

ORIGINAL ARTICLE

Androgen receptor CAG polymorphism and sporadic and early-onset prostate cancer among Mexican men

Rocío Gómez^{1,6}, Luisa Torres-Sánchez^{2,6}, Rafael Camacho-Mejorado¹, Ana I Burguete-García³, Ruth Argelia Vázquez-Salas², Gabriela A Martínez-Nava^{3,4}, Carla Santana⁵ and Gino Noris⁵

A short CAG repeat length in the gene encoding for the androgen receptor (AR) has been associated with prostate cancer (PC) risk and aggressiveness. In Latino men, information on this association is scarce. Hence, the aim of this study was to evaluate this association in Mexican males. Using fragment analysis by capillary electrophoresis, we determined the number of CAG repeats—(CAG)_n—in AR gene from 158 incident PC cases and 326 age-matched healthy controls (± 5 years), residing in Mexico City, Mexico. According to Gleason scale and age at diagnosis, cases were classified as high (≥ 7) and low grade (< 7), as well as early onset (< 60 years) or late onset PC (≥ 60 years). At diagnosis, 78% of cases were classified as high-grade and 26.6% as early onset. Men with sporadic (no family history of PC) and early-onset PC presented shorter CAG repeat length than controls (18.6 ± 2.2 vs 19.5 ± 2.5 ; $P = 0.02$). Lower number of CAG repeats (CAG) _{≤ 19} were associated with a greater risk for early-onset PC (odds ratio: 2.31; 95% confidence interval: 1.14–4.69). CAG repeat length could increase the risk for sporadic and early-onset PC. The best cutoff point for identifying at-risk subjects was (CAG)₁₉. However, further studies are necessary to replicate our findings in subjects with a family history of PC and also to evaluate the association between CAG repeats length and disease progression.

Journal of Human Genetics (2016) 61, 781–786; doi:10.1038/jhgc.2016.49; published online 19 May 2016

INTRODUCTION

In Mexico, prostate cancer (PC) is currently the first cause of morbidity (27.3 per 100 000 inhabitants) and mortality (11.3 per 100 000 inhabitants) among males. An estimated 30% of PC cases are diagnosed in subjects prior to the age of 60 years.¹ Interestingly, a high proportion of all cases, at the time of diagnosis, present a Gleason scale ≥ 7 and are classified as high-grade or poorly differentiated cancer.²

Prior studies support that age, ethnicity, and a positive PC family history are important risk factors that contribute to the increase of PC development.³ Studies in twins reinforce that genes have a remarkable influence ($\sim 42\%$) on the development of PC.⁴ Given that prostate cell growth is hormone dependent, genes that participate in the synthesis pathways of steroid hormones have a critical role.⁵ Androgen receptor (AR) is a ligand-dependent transcriptional regulator codified on the X chromosome (q11–q12). This gene comprises eight exons in which different polymorphisms were found. As with other steroid receptors, AR contains a transactivation domain (located on the first exon) that encloses a trinucleotide motif (CAG)_n.⁶ This motif has been related to receptor activity, where a lower number of CAG repeats encode for a more active receptor and greater androgenic activity, even with the same amount of androgens.^{7–8}

Preliminary studies suggest that a short CAG repeat length considerably increases risk for PC.^{9,10} Nonetheless, CAG repeat length associated with PC has been discrepant among populations, suggesting that average CAG repeat length exhibits an inter-ethnic variation.¹¹ In this scenario, Afro-descendants and Caucasians present the shortest CAG repeat length and high PC incidence rates.^{3,12} In contrast, Asian and Hispanic populations have been related to largest CAG repeat length and consequently, at low risk for PC.^{9,10,12} However, in Hispanics, there are only two available studies, with conflicting results that examine CAG repeat length and related risk for PC.^{13,14} In a case-control study carried out in Hispanic residents from San Antonio, TX, USA, fewer than (CAG)₁₈ were associated with a threefold greater probability of PC.¹³ In contrast, another study in Hispanic residents from Los Angeles and Hawaii (USA), where the cutoff points employed were < 22 or < 23 CAG repeats, no association with PC risk or its extension was observed.¹⁴ In order to add to the limited findings among Hispanic populations, our study aimed to evaluate the association between CAG repeat length and the risk of PC in a population-based case-control study carried out with Mexican male residents in Mexico City.

¹Departamento de Toxicología, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV), Mexico DF, México; ²Instituto Nacional de Salud Pública, Centro de Investigación en Salud Poblacional, Cuernavaca, México; ³Instituto Nacional de Salud Pública, Centro de Investigación en Enfermedades Infecciosas, Cuernavaca, México; ⁴Instituto Nacional de Rehabilitación 'Luis Guillermo Ibarra', Laboratorio de Líquido Sinovial, Mexico DF, México and ⁵Laboratorio BIMODI (Biología Molecular Diagnóstica), Santiago de Querétaro, México

⁶These authors contributed equally to this work.

Correspondence: Dr L Torres-Sánchez, Instituto Nacional de Salud Pública (INSP), Centro de Investigación en Salud Poblacional, Av Universidad 655, Col Sta María Ahuacatlán, Cuernavaca, CP 62100 Morelos, México.

E-mail: ltorress@insp.mx

Received 8 February 2016; revised 4 April 2016; accepted 11 April 2016; published online 19 May 2016

MATERIALS AND METHODS

Study subject selection

We analyzed the genetic data of unrelated 158 males with incident PC (cases) and 326 males without a diagnosis of PC (controls) ages 42–94 years, who resided in Mexico City, Mexico, and had no previous history of any other cancer type. Both groups comprised a random sample from a population-based case–control study (402 cases and 805 controls) performed from November 2011 to August 2014.¹⁵ Cases were recruited at three public and three social security hospitals in Mexico City. Controls were males matched by ± 5 years of age with index cases, without a previous report of prostate-specific antigen >4 ng ml⁻¹, nor with malignant PC-related symptoms (that is, dysuria, hematuria, among others). Identification of controls was performed through the master sample framework utilized in National Health Surveys after the case was detected. In agreement with the 2005 National Census of Population and Housing, we selected 33 Basic Geostatistical Areas in Mexico City, each represented by 10 city blocks, which were visited from North to East in order to find a male possessing all of the criteria for consideration as a control. From each household, only one male was included in the study; if the next home belonged to same family, this household was not visited. Males who did not accept to participate in the study responded four questions about: educational level; marital status; birthplace; and length of time living in Mexico City. For the original study, the participation rate between cases and controls was 85.9% and 87.5%, respectively.

Following the STREGA (STrengthening the REporting of Genetic Association Studies) statement, we included the study of genome control (GC) in order to support our findings and diminish false-positive results.¹⁶ GC is a useful tool to depict the genetic architecture of complex populations. The GC was composed by 300 non-related individuals (150 females and 150 males) from a representative group of 1640 unrelated Mexican mestizo individuals from the Central Valley of Mexico previously described by our research group and only persons whose eight great-grandparents were born in Mexico were eligible.¹⁷

This population-based case–control study was conducted in accordance with the principles established by the Declaration of Helsinki and was approved by the Ethics Committee of the Mexican National Institute of Public Health (INSP; CI-980) and by each committee of the participating hospitals. Also, each GC signed an informed consent validated by the Ethics Committee of the Bimodi Research Unit. Through face-to-face interviews from each case and population control, we obtained information about sociodemographic characteristics, smoking habit, and familial history of PC in first-degree relatives. From each case, we obtained information on Gleason scale and histopathological diagnosis.

Sample collection

Peripheral venous blood was collected from all subjects in Vacutainer tubes containing EDTA (Becton Dickinson (BD), Franklin Lakes, NJ, USA). Peripheral blood mononuclear cells were obtained by Ficoll-hypaque density gradients (Hystopaque; Sigma Chemical Co., Sigma-Aldrich, St Louis, MO, USA), and genomic DNA was extracted from peripheral blood mononuclear cells using TRIzol reagent (Thermo Fisher Scientific, Suwanee, GA, USA). Sample concentration and purity was evaluated by Thermo Scientific Nano-Drop 1000 Spectrophotometer (260/280), and DNA integrity was assessed by electrophoresis in 0.8% agarose gels.

Polymorphism analysis

PCR was performed with oligonucleotide primers previously reported by Westberg *et al.*¹⁸ Approximately 10 ng of target DNA was amplified. The reaction was standardized at a 6- μ l total volume, containing 0.05 μ M primers, 1 \times reaction buffer with NH₄SO₂, 20 mM MgCl₂, 200 μ M of each nucleotide (Thermo Fisher Scientific), 1 M betaine (Sigma-Aldrich), and 1 U Taq DNA polymerase (Thermo Fisher Scientific). The thermocycling procedure consisted of 30 cycles of denaturation at 94 °C for 1 min, annealing at 65 °C for 1 min, and extension at 72 °C for 45 s, followed by a final extension step of 10 min at 72 °C. The resulting amplicons were analyzed by capillary electrophoresis on the ABI Prism 3130XL Genetic Analyzer employing GeneMapper ID ver. 3.2 software (Applied Biosystems, Carlsbad, CA, USA).

Statistical analyses

Selected characteristics between cases and controls were compared. Depending on the type of studied variable, Student's *t*-test, χ^2 -test or Fisher exact tests were utilized. Family history of PC in first-degree relatives was measured as 'yes' or 'no'. According to birth state, birthplace was categorized into six regions: Mexico City, South: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, and Yucatán; West-Central: Aguascalientes, Colima, Guanajuato, Jalisco, and Michoacán; East-Central: Hidalgo, State of Mexico, Morelos, Puebla, Querétaro, and Tlaxcala; North: Chihuahua, Coahuila, Durango, San Luis Potosí, Zacatecas, Baja California, Baja California Sur, Sinaloa, Sonora, Nayarit, Nuevo León, and Tamaulipas, and East: Veracruz and Tabasco.

Average CAG repeat length between cases and population controls were compared using the Student's *t*-test, and genetic differences between populations were evaluated by means of an analysis of molecular variance test. On the basis of a search of the previous literature, we categorized CAG repeat length using the following cutoffs: ≤ 18 , 19–25, and > 25 ; < 18 vs ≥ 18 ; < 19 vs ≥ 19 ; < 21 vs ≥ 21 , and < 23 vs ≥ 23 . To evaluate the association between CAG repeat length with cancer aggressiveness and age at diagnosis, we generated two different groups of cases. Based on Gleason scale¹⁹ at diagnosis, we considered as well-differentiated or low-risk, cancers in which the Gleason scale was < 7 ; in contrast, poorly differentiated or high-risk cancers were those with a Gleason ≥ 7 . In relation to age at diagnosis, we categorized cases of < 60 years of age as 'early-onset' vs ≥ 60 years of age as 'elderly-onset'.

Allele and genotype frequencies in all studied groups were estimated using Arlequin v.3.1 software.²⁰ Hardy-Weinberg expectation was calculated by applying Weir and Cockerham F Statistic (FISW&C), using GenéTix ver. 4.05.2 software.²¹

The crude and adjusted associations between AR CAG repeat length and PC risks were estimated using independent, unconditional logistic regression models for each case type. For each model, we employed AR CAG repeat length as a continuous variable and also independently we evaluated the different previously mentioned cutoff points. Age at interview was included as a continuous variable in bivariate and multivariable models. In addition to age, we evaluated, as a potential confounder, smoking history as well as birthplace. Due to the low prevalence of PC familial history among controls, we only report the association between AR CAG repeats and PC, including all individuals and males without a PC familial history. Cases in this latter analysis were considered as sporadic PC (no family history of PC). Sensitivity and specificity in the cutoffs were evaluated using Receiver Operating Characteristic (ROC) curves. All of the analyses were performed utilizing STATA ver. 14.0 statistical software.

RESULTS

At diagnosis, 78.2% of cases were classified as high-grade and poorly differentiated PC (Gleason ≥ 7), while 26.6% were considered early-onset prostate cancer (EO-PC; data not shown in tables). Regarding study design, we did not find significant differences in the average ages between cases and controls (65.2 ± 8.9 vs 64.5 ± 9.4 years; $P = 0.19$). Compared with males born in Mexico City, those born in the East-Central region had a nearly threefold (odds ratio, 2.70; 95% CI, 1.64–4.46; $P = 0.00$) increased risk of being cases. The prevalence of former smokers and familial history of PC was significantly higher among cases than controls (Table 1).

Allele frequencies

Allele frequency distribution of the CAG polymorphism is depicted in Table 2. The most frequent alleles in the three groups were 17, 18, 19, 20 and 21 (average 19.33 ± 2.5 , range 10–29). Overall, the studied groups presented high-diversity patterns, exhibiting at least fifteen different alleles (*k*) in cases group ($k = 15$), whereas the other groups presented $k = 16$ (genome control, GC) and $k = 17$ (population controls). Comparison among populations (cases, controls and GC) suggested that the groups were similar (genetic differences less than 1%, $P \geq 0.25$). The distribution of this polymorphism was in

Table 1 Selected characteristics of the study population according to cases and controls

Characteristics	Cases		Controls		OR ^a	95% CI	P-value
	(n = 158)	(n = 326)					
Age (years)							
Mean ± s.d.	65.2 ± 8.9	64.5 ± 9.4	1.0	0.99; 1.03		0.19	
Range	43–87	42–89					
Birthplace^b (%)							
Mexico City	79 (50.0)	222 (68.3)	1.0	—			
South	13 (8.2)	18 (5.5)	2.04	0.95; 4.39		0.07	
West-Central	11 (7.0)	22 (6.8)	1.42	0.65; 3.08		0.39	
East-Central	43 (27.2)	45 (13.8)	2.70	1.64; 4.46		0.00	
North	4 (2.5)	7 (2.1)	1.60	0.46; 5.71		0.34	
East	8 (5.1)	11 (3.4)	2.04	0.79; 5.26		0.14	
Marital status (%)							
United ^c vs not united	118 (74.68)	258 (79.1)	0.77	0.50; 1.20		0.28	
Educational level (%)							
≤ Elementary school	75 (47.5)	145 (44.5)	1.0	—			
Junior-High school	24 (15.2)	79 (24.2)	0.60	0.35; 1.04		0.09	
High school	30 (19.0)	56 (17.2)	1.07	0.63; 1.83		0.76	
University or more	29 (18.4)	46 (14.1)	1.26	0.73; 2.18		0.35	
Smoking status^d (%)							
Never smoked	51 (32.3)	112 (34.4)	1.0	—			
Former smoker							
> 15 y	34 (21.5)	54 (16.6)	1.36	0.78; 2.35		0.27	
≤ 15 y	36 (22.8)	47 (14.4)	1.70	0.98; 2.93		0.06	
Current smoker	37 (23.4)	113 (34.7)	0.56	0.44; 1.20		0.21	
Familial history of PC (%)							
Yes vs No	18 (11.4)	4 (1.2)	10.35	3.44; 31.14		0.00	

Abbreviations: CI, confidence interval; PC, prostate cancer.

^aOdds ratio (OR) adjusted by age at interview.

^bBirthplace: Mexico City. South: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, and Yucatán; West-Central: Aguascalientes, Colima, Guanajuato, Jalisco, and Michoacán; East-Central: Hidalgo, State of Mexico, Morelos, Puebla, Querétaro, and Tlaxcala; North: Chihuahua, Coahuila, Durango, San Luis Potosí, Zacatecas, Baja California, Baja California Sur, Sinaloa, Sonora, Nayarit, Nuevo León, and Tamaulipas, and East: Veracruz and Tabasco.

^cUnited: common law marriage and marriage.

^dSmoking condition five years before interview.

agreement with Hardy-Weinberg equilibrium ($P > 0.05$), which was obtained from women in the GC.

Compared with population controls, PC cases in general (19.5 ± 2.5 vs 19.0 ± 2.6 ; $P = 0.06$) and PC classified as poorly differentiated at time of diagnosis (19.5 ± 2.5 vs 18.9 ± 2.5 ; $P = 0.06$) had a marginally lower average of repeats. In contrast, EO-PC cases (18.6 ± 2.2 ; $P = 0.02$) presented a significantly lower average of CAG repeats (Figures 1a-c). Birthplace was associated with the number of CAG repeats. Males born in the South region presented on average a greater number of triplets than those born in Mexico City (20.7 ± 3.5 ; $P = 0.003$); in contrast, subjects born in the North region had a marginally lower average of CAG repeats (17.8 ± 2.8 ; $P = 0.07$). The average number of CAG repeats length among those with a familial history of PC was lower (18.7 ± 3.9 ; $P = 0.11$), but not statistically significant (Table 3).

Table 2 Allele frequencies and descriptive statistics parameters regarding the locus studied in cases, controls and genome control

Allele/N	Cases		Controls		Genome Control	
	(n = 158)	(n = 314)			Men (n = 150)	Women (n = 150)
10	—	0.003	—	—	—	0.003
12	0.006	0.003	0.007	—	—	—
13	0.025	0.003	—	—	—	0.017
14	0.013	0.006	0.013	—	0.013	0.030
15	0.019	0.019	0.033	—	0.033	0.053
16	0.051	0.041	0.020	—	0.020	0.063
17	0.133	0.131	0.233	—	0.233	0.147
18	0.190	0.147	0.147	—	0.147	0.143
19	0.165	0.194	0.133	—	0.133	0.127
20	0.139	0.147	0.113	—	0.113	0.140
21	0.120	0.121	0.120	—	0.120	0.103
22	0.063	0.076	0.067	—	0.067	0.053
23	0.025	0.054	0.040	—	0.040	0.057
24	0.006	0.019	0.033	—	0.033	0.027
25	0.038	0.022	0.007	—	0.007	0.023
26	0.006	0.006	0.013	—	0.013	0.010
27	—	—	0.013	—	0.013	0.003
29	—	0.006	0.007	—	0.007	—
Descriptive Statistics						
k	15	17	16	—	16	—
H_o	0.000	0.000	0.000	—	0.000	0.893
H_e	0.875	0.876	0.870	—	0.870	0.896
HWE (P)	≤ 0.0001	≤ 0.0001	≤ 0.0001	—	≤ 0.0001	0.932
F_{IS}	1	1	1	—	1	0.003

Abbreviations: F_{IS} , weir and cockerham F statistics; H_e , expected heterosigosity; H_o , observed heterosigosity; HWE, Hardy-Weinberg expectations P -values; k , number of alleles.

PC and CAG repeat length

In order to assess a linear association between CAG repeat length and PC, we carried out a model employing CAG repeat length as a continuous variable. Overall, for each increase in number of CAG repeats, we observed a reduction in the risk of PC. However, this was only statistically significant for EO-PC (odds ratio, 0.83; 95% CI, 0.71; 0.97; $P = 0.02$). For all PC and for Gleason ≥ 7 , the reduction in the risk associated with the increase of copies was marginally significant (Table 4). On evaluating different cutoff points, solely the cutoff point in 19 repeated sequences (CAG)₁₉ exhibited the best combination of sensitivity (0.48) and specificity (0.70) for identifying the risk condition (data not shown). Lower number of (CAG)₁₉ repeats was significantly associated with a two-fold greater risk of EO-PC (odds ratio, 2.31; 95% CI, 1.14; 4.69; $P = 0.02$) and a marginal increase in the risk of PC in general or for poorly differentiated cancer (Gleason ≥ 7). With others cutoff points, although associated with a greater risk of PC in general, the associations were not statistically significant. In subjects without a familial history of PC, the association observed with EO-PC remained statistically significant and of the same magnitude (Table 4).

DISCUSSION

In agreement with previous reports, this study suggests that in Mexican men, a fewer number of CAG repeats in the gene encoding for the AR might be associated with a greater risk of sporadic and EO-PC. The cutoff point that best identified subjects at risk for presenting

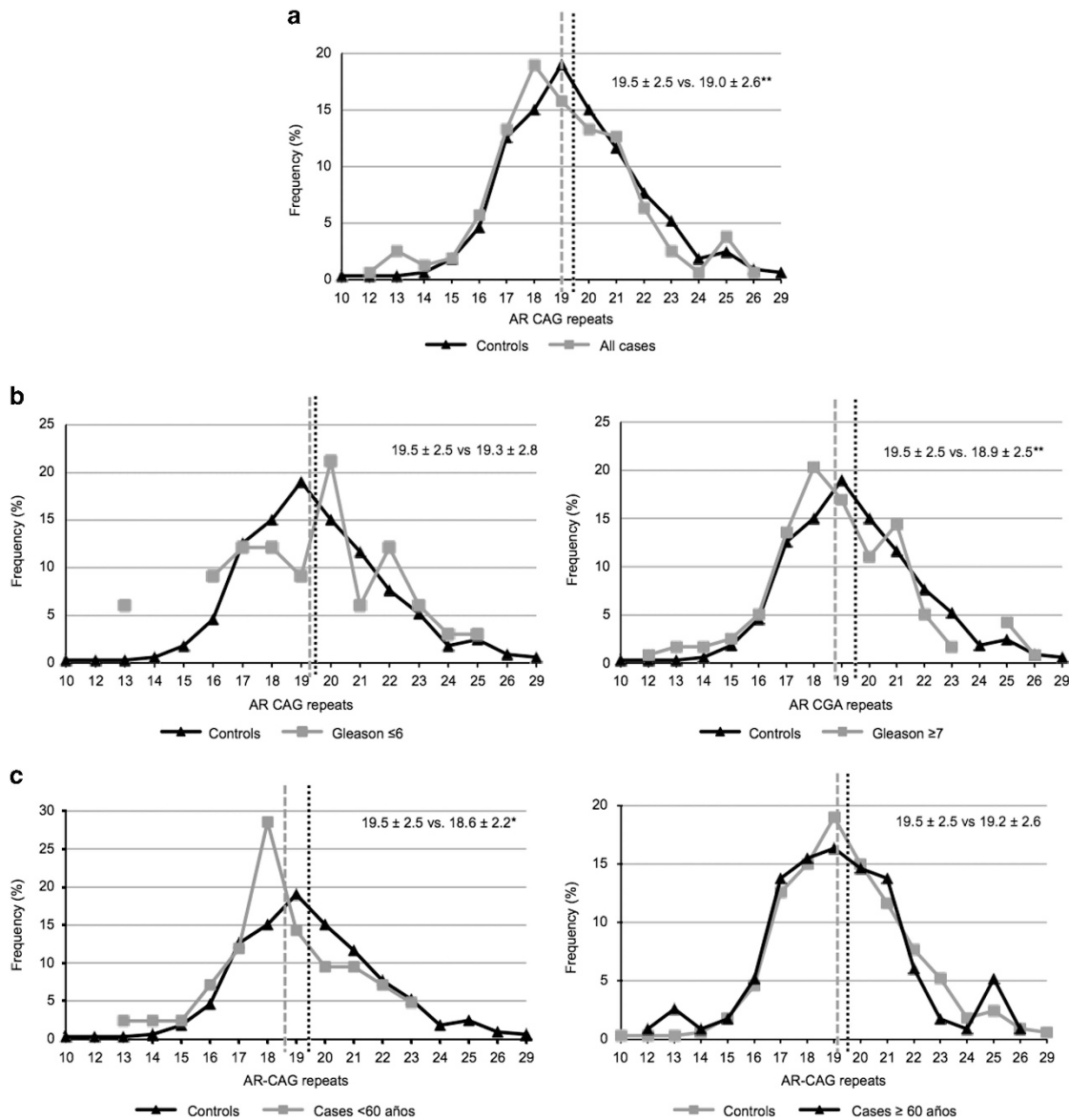


Figure 1 Distribution of CAG repeat length in *AR* gene, according to prostate cancer cases and controls. *Student's *t*-test, *P*-value for cases and controls <0.05; **Student's *t*-test, *P*-value for cases and controls=0.06.

this neoplasm was \leq (CAG)₁₉ repeat length. Likewise, these results suggest that the differential distribution of CAG repeat length across regions of Mexico could be a potential explanation for the regional distribution of PC across the country observed in the Histopathological National Register of Cancer from 1993 to 2002;²² in that study most of the prostate tumors were reported in Mexico City, followed by the northern, center and southern states of the country.

The results of this study are consistent with the majority of studies conducted in Afro-descendent, Caucasian and Asiatic populations,^{9,10} especially with a study carried out in Sweden where they found that short CAG repeats in the *AR* gene correlate with young age at diagnosis of sporadic PC.²³ Moreover, our results are in agreement with the two studies that have evaluated this association in Hispanic population. The study carried out in Hispanic residents of San Antonio, Texas, USA, reported that a CAG repeat length \leq 18 was associated with a three-fold greater risk for PC.¹³ This association was stronger when the analysis was restricted to cases of PC aged <65

years (odds ratio, 3.03; 95% CI, 1.27–7.26). The results of the multiethnic cohort in Hawaii and Los Angeles, CA, USA, were consistent with our results, due to that we did not find a significant association when similar cutoff points to those utilized by the abovementioned authors (<22 and <23 CAG repeats).¹⁴

The main proposed biological mechanism for the association between shorter CAG repeat length and PC risk is the increased receptor activity; however, several studies suggest that CAG repeat length may increase PC risk-mediating androgen-induced *TMPRSS2* and *ERG* proximity.²⁴ About a half of human PCs display a *TMPRSS2-ERG* gene fusion, which is one of the most common genomic aberrations in PC. This gene fusion is related with a poor PC prognosis.²⁵

Comparison between included and not included subjects in this analysis (Supplementary Table S1) showed that a higher proportion of included controls did not have familial history of PC (1.2 vs 3.5%). The external validity of our results is limited to subjects without

Table 3 Androgen receptor (AR) CAG repeats according to selected characteristics of the study population

Characteristics	N = 484	CAG repeats; mean ± s.d.	P-value ^a
<i>Birthplace^b</i>			
Mexico City	301	19.3 ± 2.4	
South	31	20.7 ± 3.5	0.003
West-Central	33	19.3 ± 2.4	0.86
East-Central	88	19.2 ± 2.6	0.91
North	11	17.8 ± 2.8	0.07
East	19	19.6 ± 2.2	0.55
<i>Civil status</i>			
United ^c vs Not united	376	19.3 ± 2.5	0.73
<i>Educational level</i>			
Elementary school	220	19.2 ± 2.4	
Junior-High school	103	19.6 ± 2.6	0.15
High school	86	19.4 ± 2.5	0.64
University or more	75	19.2 ± 2.9	0.89
<i>Smoking status^d</i>			
Never smoked	163	19.5 ± 2.7	
Former smoker			
> 15 years	88	19.7 ± 2.7	0.15
≤ 15 years	83	19.1 ± 2.3	
Current smoker	150	19.1 ± 2.4	
<i>Familial history of prostate cancer (PC) (%)</i>			
No	462	19.3 ± 2.5	
Yes	22	18.7 ± 3.9	0.11

^aAnalysis of variance and Student's *t*-test.

^bBirthplace: Mexico City. South: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, and Yucatán; West-Central: Aguascalientes, Colima, Guanajuato, Jalisco, and Michoacán; East-Central: Hidalgo, State of Mexico, Morelos, Puebla, Querétaro, and Tlaxcala. North: Chihuahua, Coahuila, Durango, San Luis Potosí, Zacatecas, Baja California, Baja California Sur, Sinaloa, Sonora, Nayarit, Nuevo León, and Tamaulipas, and East: Veracruz and Tabasco.

^cUnited: Common law marriage and marriage.

^dSmoking condition five years before interview.

familial history of PC. Therefore, we were not able to establish if the magnitude of this association could be similar among subjects with a family history of PC. Given that we did not find differences related to CAG repeat length between the population controls and GC, the probability of our results to be a consequence of selection bias is low. In addition, although all analysis were adjusted for place of birth, as a proxy measure for genetic ancestry, there is a possibility of having residual confounding. Some evidence suggests that body mass index is related with the risk for PC and the number of CAG repeats. However, we did not include body mass index in our analysis owing to body mass index information that we had was at the time of the interview and it did not adequately represent the subjects' body composition, mainly among cases, which can be affected by the disease. Likewise, we discarded the possibility that the results observed were the consequence of a differential measurement error in the determination of the number of CAG copies, because the person charged with carrying out this determination was unaware of case or control condition. If a measurement error was present, it would be of the non-differential type and would impact the results toward underestimation of the association.

To our knowledge this is the first study that examines the role of the lower number of CAG repeats as a risk factor for PC in Mexican males. In addition, findings from this study may help inform the establishment of cutoff points for determining potential PC risk in this population related to (CAG)₁₉. A previous small study (68 PC cases vs 48 healthy men) carried out in Mexican men suggests an association between shorter CAG repeat length and latter age at PC diagnosis;²⁶ however, some methodological limitations (selection of the study subjects and measurement of CAG repeat length) limit the validity and interpretation of these results. Finally, longitudinal studies are needed to determine and evaluate CAG cutoff points as potential risk factors for developing PC, as well as for disease progression. A high proportion of cancers respond to androgen ablation; however, part of these cancers become resistant to treatment and it is not impossible

Table 4 Androgen receptor (AR) CAG repeats length and its association with prostate cancer (PC) using different cutoff points

Characteristics	CAG continuous		CAG categories											
	OR ^{a,b}	95% CI	N	OR ^b	95% CI	N	OR ^b	95% CI	N	OR ^b	95% CI	N	OR ^b	95% CI
<i>All</i>														
Controls (n = 326)	—	—	67	—	—	116	—	—	227	—	—	290	—	—
<i>Cases</i>														
All (n = 158)	0.93	0.8; 1.00	40	1.30	0.83; 2.03	70	1.44	0.98; 2.11	116	1.20	0.78; 1.83	146	1.51	0.76; 2.98
Gleason < 7 (n = 33)	0.97	0.84; 1.12	9	1.45	0.64; 3.28	13	1.18	0.56; 2.45	23	1.00	0.46; 2.19	29	0.9	0.30; 2.71
Gleason ≥ 7 (n = 118)	0.92	0.84; 1.00	30	1.30	0.80; 2.14	54	1.53	1.00; 2.34	87	1.22	0.76; 1.96	110	1.71	0.77; 3.80
Cases < 60 years (n = 42)	0.83	0.71; 0.97^c	11	1.77	0.77; 4.04	23	2.31	1.14; 4.69^c	33	1.75	0.76; 4.02	40	2.45	0.53; 11.39
Cases ≥ 60 years (n = 116)	0.96	0.89; 1.05	29	1.18	0.71; 1.97	47	1.16	0.74; 1.82	83	1.01	0.63; 1.64	106	1.25	0.60; 2.66
<i>No familial history of prostate cancer</i>														
Controls (n = 322)	—	—	65	—	—	114	—	—	225	—	—	288	—	—
<i>Cases</i>														
All (n = 140)	0.95	0.88; 1.03	32	1.15	0.71; 1.87	58	1.29	0.86; 1.93	102	1.50	0.74; 1.79	129	1.38	0.68; 2.82
Gleason < 7 (n = 30)	1.02	0.88; 1.19	7	1.2	0.49; 2.93	10	0.91	0.41; 2.01	20	0.86	0.39; 1.91	26	0.78	0.25; 2.33
Gleason ≥ 7 (n = 103)	0.93	0.85; 1.02	24	1.18	0.69; 2.02	45	1.41	0.90; 2.22	76	1.21	0.73; 2.00	96	1.63	0.70; 3.79
Cases < 60 years (n = 40)	0.84	0.72; 0.99^c	10	1.62	0.69; 3.80	21	2.31	1.13; 4.69^c	31	1.56	0.67; 3.63	38	2.03	0.43; 9.57
Cases ≥ 60 years (n = 100)	0.99	0.91; 1.09	22	1.00	0.57; 1.76	37	0.99	0.61; 1.61	71	0.97	0.58; 1.61	91	1.15	0.52; 2.54

Abbreviations: CI, confidence interval; OR, odds ratio.

^aPer each increase of one CAG repeat.

^bAll adjusted by age and birthplace.

^cP-value < 0.05.

to think that this lack of response could be associated with the presence of this polymorphism. Finally, it is crucial to establish a National Cancer Registry in Mexico, which could provide a comprehensive repository of data to help guide related research and practice guidelines to reduce the risk of PC morbidity and mortality.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was supported by CONACyT grant number 140482 (to LT-S). We are grateful for the contribution of Arianna Ventura-Bahena, and of Ms Nicol Ventura-Bahena, who participated in the conformation of the samples bank and the coding and entering of the information, respectively. Likewise, we are grateful to each of the participating hospitals: General Hospital of Mexico (SSA), National Cancer Institute (SSA), Salvador Zubirán National Institute of Medical and Nutrition Sciences (SSA), and the Centro Médico Nacional Siglo XXI Oncology Hospital (IMSS), the Dr Carlos McGregor Sánchez Navarro Regional General Hospital No.1 (IMSS) and the Adolfo López Mateos Regional Hospital (ISSSTE). We also thank Laboratorio de Genómica, Proteómica y Metabolómica from LaNSE-Cinvestav for helping with the genotyping processes.

- 1 Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C. *et al*. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 Internet. Lyon, France (2013): International Agency for Research on Cancer. <http://globocon.iarc.fr>. Accessed 13 March 2016.
- 2 Gomez-Guerra, L. S., Martinez-Fierro, M. L., Alcantara-Aragon, V., Ortiz-Lopez, R., Martinez-Villarreal, R. T., Morales-Rodriguez, I. B. *et al*. Population based prostate cancer screening in north Mexico reveals a high prevalence of aggressive tumors in detected cases. *BMC Cancer* **9**, 91–96 (2009).
- 3 Grönberg, H. Prostate cancer epidemiology. *Lancet* **361**, 859–864 (2003).
- 4 Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M. *et al*. Environmental and heritable factors in the causation of cancer- analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78–85 (2000).
- 5 Bosland, M. C. & Mahmoud, A. M. Hormones and prostate carcinogenesis: androgens and estrogens. *J. Carcinog.* **10**, 33 (2011).
- 6 Koochekpour, S. Androgen receptor signaling and mutations in prostate cancer. *Asian J. Androl.* **12**, 639–657 (2010).
- 7 Buchanan, G., Yang, M., Cheong, A., Harris, J. M., Irvine, R. A., Lambert, P. F. *et al*. Structural and functional consequences of glutamine tract variation in the androgen receptor. *Hum. Mol. Genet.* **13**, 1677–1692 (2004).
- 8 Beilin, J., Ball, E. M., Favaloro, J. M. & Zajac, J. D. 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 (2000).
- 9 Gu, M., Dong, X., Zhang, X. & Niu, W. The CAG repeat polymorphism of androgen receptor gene and prostate cancer: a meta-analysis. *Mol. Biol. Rep.* **39**, 2615–2624 (2012).
- 10 Sun, J. H. & Lee, S. A. 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 (2013).
- 11 Williams, H. & Powell, I. J. Epidemiology, pathology, and genetics of prostate cancer among African Americans compared with other ethnicities. *Methods Mol. Biol.* **472**, 439–453 (2009).
- 12 Ackerman, C. M., Lowe, L. P., Lee, H., Hayes, M. G., Dyer, A. R., Metzger, B. E. *et al*. Ethnic variation in allele distribution of the androgen receptor (AR) (CAG)_n repeat. *J. Androl.* **33**, 210–215 (2012).
- 13 Balic, I., Graham, S. T., Troyer, D. A., Higgins, B. A., Pollock, B. H., Johnson-Pais, T. L. *et al*. Androgen receptor length polymorphism associated with prostate cancer risk in Hispanic men. *J. Urol.* **168**, 2245–2248 (2002).
- 14 Freedman, M. L., Pearce, C. L., Penney, K. L., Hirschhorn, J. N., Kolonel, L. N., Henderson, B. E. *et al*. Systematic evaluation of genetic variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study. *Am. J. Hum. Genet.* **76**, 82–90 (2005).
- 15 Vázquez-Salas, R. A., Torres-Sánchez, L., López-Carrillo, L., Romero-Martínez, M., Manzanilla-García, H. A., Cruz-Ortiz, C. H. *et al*. History of gonorrhea and prostate cancer in a population-based case-control study in Mexico. *Cancer Epidemiol.* **40**, 95–101 (2016).
- 16 Silviu-Alin, B., Devlin, B. & Roeder, K. The power of genomic control. *Am. J. Hum. Genet.* **66**, 1933–1944 (2000).
- 17 Noris, G., Santana, C., Meraz-Ríos, M. A., de Lourdes Muñoz, M., Majluf-Cruz, A., Magaña, J. J. *et al*. Mexican mestizo population sub-structure: effects on genetic and forensic statistical parameters. *Mol. Biol. Rep.* **39**, 10139–10156 (2012).
- 18 Westberg, L., Henningsson, S., Landén, M., Annerbrink, K., Melke, J., Nilsson, S. *et al*. Influence of androgen receptor repeat polymorphisms on personality traits in men. *J. Psychiatry Neurosci.* **34**, 205–213 (2009).
- 19 Humphrey, P. A. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod. Pathol.* **17**, 292–306 (2004).
- 20 Excoffier, L., Laval, G. & Schneider, S. Arlequin ver. 3.1: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50 (2005).
- 21 Belkhir, L. C., Raufaste, N. & Bonhomme, F. in *GENETIX 4.05 Logiciel Sous Windows TM Pour la Génétique des Populations*. (eds Laboratoire Génome P) Interactions, CNRS UMR 5171 (Université de Montpellier II, Montpellier, 1996–2004).
- 22 Meneses-García, A., Ruiz-Godoy, L. M., Beltrán-Ortega, A., Sánchez-Cervantes, F., Tapia-Conyer, R. & Mohar, A. Principales neoplasias malignas en México y su distribución geográfica (1993-2002). *Rev. Inv. Clin.* **64**, 322–329 (2012).
- 23 Bratt, O., Borg, A., Kristofferson, U., Lundgren, R., Zhang, Q. X. & 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* **81**, 672–676 (1999).
- 24 Mao, X., Li, J., Xu, X., Boyd, L. K., He, W., Stankiewicz, E. *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.* **4**, 886–896 (2014).
- 25 Yoshimoto, M., Joshua, A. M., Cunha, I. W., Coudry, R. A., Fonseca, F. P., Ludkovski, O. *et al*. Absence of TMPRSS2:ERG fusions and PTEN losses in prostate cancer is associated with a favorable outcome. *Mod. Pathol.* **21**, 1451–1460 (2008).
- 26 Patiño-García, B., Arroyo, C., Rangel-Villalobos, H., Soto-Vega, E., Velarde-Félix, J. S., Gabilondo, F. *et al*. Association between polymorphisms of the androgen and vitamin D receptor genes with prostate cancer risk in a Mexican population. *Rev. Invest. Clin.* **59**, 25–31 (2007).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)