictive model, based on simple line regression, however, it is suggested that this model has to be performed in a larger study.

Acknowledgements

This work was supported by the Universidad Autónoma de Sinaloa and the Universidad Politécnica de Sinaloa. We thank Fernando Coronado Nieves and Yolanda Galvez Vazquez for their help during the collection of samples, as well as Andrea Ross Orozco, for her help during the experiments.

References

Albà M. M. and Guigó R. 2004 Comparative analysis of amino acid repeats in rodents and humans. *Genome Res.* **14**, 549–554.

Álvarez-Castro P., Sangiao-Alvarellos S., Brandón-Sandá I. and Cordido F. 2011 Función endocrina en la obesidad. *Endocrinol. Nutr.* **58**, 422–432.

Aurilio G., Cimadamore A., Mazzucchelli R., Lopez-Beltran A., Verri E., Scarpelli M. *et al*. 2020 Androgen receptor signaling pathway in prostate cancer: from genetics to clinical applications. *Cells* **9**, 2653.

Beilin J., Ball E., Favaloro J. and Zajac J. 2000 Effect of the androgen receptor CAG repeat polymorphism on transcriptional activity: specificity in prostate and non-prostate cell lines. *J. Mol. Endocrinol.* **25**, 85–96.

Ben-Salem S. and Balaji Venkadakrishnan V. 2021 Novel insights in cell cycle dysregulation during prostate cancer progression. *Endocr. Relat. Cancer* **28**, 141–155.

Buchanan G. 2004 Structural and functional consequences of glutamine tract variation in the androgen receptor. *Hum. Mol. Genet.* **13**, 1677–1692.

Cai C., He H. H., Sen Chen, Coleman I., Wang H., Fang Z. *et al*. 2011 Androgen receptor gene expression in prostate cancer is directly suppressed by the androgen receptor through recruitment of lysine-specific demethylase 1. *Cancer Cell* **20**, 457–471.

Correa-Cerro L., Wöhr G., Häussler J., Berthon P., Drelon E., Mangin P. *et al*. 1999 (CAG)nCAA 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.

Crumbaker M., Khoja L. and Joshua A. 2017 AR Signaling and the PI3K pathway in prostate cancer. *Cancers* **9**, 34.

Davey R. A. and Grossmann M. 2016 Androgen receptor structure, function and biology: from bench to bedside. *Clin. Biochem. Rev.* **37**, 3–15.

Edwards S. M., Badzioch M. D., Minter R., Hamoudi R., Collins N., Ardern-Jones A. *et al*. 1999 Androgen receptor polymorphisms: association with prostate cancer risk, relapse and overall survival. *Int. J. Cancer* **84**, 458–465.

Ellegren H. 2004 Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* **5**, 435–445.

Gómez R., Torres-Sánchez L., Camacho-Mejorado R., Burguete-García A. I., Vázquez-Salas R. A., Martínez-Nava G. A. *et al*. 2016 Androgen receptor CAG polymorphism and sporadic and early-onset prostate cancer among Mexican men. *J. Hum. Genet.* **61**, 781–786.

Gsur A., Preyer M., Haidinger G., Zidek T., Madersbacher S., Schatzl G. *et al*. 2002 Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis* **23**, 1647–1651.

Gu M., Dong X., Zhang X. and Niu W. 2012 The CAG repeat polymorphism of androgen receptor gene and prostate cancer: a meta-analysis. *Mol. Biol. Rep.* **39**, 2615–2624.

Gustincich S., Manfioletti G., Del Sal G., Schneider C. and Carninci P. 1991 A fast method for high-quality genomic DNA

extraction from whole human blood. *BioTechniques* **11(298–300)**, 302.

Heinlein C. A. and Chang C. 2004 Androgen receptor in prostate cancer. *Endocr. Rev.* **25**, 276–308.

Hsing A. W., Gao Y. T., Wu G., Wang X., Deng J., Chen Y. L. *et al*. 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.

Kim J. and Coetzee G. A. 2004 Prostate specific antigen gene regulation by androgen receptor. *J. Cell. Biochem.* **93**, 233–241.

Lindström S., Ma J., Altshuler D., Giovannucci E., Riboli E., Albanes D. *et al*. 2010 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. *J. Clin. Endocrinol. Metab.* **95**, 121–127.

Livak K. J. and Schmittgen T. D. 2001 Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* **25**, 402–408.

Mitsis T., Efthimiadou A., Bacopoulou F., Vlachakis D., Chrousos G. and Eliopoulos E. 2020 Transcription factors and evolution: an integral part of gene expression (Review). *World Acad. Sci. J.* **2**, 3–8.

Nam R. K., Elhaji Y., Krahn M. D., Hakimi J., Ho M., Chu W. *et al*. 2000 Significance of the CAG repeat polymorphism of the androgen receptor gene in prostate cancer progression. *J. Urol.* **164**, 567–572.

Nelson K. A. and Witte J. S. 2002 Androgen receptor CAG repeats and prostate cancer. *Am. J. Epidemiol.* **155**, 883–890.

Nenonen H., Bjork C., Skjaerpe P. A., Giwercman A., Rylander L., Svartberg J. *et al*. 2010 CAG repeat number is not inversely associated with androgen receptor activity in vitro. *Mol. Hum. Reprod.* **16**, 153–157.

Pasquali R. 2006 Obesity and androgens: facts and perspectives. *Fertil. Steril.* **85**, 1319–1340.

Patiño-García B., Arroyo C., Rangel-Villalobos H., Soto-Vega E., Velarde-Félix J. S., Gabilondo F. *et al*. 2007 Association between polymorphisms of the androgen and vitamin D receptor genes with prostate cancer risk in a Mexican population. *Rev. Investig. Clin. Organo Hosp. Enfermedades Nutr.* **59**, 25–31.

Paz-Y-Miño C., Robles P., Salazar C., Leone P. E., García-Cárdenas J. M., Naranjo M. *et al*. 2016 Positive association of the androgen receptor CAG repeat length polymorphism with the risk of prostate cancer. *Mol. Med. Rep.* **14**, 1791–1798.

Price D. K., Chau C. H., Till C., Goodman P. J., Baum C. E. and Ockers S. B. 2010 Androgen receptor CAG repeat length and association with prostate cancer risk: results from the prostate cancer prevention trial. *J. Urol.* **184**, 2297–2302.

Sun J. H. and Lee S. A. 2013 Association between CAG repeat polymorphisms and the risk of prostate cancer: a meta-analysis by race, study design and the number of (CAG)n repeat polymorphisms. *Int. J. Mol. Med.* **32**, 1195–1203.

Taylor S. C., Nadeau K., Abbasi M., Lachance C., Nguyen M. and Fenrich J. 2019 The ultimate qPCR experiment: producing publication quality, reproducible data the first time. *Trends Biotechnol.* **37**, 761–774.

Urbanucci A., Waltering K. K., Mills I. G. and Visakorpi T. 2013 The effect of AR overexpression on androgen signaling in prostate cancer. In *Androgen-responsive genes in prostate cancer* (ed. Z. Wang), Springer, New York, pp. 187–200..

Xu J., Meyers D. A., Sterling D. A., Zheng S. L., Catalona W. J., Cramer S. D. *et al*. 2002 Association studies of serum prostate-specific antigen levels and the genetic polymorphisms at the androgen receptor and prostate-specific antigen genes. Cancer Epidemiol. *Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **11**, 664–669.

Xue W. M., Coetzee G. A., Ross R. K., Irvine R., Kolonel L. and Henderson B. E. 2001 Genetic determinants of serum prostate-specific antigen levels in healthy men from a multiethnic cohort. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored. Am. Soc. Prev. Oncol.* **10**, 575–579.

Corresponding editor: Durgadas P. Kasbekar