

Allelic frequencies of six polymorphic markers for risk of prostate cancer

M.L. Ribeiro¹,
A. Santos¹,
A.B. Carvalho-Salles³
and C. Hackel^{1,2}

¹Centro de Biologia Molecular e Engenharia Genética, and
²Departamento de Genética Médica, Faculdade de Ciências Médicas,
Universidade Estadual de Campinas, Campinas, SP, Brasil
³Hemocentro, São José do Rio Preto, SP, Brasil

Abstract

The aim of the present study was to evaluate the distribution of polymorphisms for the androgen receptor (*AR*) (CAG, *StuI*, GGN), *SRD5A2* (Ala49Thr, Val89Leu) and *CYP17* (*MspA1*) genes that are considered to be relevant for risk of prostate cancer. We studied 200 individuals from two cities in the State of São Paulo, by PCR, PCR-RFLP and ASOH techniques. The allelic frequencies of the autosomal markers and the *StuI* polymorphism of the *AR* gene were very similar to those described in most North American and European populations. In relation to the CAG and GGN number of repeats, the study subjects had smaller repeat lengths (mean of 20.65 and 22.38, respectively) than those described in North American, European and Chinese populations. In the present study, 30.5% of the individuals had less than 22 CAG repeats and 45.5% had less than 23 GGN repeats. When both repeat lengths are considered jointly, this Brazilian population is remarkably different from the others. Further studies on prostate cancer patients need to be conducted to assess the significance of these markers in the Brazilian population.

Key words

- Prostate cancer
- Polymorphisms
- *AR*
- *SRD5A2*
- *CYP17*

Correspondence

C. Hackel
Departamento de Genética Médica
FCM, UNICAMP
Caixa Postal 6111
13081-970 Campinas, SP
Brazil
Fax: 55-19-3788-1089
E-mail: hackel@unicamp.br

Research supported by FAPESP
to M.L. Ribeiro and A. Santos
(Nos. 99/4512-7 and 00/04251-8,
respectively).

Received April 26, 2001
Accepted November 7, 2001

Introduction

In Brazil some 20,820 men will be diagnosed with prostate cancer this year alone and 7,320 men will die of the disease (1). In fact, prostate cancer is the third most common cancer in incidence among men in this country. Studies of risk factors such as occupation, diet, smoking, alcohol and sexual activity are still inconclusive (2-5). However, age, ethnicity and family history clearly affect the risk of prostate cancer (6).

There is evidence to support the hypothesis of hormonal etiology of prostate cancer involving androgen action (7,8). Andro-

gen is required for differentiation and growth of the prostate *in utero* and at puberty (8). Testosterone is synthesized from cholesterol by a series of reactions involving cytochrome P450 enzymes. Testosterone is converted to dihydrotestosterone (DHT), a more potent androgen, by 5 α -reductase type 2 with NADPH as the cofactor in many androgen-dependent target tissues. DHT binds to the androgen receptor (AR), and the DHT-AR complex transactivates a number of genes with AR-responsive elements (9). These events ultimately result in cell division in the prostate.

In the steroid metabolic pathway, the cy-

tochrome P450c17 α (*CYP17*) gene is the rate-limiting step in androgen biosynthesis. The *CYP17* gene maps to chromosome 10q24.3 (10) and encodes the cytochrome P450c17 α enzyme which catalyzes steroid 17 α -hydroxylase/17,20 lyase activities. A T \rightarrow C transition polymorphism creates an additional Sp1-type (CCACC box) promoter site, suggesting that this variant (A2 allele) may have an increased rate of transcription (11).

The *AR* gene, located on chromosome Xq11-12, encodes the androgen receptor, a ligand-activated transcription factor that mediates the androgenic response and stimulates the expression of genes associated with the differentiated phenotype of the prostate, such as prostate-specific antigen (12). The large exon 1, that encodes the transactivation domain, contains a highly polymorphic CAG microsatellite, a *StuI* single nucleotide polymorphism at codon 211 (G1733A), and a less polymorphic GGN repeat (13,14).

In transfection experiments, the CAG repeat length is inversely and linearly correlated with transcriptional transactivation activity, suggesting that a long polyglutamic chain interferes with AR function (15). It has been shown that shorter *AR* gene repeat lengths are related to a higher risk of prostate cancer development or progression (16,17). The effect of the length of the second *AR* gene microsatellite, consisting of GGN repeats encoding polyglycine, is unknown. However, Platz et al. (18) speculated that 23 repeats might represent the coding sequence for optimal AR protein conformation and activity. A deviation in repeat number in either direction away from the mean may represent slightly diminished AR activity. The *StuI* polymorphism allows the identification of two alleles (S1/S2), for which an association with increased risk of prostate cancer has been investigated (7).

The cloning and characterization of the 5 α -reductase type 2 gene (*SRD5A2*, mapped at 2p23) has allowed the description of ex-

tensive genetic polymorphisms in this gene (19). A valine to leucine polymorphism at codon 89 (Val89Leu) has been reported, with the Leu allele associated with lower steroid 5 α -reductase activity. Since the Leu allele would lead to lower levels of intraprostatic DHT, a protective effect towards prostate cancer has been suggested (20). Another variant that changes an alanine to a threonine at amino acid 49 (Ala49Thr) was correlated with an increase in 5 α -reductase activity and an association of the Thr allele with prostate cancer risk has been reported (21).

In the present study, we determined the frequencies of the aforementioned polymorphisms among 200 individuals from two cities in the State of São Paulo in order to evaluate the genotypic distribution of these prostate cancer markers.

Material and Methods

Polymorphisms of the *AR*, *SRD5A2* and *CYP17* genes were analyzed in DNA samples from 200 blood donors from the Hematology-Hemotherapy Center of the State University of Campinas (N = 118; mean age = 36.67 \pm 9.76 years) and Hematology-Hemotherapy Center of the State University of São José do Rio Preto (N = 82; mean age = 39 \pm 9.10 years). The protocol was approved by the Ethics Committee of the Faculdade de Ciências Médicas, Universidade Estadual de Campinas. All individuals were included in this study after giving informed written consent.

Information about the ancestors of all individuals investigated was obtained in order to characterize the ethnic composition of both samples. Among the blood donors from Campinas, 61% (N = 72) considered themselves to be white and 39% (N = 46) to be black or reported parents or grandparents of black ancestry. In the São José do Rio Preto sample, 67.3% (N = 56) considered themselves to be white and 31.7% (N = 26) to be

black or reported black ancestry. Neither the São José do Rio Preto individuals nor those from Campinas reported Amerindian or Asian ancestry in the last two generations.

Genomic DNA from leukocytes was used as a template in the polymerase chain reaction (PCR). The fragment of the *AR* gene comprising the polymorphic CAG repeat was amplified with sense and antisense primers according to Yee et al. (22). The reaction was performed in 25 μ l total volume containing 50 ng DNA, 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 1.5 mM MgCl₂, 10 mM dNTP mixture (0.2 mM dATP + 0.2 mM dTTP + 0.2 mM dGTP + 0.1 mM dCTP), 20 pmol of each primer, 2.5 units Taq DNA polymerase (Gibco-BRL, Gaithersburg, MD, USA), and 1 μ Ci [α -³²P]-dCTP (10 mCi/ml). The amplification consisted of an initial 5-min denaturation step followed by 10 cycles of 1 min at 94°C, 1 min at 55°C and 1 min at 72°C and of 25 cycles of 1 min at 94°C, 1 min at 60°C and 1 min at 72°C. The PCR product was electrophoresed on a 6% denaturing polyacrylamide DNA sequencing gel, and the length was determined after comparison to standards derived from previously sequenced CAG repeats. Gels were transferred to filter paper (Whatman 3MM), dried and exposed to X-ray film for 16-20 h.

GGN amplification was performed with the sense and antisense primers described by Sleddens et al. (14). Reaction conditions were as described above, with the following modifications: 10 mM dNTP mixture (0.2 mM dATP + 0.2 mM dTTP + 0.1 mM dCTP + 1 mM total with 3:1 mixture of 7-deaza-dGTP:dGTP) and 5% DMSO. Thermocycling consisted of 12 min at 95°C, 21 cycles of 1 min at 96°C, 1 min at 64°C and 1 min at 72°C, followed by a touchdown reduction of annealing temperature of 0.5°C per cycle until 54°C and 25 cycles with annealing temperature at 54°C. After electrophoresis and exposition to X-ray film, the length was determined by comparison with previously sequenced GGN repeats.

PCR-restriction fragment length polymorphism (PCR-RFLP) analysis was used to identify the S1/S2 alleles of the *StuI* polymorphism. A 416-bp DNA fragment was amplified with the primers reported by Lu and Danielsen (23). The PCR consisted of an initial 5-min denaturation step followed by 35 cycles of 1 min at 94°C, 1 min at 62°C and 1 min at 72°C. The presence of an A at nucleotide 1733 (S2 allele) results in a *StuI* site, which is abolished on the S1 allele. The PCR fragments were digested with *StuI* for 2 h at 37°C and the digestion products were separated by 1.5% agarose gel electrophoresis. Specific digestion products consisted of 416 and 329/87 bp for the S1 and S2 genotypes, respectively.

The *CYP17* A1/A2 alleles were detected by RFLP analysis of 629-bp DNA fragments amplified by PCR (35 cycles, annealing temperature of 58°C) with the primers described by Lunn et al. (24). The A2 allele can be identified by the presence of a second *MspA1* site created by the T→C substitution. The PCR fragments were digested with *MspA1* for 2 h at 37°C and the digestion products were separated by 3% agarose 1000 (Gibco-BRL) gel electrophoresis. Specific digestion products consisted of 577-, 577/305- and 272-, and 305/272-bp fragments for the A1/A1, A1/A2 and A2/A2 genotypes, respectively. A 52-bp fragment was present in all samples due to an invariant *MspA1* site that served as an internal control for complete digestion.

PCR-RFLP was also used to identify the Val89Leu alleles of the *SRD5A2* gene. A pair of primers was used to produce a 369-bp DNA fragment (25) after 35 cycles of PCR with an annealing temperature of 65°C. The presence of a G at nucleotide 296 (Val allele) creates an *RsaI* site, which is absent in the Leu allele. The PCR fragments were digested with *RsaI* for 2 h at 37°C and the digestion products were separated by 3% agarose 1000 (Gibco-BRL) gel electrophoresis. Specific digestion products consisted

of 169/106/73- and 21-, 169/106/94/73- and 21-, 169/106- and 94-bp fragments for the Val/Val, Val/Leu and Leu/Leu genotypes, respectively.

The *SRD5A2* Ala49Thr alleles were analyzed by PCR-allele-specific oligonucleotide hybridization. A 309-bp DNA fragment was amplified with a pair of primers described by Makridakis et al. (21) after 35 cycles of PCR with an annealing temperature of 55°C. The selected probes (ALA49-CTACCCGCCTGCCAGCCC and THR49-CTACCCGCCTACCAGCCC) were labeled with [γ - 32 P]-dATP and used to identify the Ala49Thr alleles. The THR49 oligonucleotide was used as a sense primer to produce PCR fragments containing the G→A substi-

tution in order to provide a positive control for this rare allele.

Statistical analysis

The allelic frequencies of polymorphisms involving the autosomal genes were calculated by gene counting. The goodness of fit to Hardy-Weinberg equilibrium as well as statistical analysis of differences between population samples were performed using the χ^2 test. For the *AR* gene, the distribution of the number of CAG and GGN repeats in both samples was compared using the independent *t*-test. The haplotype frequencies and the likelihood ratio test of linkage disequilibrium for genotypic data of unknown gametic phase were estimated using the Arlequin (26) program.

Preliminary comparative analysis showed no difference concerning ancestral origin in either sample. Since the Brazilian population is known to be highly miscegenated, further comparisons were not adjusted by ethnic background.

Results

The number of *AR* gene CAG repeats varied from 11 to 27 among blood donors from Campinas and from 13 to 28 in the individuals from São José do Rio Preto (Figure 1), with average numbers of 20.85 (SD \pm 2.61) and 20.35 (SD \pm 2.85), respectively. The distribution of CAG repeats was not significantly different ($t = 1.28$; $P = 0.20$). The GGN repeat length ranged from 16 to 27 (mean \pm SD, 22.55 \pm 1.33) in the Campinas sample and from 10 to 27 (22.15 \pm 2.47) among blood donors from São José do Rio Preto (Figure 2), with no difference between these two samples ($t = 1.48$; $P = 0.14$). The results concerning *StuI* polymorphism analyses are shown in Table 1. No significant differences in S1/S2 allelic frequencies were found between samples ($\chi^2 = 3.10$; $P = 0.08$). Since there were no differences for the *AR*

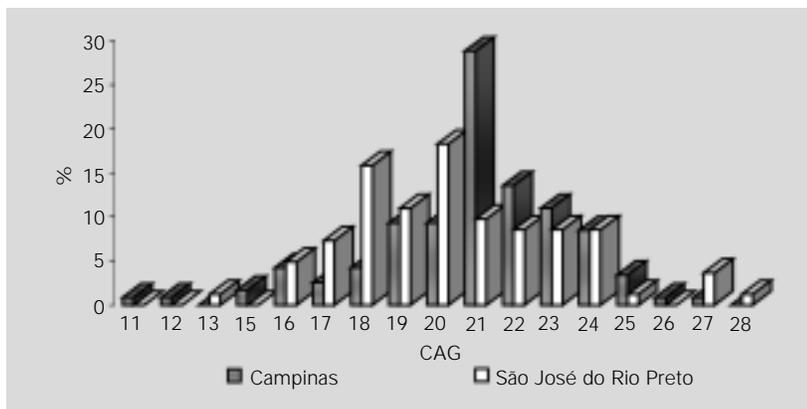


Figure 1. Distribution of CAG repeats among individuals from Campinas and São José do Rio Preto.

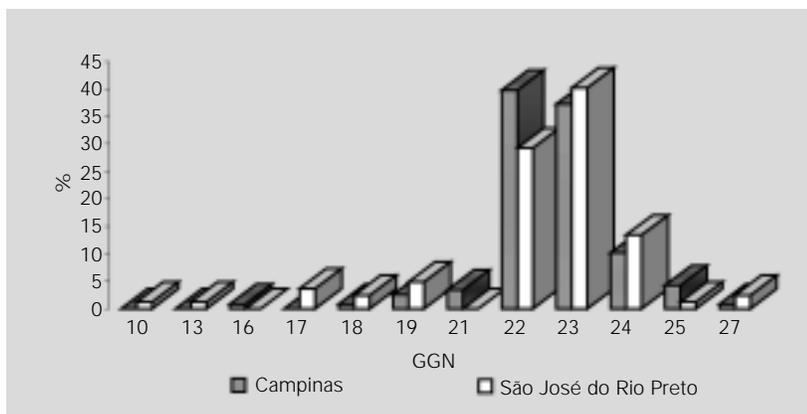


Figure 2. Distribution of GGN repeats among individuals from Campinas and São José do Rio Preto.

gene markers in either sample, these data were pooled together for haplotype analysis. Seventy-six haplotypes were estimated and the most frequent ones ($\geq 4\%$) are shown in Table 2. When a possible association (linkage disequilibrium) was tested, the results were negative for CAG and GGN repeats ($P = 0.14$) as well as for CAG repeats and *StuI* polymorphism ($P = 0.27$). Linkage disequilibrium was found between GGN repeats and *StuI* polymorphism ($P = 0.01$).

With respect to the *SRD5A2* gene, the G→A mutation that leads to the Ala49Thr substitution was not detected in the present study. The genotype frequencies of *SRD5A2* Val89Leu and *CYP17* A1/A2 polymorphisms were estimated in both population samples (Table 1). The distributions of *SRD5A2* and *CYP17* genotypes were found to be in Hardy-Weinberg equilibrium for each sample. However, the *SRD5A2* Val89Leu genotypic frequencies were different between samples ($\chi^2 = 8.41$; $P = 0.015$), with a lower than expected frequency of the Val/Leu genotype in the Campinas sample. Although no differences were detected when the allelic frequencies were compared ($\chi^2 = 1.84$; $P = 0.17$), the Val allele was more frequent in the Campinas sample (0.68) than in the São José do Rio Preto sample (0.60).

No differences were observed in the genotypic distribution of the *CYP17* A1/A2 polymorphism ($\chi^2 = 0.38$; $P = 0.82$) and identical allelic frequencies were estimated ($A1 = 0.66$).

Discussion

The “ethnic/racial” classification has been used by many investigators to denote origin by birth or descent rather than nationality. This is a difficult task, since the Brazilian population in general and especially that of the southeastern region has a high degree of admixture, even in the so-called “white” individuals for whom a significant contribution of Amerindian and African matrilineages

has been demonstrated (27).

According to Templeton (28), human races do not exist under the traditional concept of a subspecies as a geographically circumscribed population showing sharp genetic differentiation. However, many traits and their underlying polymorphic genes show independent patterns of geographical variation. As a result, some combination of characters will distinguish virtually each population from all others. Thus, it is important to define the polymorphic spectrum of genes that may be involved in cancer diseases in each population in order to evaluate its usefulness in the prediction of risk.

The six polymorphic markers evaluated in the present study showed a closely similar distribution in these two samples of the Brazilian population. Hence, for further comparisons these data were pooled together.

With respect to *AR* markers, the mean number of CAG repeats ($x = 20.6$) observed in the present study was the lowest one among different populations, with a narrow range of distribution (11 to 28 repeats), while the highest value ($x = 23$) was reported in a Chinese population sample (Table 3) (29-33). The mean GGN number of 22.4 repeats in the Brazilian population was also lower than those found in North American and

Table 1. Genotype distribution of *AR*, *SRD5A2* and *CYP17* polymorphisms.

Gene (polymorphism)	Genotypes	Campinas (%) (N = 118)	São José do Rio Preto (%) (N = 82)
<i>AR</i> (<i>StuI</i>)	S1	29 (24.57)	11 (13.41)
	S2	89 (75.43)	71 (86.59)
<i>SRD5A2</i> (Val89Leu)	Val/Val	61 (51.69)	27 (32.93)
	Val/Leu	40 (33.90)	44 (53.66)
	Leu/Leu	17 (14.41)	11 (13.41)
χ^2 HW (P)		5.32 (ns)	1.16 (ns)
<i>CYP17</i> (A1/A2)	A1/A1	51 (43.22)	34 (41.46)
	A1/A2	53 (44.92)	40 (48.78)
	A2/A2	14 (11.86)	8 (9.76)
χ^2 HW (P)		0.001 (ns)	0.35 (ns)

HW = Hardy-Weinberg; ns = not significant at the 5% level.

Chinese populations (Table 3) (18,32). The allelic frequencies of *StuI* polymorphism (S1 = 0.2 and S2 = 0.8) were close to those reported by Lu and Danielsen (23) (S1 = 0.13 and S2 = 0.87) for white North Americans.

It has been suggested that part of the large “ethnic/racial” difference in prostate cancer risk may be explained by the observed variations in the CAG repeat length between populations. Previous studies have shown that the CAG repeat length is shortest in African Americans, intermediate in whites and longest in Asians (13,34), a fact probably related to higher, intermediate and lower prostate cancer risks. Indeed, 55% of African North Americans are reported to have a CAG re-

peat length shorter than 20 (13,34) versus 22% of the white men from two North American studies (16,17) versus only 10% of Chinese men (32). In the present study, 30.5% of the individuals had less than 20 CAG repeats, suggesting an intermediate-to-high risk condition in our highly miscegenated population.

According to Stanford et al. (17), men with a CAG repeat number lower than 22 have a 23% increase in risk of prostate cancer when compared to men with ≥ 22 repeats. Findings related to GGN number revealed that men who had ≤ 16 repeats experienced a 60% higher risk than men with longer repeat lengths (17). In the present study, almost 44% of the individuals had less than 22 CAG repeats, but only 1.5% had ≤ 16 GGN repeats. When both repeat lengths are considered jointly, 63.5% of the individuals in the Brazilian sample showed the combined < 22 CAG > 16 GGN repeat numbers and 35% had ≥ 22 CAG > 16 GGN repeats. Only 1 and 0.5% of the individuals exhibited the < 22 CAG ≤ 16 GGN and ≥ 22 CAG < 16 GGN lengths, respectively (Table 4). This distribution is strikingly different from that for the North American population described by Stanford et al. (17), where the same stratification was done ($\chi^2 = 125.72$; $P = 0$). In addition, when the present data were classified according to the categories proposed by Platz et al. (18) or by Hsing et al. (32) (Table 4), the combined repeat lengths showed remarkably significant differences in their distributions among these populations ($\chi^2 = 169.49$; $P = 0$ and $\chi^2 = 72.44$; $P = 0$, respectively).

These statistical differences are obviously a consequence of the different mean CAG and GGN repeat values in each population and also of the different criteria used for the classification of combined CAG and GGN repeats that are based either on the mean (18) or on the median lengths within the populations (17,32). In our population, median lengths (CAG = 21 repeats and GGN = 23 repeats) would allow us to classify our

Table 2. Estimated haplotype frequencies for the AR gene.

Haplotypes			Frequencies
CAG	GGN	StuI	
21	22	S2	0.085
20	23	S2	0.065
22	23	S2	0.045
21	23	S2	0.040
23	22	S2	0.040
24	22	S2	0.040
19	22	S2	0.040

Table 3. CAG and GGN repeats in different populations.

Repeat	Minimum	Maximum	Mean	n	Study group
CAG	6	39	21	588	North American (16)
	11	32	21.98	266	North American (17)
	12	30	21.7	186	European (Sweden) (29)
	13	30	21	390	European (UK) (30)
	8	37	22	882	North American (31)
	10	33	23	300	Chinese (32)
	11	32	21.9	305	North American (33)
	11	28	20.65	400	Brazilian
GGN	10	29	23	794	North American (18)
	14	27	23	295	Chinese (32)
	10	26	22.38	400	Brazilian

n = number of chromosomes.

data according to Platz et al. (18). In this situation, the differences between these two populations can be shown by the fact that only 8.9% of the North Americans (Table 4) (18) had less than 23 GGN repeats versus 45.5% of the Brazilian individuals. These data may be of concern, since 23 GGN repeats are considered to represent the coding sequence for optimal AR protein activity and shorter GGN repeat lengths appear to be associated with a moderate increase in the risk of prostate cancer.

Linkage disequilibrium between CAG and GGN repeats has been reported by Platz et al. (18), but has not been detected by other authors (34,35) in male base populations or in the present data. The relevance of linkage disequilibrium between GGN repeats and *StuI* polymorphism observed in the present study has to be further evaluated in prostate cancer patients.

Concerning the *SRD5A2* Val89Leu polymorphism, the observed allelic frequencies do not differ from North American data (Table 5) (20,24,36,37). The absence of the Thr49 allele of the same gene may be attributed to its low frequency in our population. Indeed, allelic frequencies lower than 2.5% have been previously reported for this allele among the so-called African and Latino North Americans (21). It is noteworthy that a frequency of 3.5% for the Thr49 allele has been reported only in a sample of white North American prostate cancer patients (37), suggesting that this allele may be of relevance for this disease.

The *CYP17* A1/A2 allelic frequencies (Table 5) are similar to those described for a mixed North American population (24) and do not differ from a European sample (38). On the other hand, significant differences are observed concerning the reported frequencies from Swedish ($\chi^2 = 7.73$; $P = 0.02$) and Japanese populations ($\chi^2 = 18.93$; $P = 0.0001$).

To our knowledge, this is the first report concerning the frequencies of these six poly-

Table 4. Prevalent combined numbers of CAG and GGN repeats of the AR gene in different populations.

CAG	GGN	N (%)	Study group
<22	>16	93 (38.3)	North American (17)
<22	≤16	77 (31.7)	
≥22	≤16	41 (16.9)	
≥22	>16	32 (13.2)	
21-23	≥23	279 (35.1)	North American (18)
<21	≥23	212 (26.7)	
>23	23	133 (16.7)	
>23	>23	99 (12.4)	
All	<23	71 (8.9)	
≥23	≥23	120 (41.4)	Chinese (32)
<23	≥23	115 (39.6)	
≥23	<23	29 (10.0)	
<23	<23	26 (9.0)	
<22	>16	127 (63.5)	Brazilian
<22	≤16	2 (1.0)	
≥22	≤16	1 (0.5)	
≥22	>16	70 (35.0)	

N = number of individuals.

Table 5. Distribution of *SRD5A2* and *CYP17* genotypes in different populations.

Gene (polymorphism)	Genotype distribution				N	Study group
	Val/Val	Val/Leu	Leu/Leu	Leu89		
<i>SRD5A2</i> (Val89Leu)	0.46	0.42	0.12	0.32	286	North American (20)
	0.47	0.43	0.10	0.31	386	North American (36)
	0.43	0.49	0.08	0.32	108	North American (24)
	0.50	0.40	0.10	0.30	245	North American (37)
	0.44	0.42	0.14	0.35	200	Brazilian
<i>CYP17</i> (A1/A2)	A1/A1	A1/A2	A2/A2	A2		
	0.43	0.46	0.11	0.35	167	North American (24)
	0.37	0.53	0.10	0.36	252	European (Austria)(38)
	0.29	0.55	0.16	0.43	160	European (Sweden) (39)
	0.25	0.47	0.28	0.51	262	Japanese (40)
	0.42	0.47	0.11	0.34	200	Brazilian

N = number of individuals.

morphic markers in South America. The allelic frequencies of the autosomal markers are closely similar to those described in most North American and European populations. In relation to these polymorphisms, further studies should be conducted on prostate cancer patients to verify if the alleles that

are considered of relevance in most populations can also be taken into account in Brazilian urban populations, where significant contributions from European and African gene pools can be found. In contrast, remarkable differences from reported literature data were observed for the CAG and GGN repeat number combinations. Because of the importance of AR in prostate cancer etiology, further investigations are needed to evaluate the combined effect of CAG

and GGN repeats and their use as molecular markers for the identification of men at higher risk of developing prostate cancer in the Brazilian population.

Acknowledgments

We gratefully acknowledge the Hematology-Hemotherapy Center, State University of Campinas, for allowing us access to the blood donors.

References

- Instituto Nacional do Câncer (INCA) (2001). Estimativa de incidência e mortalidade por câncer no Brasil. Ministério da Saúde, Rio de Janeiro, RJ, Brazil.
- Van der Gulden JW, Kolk JJ & Verbeek AL (1992). Prostate cancer and work environment. *Journal of Occupational and Environmental Medicine*, 34: 402-409.
- Daviglus ML, Dyer AR, Persky V, Chavez N, Drum M, Goldberg J, Liu K, Morris DK, Shekelle RB & Stamler J (1996). Dietary betacarotene, vitamin C, and risk of prostate cancer: results from the Western electric study. *Epidemiology*, 7: 472-477.
- Ilic M, Vlijinac H & Marinkovic J (1996). Case-control study of risk factors for prostate cancer. *British Journal of Cancer*, 74: 1682-1686.
- Lumey LH (1996). Prostate cancer and smoking: a review of case-control and cohort studies. *Prostate*, 29: 249-260.
- Parkin DM, Pisani P & Ferlay J (1993). Estimates of the worldwide incidence of eighteen major cancers in 1985. *International Journal of Cancer*, 54: 594-606.
- Ross RK, Pike MC, Coetzee GA, Reichardt JKV, Yu MC, Feigelson H, Stanczyk FZ, Kolonel LN & Henderson BE (1998). Androgen metabolism and prostate cancer: Establishing a model of genetic susceptibility. *Cancer Research*, 58: 4497-4504.
- Wilding G (1992). The importance of steroid hormones in prostate cancer. *Cancer Surveys*, 14: 113-130.
- Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson E & French FS (1995). Androgen receptor defects: historical, clinical, and molecular perspectives. *Endocrine Reviews*, 16: 271-321.
- Picado-Leonard J & Miller WL (1987). Cloning and sequence of the human gene for P450c17 (steroid 17 α -hydroxylase/17,20 lyase): similarity with the gene for P450c21. *DNA*, 6: 439-448.
- Carey AH, Waterworth D, Patel K, White D, Little J, Novelli P, Franks S & Williamson R (1994). Polycystic ovaries and premature male pattern baldness are associated with one allele of the steroid metabolism gene CYP17. *Human Molecular Genetics*, 3: 1873-1876.
- Riegman PH, Vlietstra RJ, van der Korput JA, Brinkmann AO & Trapman J (1991). The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. *Molecular Endocrinology*, 5: 1921-1930.
- Edwards A, Hammond HA, Jin L, Caskey CT & Chakraborty R (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, 12: 241-253.
- Sleddens HFBM, Oostra BA, Brinkmann AO & Trapman J (1993). Trinucleotide (GGN) repeat polymorphism in the androgen receptor gene. *Human Molecular Genetics*, 2: 493.
- Kazemi-Esfarjani P, Trifiro MA & Pinsky L (1995). Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: Possible pathogenic relevance for the (CAG)_n-expanded neuropathies. *Human Molecular Genetics*, 4: 523-527.
- Giovannucci E, Stampfer MJ, Krithivas K, Brown M, Dahl D, Brufsky A, Talcott J, Hennekens CH & Kantoff PW (1997). The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proceedings of the National Academy of Sciences, USA*, 94: 3320-3323.
- Stanford JL, Just JJ, Gibbs M, Wicklund KG, Leal CL, Blumenstein BA & Ostrander EA (1997). Polymorphic repeats in the androgen receptor gene: Molecular markers of prostate cancer risk. *Cancer Research*, 57: 1194-1198.
- Platz EA, Giovannucci E, Dahl DM, Krithivas K, Hennekens CH, Brown M, Stampfer MJ & Kantoff PW (1998). The androgen receptor gene GGN microsatellite and prostate cancer risk. *Cancer Epidemiology, Biomarkers and Prevention*, 7: 379-384.
- Thigpen AE, Davis DL, Milatovich A, Mendonca BB, Imperato-McGinley J, Griffin JE, Francke A, Wilson JD & Russell DW (1992). Molecular genetics of steroid 5 α -reductase 2 deficiency. *Journal of Clinical Investigation*, 90: 799-809.
- Makridakis N, Ross RK, Pike MC, Chang L, Stanczyk FZ, Kolonel LN, Shi CY, Yu MC, Henderson BE & Reichardt JK (1997). A prevalent missense substitution that modulates activity of prostatic steroid 5 α -reductase. *Cancer Research*, 57: 1020-1022.
- Makridakis NM, di Salle E & Reichardt JK (2000). Biochemical and pharmacogenetic dissection of human steroid 5 α -reductase type II. *Pharmacogenetics*, 10: 407-413.
- Yee CJ, Roodi N, Verrier CS & Parl FF (1994). Microsatellite instability and loss of heterozygosity in breast cancer. *Cancer Research*, 54: 1641-1644.
- Lu J & Danielsen M (1996). A Stul polymorphism in the human androgen receptor gene (AR). *Clinical Genetics*, 49: 323-324.
- Lunn RM, Bell DA, Mohler JL & Taylor JA (1999). Prostate cancer risk and polymorphisms in 17 hydroxylase (CYP17) and steroid reductase (SRD5A2). *Carcinogenesis*, 20: 1727-1731.
- Katz MD, Cai LQ, Zhu Y, Herrera C,

- DeFilloRicart M, Shackleton HL & Imperato-McGinley J (1995). The biochemical and phenotypic characterization of females homozygous for 5 α -reductase deficiency. *Journal of Clinical Endocrinology and Metabolism*, 80: 3160-3167.
26. Schneider S, Kueffer JM, Roesli D & Excoffier L (1997). Software Arlequin Ver.1.1: A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.
27. Alves-Silva J, Santos MS, Pena SDJ & Prado VF (1999). Multiple geographic sources of region V 9-bp deletion haplotypes in Brazilians. *Human Biology*, 71: 245-259.
28. Templeton AR (1998). Human races: a genetic and evolutionary perspective. *American Anthropologist*, 100: 632-650.
29. Bratt O, Borg A, Kristofferson U, Lundgren R, Zhang QX & Olsson H (1999). 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. *British Journal of Cancer*, 81: 672-676.
30. Edwards SM, Badzioch MD, Minter R, Hamoudi R, Collins N, Ardern-Jones A, Dowe A, Osborne S, Kelly J, Shearer R, Easton DF, Saunders GF, Dearnaley DP & Eeles RA (1999). Androgen receptor polymorphisms: Association with prostate cancer risk, relapse and overall survival. *International Journal of Cancer*, 84: 458-465.
31. Krithivas K, Yurgalevitch SM, Mohr BA, Wilcox CJ, Batter SJ, Brown M, Longcope C, McKinlay JB & Kantoff PW (1999). Evidence that CAG repeat in the androgen receptor gene is associated with the age-related decline in serum androgen levels in men. *Journal of Endocrinology*, 162: 137-142.
32. Hsing AW, Gao YT, Wu G, Wang X, Deng J, Chen YL, Sesterhenn A, Mostofi FK, Benichou J & Chang C (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 Research*, 60: 5111-5116.
33. Lange EM, Chen H, Brierley K, Livermore H, Wojno KJ, Langefeld CD, Lange K & Cooney KA (2000). The polymorphic exon 1 androgen receptor CAG repeat in men with a potential inherited predisposition to prostate cancer. *Cancer Epidemiology, Biomarkers and Prevention*, 9: 439-442.
34. Irvine RA, Yu MC, Ross RK & Coetzee GA (1995). The CAG and GGC microsatellites of androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Research*, 55: 1937-1940.
35. Macke JP, Hu N, Hu S, Bailey M, King VL, Brown T, Hamer D & Nathans J (1993). Sequence variation in the androgen receptor gene is not a common determinant of male sexual orientation. *American Journal of Human Genetics*, 53: 844-852.
36. Febbo PG, Kantoff PW, Platz EA, Cassey D, Batter S, Giovannucci E, Hennekens CH & Stampfer MJ (1999). The V89L polymorphism in the 5 α -reductase type 2 gene and risk of prostate cancer. *Cancer Research*, 59: 5878-5881.
37. Jaffe JM, Malkowicz SB, Walker AH, MacBride S, Peschel R, Tomaszewski J, Aradalem KV, Wein AJ & Rebeck TR (2000). Association of SRD5A2 genotype and pathological characteristics of prostate tumors. *Cancer Research*, 60: 1626-1630.
38. Gsur A, Bernhofer G, Hinteregger S, Haidinger G, Schatzl G, Madersdacher S, Marberger M, Vutuc C & Micksche M (2000). A polymorphism in the CYP17 gene is associated with prostate cancer risk. *International Journal of Cancer*, 87: 434-437.
39. Wadelius MA, Andersson SO, Johansson JE, Wadelius C & Rane A (1999). Prostate cancer associated with CYP17 genotype. *Pharmacogenetics*, 9: 635-639.
40. Habuchi T, Liqing Z, Suzuki T, Sasaki R, Tsuchiya N, Tachiki H, Shimoda N, Satoh S, Sato K, Takechi Y, Kamoto T, Ogawa O & Kato T (2000). Increased risk of prostate cancer and benign prostatic hyperplasia associated with a CYP17 gene polymorphism with a gene dosage effect. *Cancer Research*, 60: 5710-5713.