



# CAG repeats in the androgen receptor gene are shorter in patients with pulmonary, esophageal or bladder carcinoma and longer in women with uterine leiomyoma

XIAO-YING TENG<sup>1</sup>, GUI-QIU LIU<sup>2</sup>, XIAO-LI DIAO<sup>2</sup>, ZONG-YONG WU<sup>3</sup>,  
LAN LI<sup>1</sup>, WEI ZHANG<sup>2</sup>, XUN ZHANG<sup>1</sup> and QIN SU<sup>1,2</sup>

<sup>1</sup>Department of Pathology, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021; <sup>2</sup>Department of Pathology, Tangdu Hospital, The Fourth Military Medical University, Xi'an, Shaanxi 710038; <sup>3</sup>Clinical Laboratory, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, P.R. China

Received October 9, 2009; Accepted November 9, 2009

DOI: 10.3892/or\_00000702

**Abstract.** Preferential occurrence of pulmonary, esophageal and bladder carcinomas in males indicate a possible involvement of androgen receptor (AR)-mediated functions. We evaluated the roles of the CAG repeat polymorphism in AR exon 1 in development of these lesions. The exon 1 of AR gene was amplified in samples from 198 male patients with lung carcinoma, 183 with esophageal carcinoma, 95 with bladder carcinoma and 94 males with appendicitis, as a reference group. Mean numbers of the CAG repeat in these 3 cancer groups were determined to be 20.2, 20.0 and 20.0, respectively, all being significantly smaller than that of the reference group (21.1;  $P < 0.05$ ). Samples from 118 female patients with lung carcinoma and 154 females with appendicitis, as a reference group, were examined, with the mean CAG repeat number significantly smaller (19.8) than that of the female reference group (20.7;  $P < 0.01$ ). Samples from 108 patients with uterine leiomyoma were also examined, and their CAG repeat numbers were found to be markedly expanded (23.4;  $P < 0.01$ ). The patients with multiple leiomyomas tend to carry a longer CAG repeat structure, with the mean CAG repeat number longer in the multicentric multiple cases (24.1) compared to that of the unicentric, multinodular cases (22.2) and those with solitary lesions (23.1;  $P < 0.01$ ). These results indicate that a shorter CAG repeat structure may predispose individuals to a higher risk to some male-predominant neoplasms including pulmonary, esophageal

and bladder carcinomas and a longer one confers women greater susceptibility to leiomyoma development in the uterus.

## Introduction

Carcinomas of the lung, esophagus and urinary bladder are among the most common malignancies in man worldwide. While their pathogenic mechanisms are not fully understood, it is noteworthy that all of them occur more frequently in males, with their male/female ratios estimated to be 2.7, 3.5 and 5.0, respectively (1-3). The preferential occurrence of these diseases in males strongly indicates involvement of some androgenic functions in cancer development. Elevation of circulating androgen level has been observed in these patients (4) and androgen receptor (AR) was detected in tissue samples of these carcinomas (5-7). It is largely unknown, whether the male predilection indicates a role of androgen-mediated functions in the development and/or progression of these neoplasms. Androgen play its function through AR. It is reasonable to think that AR is even more important in modulation of the carcinogenesis than its ligands, considering the persistent existence and functional activation of AR molecules in reproductive and some somatic tissues.

AR is a member of the steroid receptor family, encoded by a large gene located at q11-12 on X chromosome (8,9). Within its exon 1, there is a highly polymorphic CAG short-tandem repeat (STR), which encodes for a polyglutamine chain of variable length in the aminoterminal transactivation domain of the protein (10,11). An inverse relationship was demonstrated between the STR and AR activity, with the tracts of shorter size conferring greater activity than those of larger size (10). The CAG STR polymorphism has been found to be involved in several androgen-involved pathogenic processes in both men and women (12). A shorter CAG STR was observed in patients with coronary artery disease (13), prostate cancer (14-16), ovarian cancer (17-19) and hepatitis B virus-related hepatocellular carcinoma (20) and a longer one was noted in female patients with breast cancer (21) and uterine endometrial carcinoma (22). It is conceivable to

*Correspondence to:* Professor Qin Su, Department of Pathology, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Panjiayuan Nanli 17, Beijing 100021, P.R. China

E-mail: q.su@wjgnet.com

*Key words:* CAG repeat polymorphism, androgen receptor gene, carcinoma, lung, esophagus, bladder, uterine leiomyoma

consider that the CAG repeat polymorphism is one of the genetic factors predisposing individuals to some diseases including various types of cancer.

During this survey, the CAG repeat numbers in patients with lung, bladder and esophageal carcinomas were measured and their relationships with the diseases were evaluated. Samples from female patients with uterine leiomyoma, whose development was ascribed to persistent stimulation by estrogen (23), were also evaluated as an additional disease group.

## Materials and methods

**Samples.** All samples used were obtained from Cancer Hospital, Chinese Academy of Medical Sciences in Beijing and Tangdu Hospital, The Fourth Military Medical University in Xi'an. The study protocol was approved by the Medical Ethics Commission of the Fourth Military Medical University. The lung carcinoma cases enrolled included 198 male and 118 female patients, with their ages ranging from 29 to 81 years (median, 60 years) and 19 to 75 years (median, 56 years), respectively. All of the specimens were reviewed independently by three pathologists (Q. Su, X.-Y. Teng and X.-L. Diao), with 151 (119 from males and 32 from females) diagnosed as squamous cell carcinoma, 131 (65 from males and 66 from females) as adenocarcinoma, 27 (12 from males and 15 from females) as small cell lung carcinoma (SCLC) and 6 (1 from males 5 from females) as other cancer types. The esophageal carcinoma group included 183 male patients, with their ages ranging from 35 to 78 years (median, 59 years). A total number of 91 bladder urothelial carcinoma specimens from male patients were assessed, with their ages ranging from 17 to 84 years (median, 63 years). Samples from 108 patients with uterine leiomyomas were examined, including 31 cases with single tumor, 43 with 2 or 3 tumors, 24 with >3 tumor nodules and 10 multinodular cases whose tumor numbers were not recorded. Among the 77 multinodular cases, 25 were classified as unicentric type and 23 as multicentric type by assays on their X chromosome inactivation patterns (24). In addition, 248 consecutive appendix samples, resected from 94 male and 154 female adults for appendicitis, were examined as references.

**DNA isolation and amplification of the AR gene exon 1.** Paraffin-embedded tissues were retrieved from the files and sections of 10  $\mu\text{m}$  in thickness were prepared. The sections were deparaffinized in xylene and rehydrated by rinsing in graded ethanol. Genomic DNA was isolated from the samples using an extraction kit (Qiagen, Hilden, Germany). DNA of peripheral leukocytes was also extracted for the cases with pulmonary and esophageal carcinomas.

Exon 1 of AR gene, containing the CAG STR region, was amplified by nested polymerase chain reaction (PCR) as described previously (24). Two pairs of primers (25) were used, including 5'-GAG GAG CTT TCC AGA ATC TG-3' (AR1A), 5'-CAT GGG CTT GGG GAG A-3' (AR1B), 5'-TCC AGA ATC TGT TCC AGA GC-3' (AR2A) and 5'-TGG GGA GAA CCA TCC TCA CC-3' (AR2B). The first-round reaction was in a volume of 50  $\mu\text{l}$  containing DNA template 5  $\mu\text{l}$ , 10 mmol/l deoxynucleotide triphosphate 4  $\mu\text{l}$ , 50 mmol/l  $\text{MgCl}_2$  1.5  $\mu\text{l}$ , 20 pmol/l primers AR1A and AR1B, 1  $\mu\text{l}$  each

and 0.25 unit of Taq DNA polymerase (Takara Biotechnology Co., Ltd., Dalian, China). After preheating, 25 cycles were performed under the condition of 97°C for 40 sec, 56°C for 50 sec and 72°C for 1 min, with a final extension at 72°C for 15 min. The amplification products were diluted at 1:10 and subjected to second-round reaction using the primer pair AR2A/AR2B and the same conditions as for the first one.

**Size assessment of AR amplification products by denaturing polyacrylamide gel electrophoresis and sequencing.** Efficacy of the PCR for the AR gene was assessed by electrophoresis on 2% agarose gels. Amplification products, 3  $\mu\text{l}$  each, were mixed with the same volume of sample buffer (99% formamide, 1 mg/ml bromophenol blue and 1 mg/ml xylene cyanol), and loaded onto a 0.75-mm thick gel (10% polyacrylamide gel containing 8.0 M urea) as long as 26 cm and resolved at a voltage of 150 V for 26 h. Then, the gels were fixed for 4 min in 10% ethanol containing 0.5% acetic acid, treated with 10% nitric acid for 4 min, washed in deionized water for 4 min and impregnated in 0.6% silver nitrate for 45 min. After rinsing the gel in deionized water for 5 min, the reaction was visualized in a solution containing 3% sodium carbonate and 0.1% formaldehyde for 5 min and terminated by soaking gels in 10% acetic acid for 5 min. Sizes of products were determined using 10-bp DNA ladders (Gibco BRL).

Representative PCR products for AR exon 1 were also analyzed by direct sequencing, as described previously (24), in order to correlate their migration behaviors, sizes and numbers of CAG repeat within the gene.

**Evaluation of data and statistical analysis.** Numbers of the CAG repeats ( $n$ ) were calculated from the size of the amplification products, in relation to a series of standards obtained by direct sequencing of representative PCR products. Statistical computations were performed using the Statistical Package for Social Science 13.0 for Windows (SPSS, Inc, Chicago, IL, USA). Differences in CAG repeat numbers between two groups were described using the t-test and the analysis of variance (ANOVA) test. P-values <0.05 were regarded as being statistically significant.

## Results

**Assessment of CAG STR sizes of AR gene.** Amplification was successful for all of the samples examined. Single band was demonstrated in gels in all of the samples from males. For the heterozygous samples from females, only shorter alleles were assessed. Sizes of the products were measured by their mobility on denaturing polyacrylamide gels. Accuracy of the assay was confirmed by fully consistent results obtained through electrophoresis and sequencing analyses in 11 representative samples (Table I). The  $n$  values were calculated according to sizes of the products (Fig. 1). Results of the assay are listed in Table II, with the  $n$  values assessed separately in patients with pulmonary, esophageal and bladder carcinomas, uterine leiomyomas and appendicitis.

**Shorter CAG structures observed in males with lung carcinoma.** As shown in Fig. 2A and D,  $n$  values in male patients with lung cancer ranged from 7 to 31, with their mean



Measurement of sizes of AR gene amplification products by denaturing gel electrophoresis and sequencing and determination of the CAG repeat numbers.

Codes of samples	Sizes measured by mobility (bp)	Sizes measured by sequencing (bp)	CAG repeat numbers ( <i>n</i> )
1	193	193	11
2	202	202	14
3	208	208	16
4	211	211	17
5	220	220	20
6	226	226	22
7	229	229	23
8	232	232	24
9	235	235	25
10	247	247	29
11	250	250	30

Table II. Numbers of CAG repeat structures in the androgen receptor gene in male and female patients with neoplastic and non-neoplastic diseases.

Patient groups	Case numbers	CAG repeat numbers		P-values <sup>a</sup>
		Ranges	Mean ± SD	
Males with				
appendicitis	94	14-30	21.1±3.4	
lung cancer	198	7-31	20.2±3.5	0.043
esophageal cancer	183	6-32	20.0±5.0	0.043
bladder cancer	91	9-27	19.8±3.3	0.013
Females with				
appendicitis	154	14-27	20.7±2.9	
lung cancer	118	9-24	19.8±2.5	0.006
uterine leiomyoma	108	17-31	23.4±2.5	0.001

<sup>a</sup>As compared with the reference group admitted for appendicitis and without a detectable tumor.

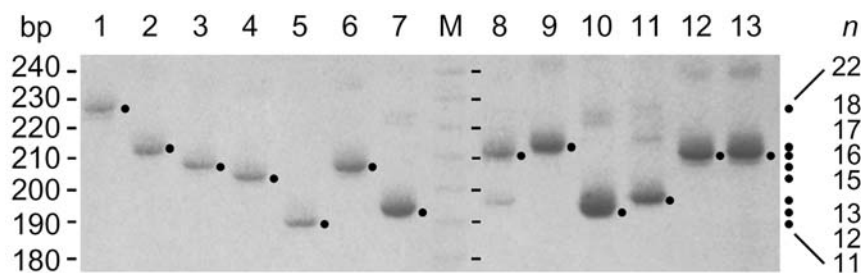


Figure 1. Representative data of measuring CAG STR length of the AR gene by denaturing gel electrophoresis. The products, amplified from 11 samples from male patients with urothelial carcinoma in the bladder, were loaded onto a polyacrylamide gel containing 8 M urea (lanes 1-11) and resolved for 26 h. Their sizes were assessed by correlating their mobility to that of 10-bp DNA markers (-, on the lane M). Locations of the markers were denoted by short bars, and their sizes, from 180 to 240 base pairs (bp), are on the left side. Numbers of the CAG repeat (*n*), calculated by sizes of these products (●), are on the right side.

being 20.2±3.5, while those in male references ranged from 14 to 30, with their mean being 21.1±3.4. The *n* values of CAG repeat in both the cancer patients (12.6%, 25/198) and reference group (18.1%, 17/94) peaked at 20, but the STR was found to be significantly shorter in the former group than in the latter ( $P<0.05$ ; Table II). Occurrence of shorter CAG

STR ( $n<22$ ) was observed in 64.1% (127/198) of the male carcinoma patients, being more frequent than in the reference group (59.6%, 56/94;  $P<0.05$ ).

While the patients with lung carcinoma were shown to have a shorter CAG STR, it is conceivable to consider that cancer may develop earlier in these patients. The mean

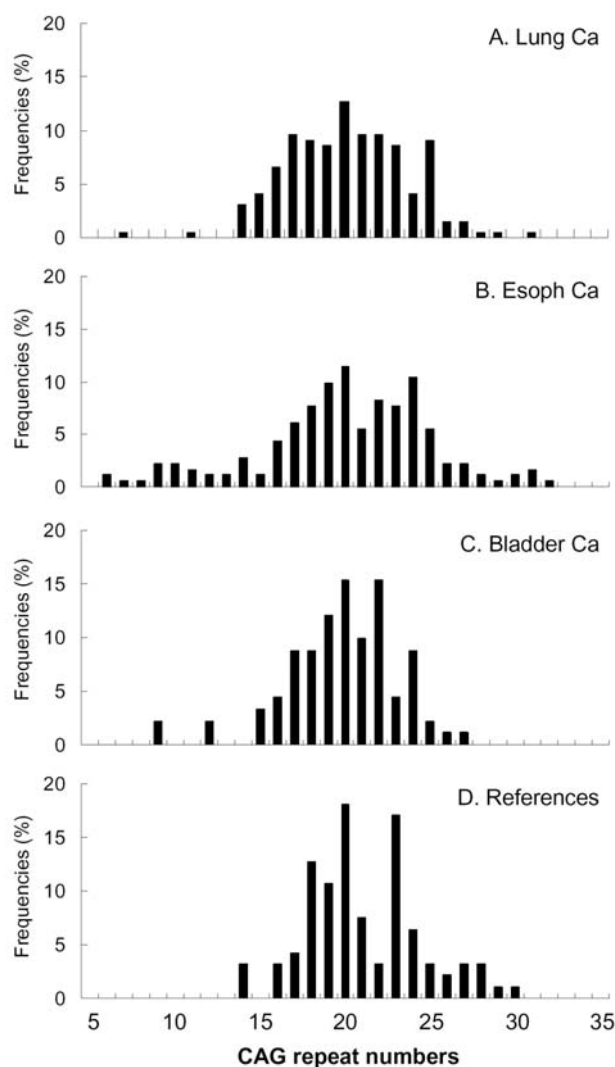


Figure 2. Distribution patterns of CAG repeat numbers of the AR gene in males with pulmonary (Lung Ca; A), esophageal (Esoph Ca; B) and urinary bladder carcinomas (Bladder Ca, C), as well as those with appendicitis as a reference group (References; D).

repeat numbers of the cancer patients aging <55 years and  $\geq 55$  years at diagnosis were determined to be  $20.0 \pm 3.3$  and  $20.3 \pm 3.7$ , respectively. The difference did not attain statistical significance ( $P > 0.05$ ).

**Shorter CAG repeat structures observed in females with lung carcinoma.** As shown in Fig. 3A and C,  $n$  values of the CAG repeat in the female patients ranged from 9 to 24, with the mean of  $19.8 \pm 2.5$  and the most frequent number located at 21, while those in female reference group ranged from 14 to 27, with a mean of  $20.7 \pm 2.9$  and the most frequent number located at 20. The CAG STR was also shorter in female cancer patients than those without a carcinoma ( $P < 0.01$ ; Table II). Occurrence of shorter CAG STR ( $n < 22$ ) was observed in 78% (92/118) of the cancer patients, being more frequent than in the reference individuals without a detectable neoplasm (64%, 98/154;  $P < 0.05$ ). The mean CAG repeat number was  $20.1 \pm 3.4$  in the patients aging <55 years and  $19.5 \pm 2.8$  in those up to 55 years. The difference did not attain statistical significance ( $P > 0.05$ ).

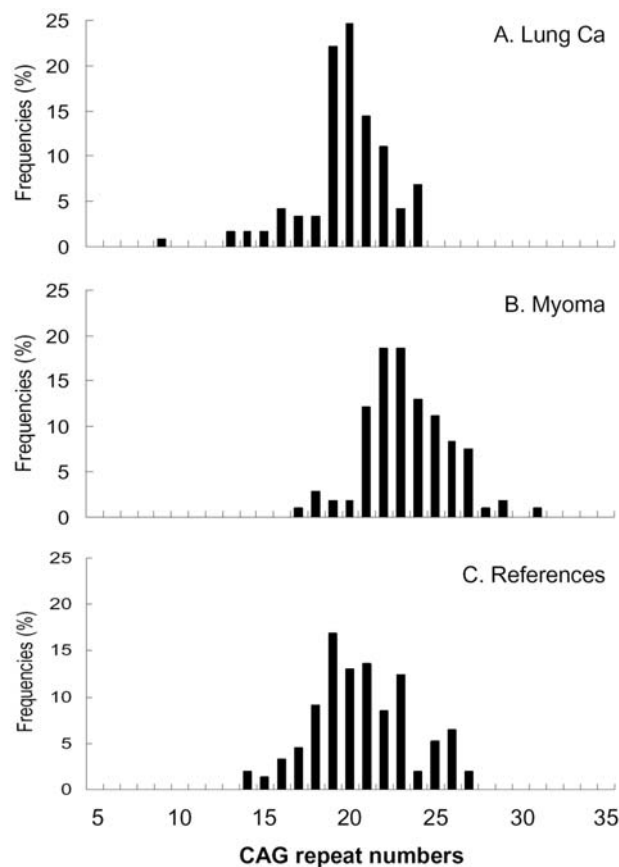


Figure 3. Distribution patterns of CAG repeat numbers of the AR gene in females with pulmonary carcinoma (Lung Ca; A) and uterine leiomyoma (Myoma; B), as well as those with appendicitis as a reference group (References; C).

**Shorter CAG repeat structures observed in males with esophageal carcinoma.** Distribution patterns of the  $n$  values of the 183 male patients with esophageal carcinoma and the corresponding reference groups are shown in Fig. 2B and D. The values of the cancer patients ranged from 6 to 32, with a mean of  $20.0 \pm 5.0$ . The STR was significantly shorter in the carcinoma patients than in corresponding reference group ( $21.1 \pm 3.4$ ,  $P < 0.05$ ; Table II). However, shorter CAG repeats ( $n < 22$ ) were detected at a similar frequency (59.0%, 108/183) as in the corresponding reference individuals (59.6%, 56/94;  $P > 0.05$ ). The mean repeat number of the cancer patients aged <55 years at diagnosis was  $20.6 \pm 4.6$ , whereas that of patients  $\geq 55$  years was  $19.8 \pm 5.3$ . The difference was not of statistical significance ( $P > 0.05$ ).

**Shorter CAG repeat structures observed in males with bladder carcinoma.** A total number of 91 male patients with urothelial carcinoma patients were examined. The  $n$  values ranged from 9 to 27, with the most frequent numbers located at 20 (Fig. 2C). Their mean was  $19.8 \pm 3.3$ , being significantly shorter than that of the male reference group without a detectable neoplasm ( $21.1 \pm 3.4$ ,  $P < 0.05$ ; Table II). Percentage of the cancer patients with shorter CAG repeats ( $n < 22$ ) was shown to be higher (67.0%, 61/91) than that of the reference group (59.6%, 56/94;  $P < 0.05$ ). The  $n$  values in different age groups were also determined. Mean number of the cancer patients aged <55 years was  $20.9 \pm 4.0$  and that of the patients aged

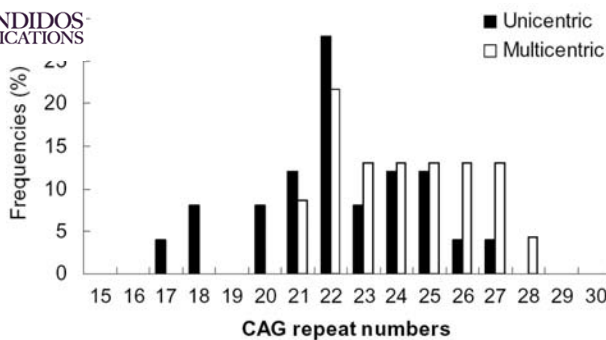


Figure 4. Distribution patterns of CAG repeat numbers of the AR gene in female patients with unicentric (Unicentric) and multicentric leiomyomas (Multicentric) in the uterus.

$\geq 55$  years was  $19.5 \pm 3.1$ . The difference did not attain statistical significance ( $P > 0.05$ ).

*Longer CAG repeat structures observed in patients with uterine leiomyoma.* Samples from 108 patients with uterine leiomyoma were also examined. The  $n$  values ranged from 17 to 31 in the leiomyoma patients and 14 to 27 in the reference cases, peaking at 19 in the former and at 22, 23 in the latter group (Fig. 3B and C). The mean was  $23.4 \pm 2.5$  in the leiomyoma group, being significantly longer than that of the reference group ( $20.7 \pm 2.9$ ,  $P < 0.01$ ; Table II).

The relationship between tumor numbers and the CAG STR length was also assessed in 98 patients with uterine leiomyomas. Mean of the  $n$  values was determined to be  $23.1 \pm 2.2$  (19-29),  $23.3 \pm 2.8$  (18-31) and  $23.8 \pm 2.2$  (17-28), respectively, in cases with single, 2 or 3 and  $> 3$  tumor nodules. Frequencies of longer CAG repeats ( $n > 22$ ) were 58.1% (18/31), 58.1% (25/43) and 71.6% (14/24) in these 3 groups, respectively. The patients with  $> 3$  tumor nodules tend to have a longer CAG STR at the AR gene than the former two groups, but the differences did not attain statistical significance ( $P > 0.05$ ).

The unicentric and multicentric types of multinodular uterine leiomyomas were proposed recently to develop in different mechanisms (24) and  $n$  values of these patients were also assessed in this study (Fig. 4). Mean of the  $n$  values was  $22.2 \pm 2.5$  (17-27) in the unicentric cases and  $24.1 \pm 2.1$  (21-28) in the multicentric cases, and their frequencies with longer CAG repeats ( $n > 22$ ) estimated to be 40% (10/25) and 65.7% (16/23), respectively. The cases with multicentric type of multiple uterine leiomyomas were shown to have a longer CAG repeat than those with unicentric, multinodular lesions ( $P = 0.007$ ), the difference between the unicentric type of multiple leiomyomas and the solitary tumor did not attain statistical significance ( $P > 0.05$ ).

## Discussion

AR gene is composed of 75,000 to 90,000 nucleotides, including 8 exons (26). The encoded protein contains 910 to 919 amino acid residues, consisting of a NH<sub>2</sub>-terminal transactivation domain, a central DNA-binding domain, a nuclear-localizing short hinge region and a COOH-terminal ligand-binding domain (27,28). The transactivation domain is

encoded by exon 1. In this region, several polymorphic structures have been demonstrated to be responsible for variation of the transcription-activating activity of AR, among which the polyglutamine tract encoded by the CAG STR is believed to be most important (12,29). Functional studies demonstrated an inverse relationship between size of the polyglutamine structure and transactivation capacity of AR (30-32), that the CAG STR polymorphism may exert influences on some physiological and pathological processes by modulating AR activity.

Recent surveys have associated variations of the CAG repeat numbers to several clinical phenotypes including some reproductive, metabolic disorders (33,34) and cancer formation in several sites (12,29,34). Expansion of its size, but still within the normal range ( $n < 37$ ), was linked to reduced androgenic function (35,36). Markedly increased  $n$  values of the CAG repeat (40 to 52) have been causally linked to spinal-bulbar muscular atrophy, a neurodegenerative disorder that only affects males (37). On contrary, reduced  $n$  values have been associated with clinical features similar to that observed with increased androgenic function (29). In addition, the CAG repeat length polymorphism was associated to development of several malignancies. Males carrying a shorter CAG STR were found to have an increased risk to a few neoplasms including prostate (12,14,15,38) and hepatocellular carcinomas (20), while a longer CAG STR was associated to cancer development in the breast (39-42) and endometrium (43,44). It seems to be reasonable that relatively reduced sizes of the CAG repeat structure may confer individuals increased susceptibility to some male or male-predominant disorders and expansion of its size may result in greater predisposition to some female or female-predominant diseases including mammary and genital tract tumors.

A marked male predilection is found in pulmonary, esophageal and bladder carcinomas (1-3). While pathogenesis of these cancers is not fully understood, cigarette smoking is believed to be one of the causative factors for all of them (45-48). Data of haplotype analysis from South Africa indicated increased susceptibility to esophageal cancer in individuals with a (CAG)  $\leq 21$ /(GGC)  $\leq 16$  pattern, but they failed to show an association between CAG repeat numbers and cancer development (49). In the present study, we assessed the CAG STR length in 183 male patients with esophageal carcinoma, revealed a significantly smaller  $n$  value in cancer patients (mean, 20.0) compared to that of male reference group (mean, 21.1). A possible link was described in a recent survey between the CAG and GGN repeat length polymorphism of AR exon 1 and urothelial carcinoma in bladder (50), but the results were contradictory. In our study, the CAG STR length was assessed in 95 patients with bladder carcinoma, with their mean determined to be 19.8, being 1.3 smaller than that of the reference group. Our results demonstrate that a shorter CAG STR confers individuals a greater risk to esophageal and bladder carcinomas. In addition, we also measured the CAG ATR length in male and female patients with pulmonary carcinoma, with their  $n$  values determined to be  $\sim 1$  nucleotide triplet smaller than those of the male and female reference groups. Our study, for the first time, provided evidence for the involvement of the CAG repeat polymorphism in lung cancer development in

both men and women. These results may partly explain the marked male predilection of these three malignancies.

Uterine leiomyoma is one of the most common female genital tract tumors. The lifetime risk for a woman to the tumor may exceed 60% (51). Although its pathogenesis has not been elucidated, it is well known that its growth is dependent on ovarian hormones, particularly estradiol (52). Tumor regression occurs in most of the cases after menopause, and tumor growth was enhanced in these patients receiving hormone replacement therapy with estrogen analogues (53,54). Administration of oral contraceptives containing progestin may counterbalance the effect of estrogens and result in growth inhibition of the tumor (55). However, a growth-promoting effect was also observed in leiomyoma cells treated with progestin (56). The difference in the actions of progesterones on uterine leiomyoma may be explained by the ratio of progesterone receptor (PR) A and B (57). High-level expression of estrogen receptor (ER) and PR is demonstrated in the leiomyoma tissue and myometrium of the uterus (55,58-60), confirming the view that dysfunction of ER- and PR-mediated signal pathways is involved in development and progression of the tumor. AR expression was also demonstrated in uterine leiomyoma (58-60), indicating that AR-mediated functions may also play a role in uterine leiomyogenic tumorigenesis.

Family clustering was noted in some patients with uterine leiomyoma (61-63), indicating presence of some inheritable factors in development of the neoplasm. The familial and non-familial cases were shown to differ in some clinical phenotypes, with multiple leiomyomas found preferentially in the former group and solitary lesions in the latter (62-64). Polymorphisms of ER (65,66) and PR genes (67,68) have been linked to susceptibility to uterine leiomyoma. Two research groups also assessed *n* values of the CAG repeat structure in these patients, but they failed to associate the polymorphism to the tumor development (69,70). The current study revealed longer CAG repeat structures in patients with uterine leiomyomas, demonstrating an inverse relationship between the CAG STR length and susceptibility to uterine leiomyomas. These results, combined with the previous observations on ER and PR polymorphism (65-68), may provide a possible explanation for the family clustering of the tumor. A longer CAG STR in AR exon 1 may predispose females a greater risk to uterine leiomyomas, particularly to occurrence of multiple leiomyomas.

Our previous study provided an approach to divide multiple leiomyomas of the uterus, based on their clonal origins, into multicentric and unicentric lesions (24). In the former type, multiple lesions were proven to develop independently. In the latter type, multiple tumor nodules were revealed to develop from a parent tumor by neoplastic cell spreading. In the present study, a difference was demonstrated in the CAG repeat numbers between these two types of cases. The CAG STR in patients with multicentric lesions was found to be longer, by a factor of 1.9 in the *n* value, compared to those with unicentric lesions. These results provide further evidence for the presence of different types of multinodular uterine leiomyomas, which may be associated with some inheritable traits including distinct genotypes of these hormonal receptors and may show different biological behavior.

## Acknowledgements

The authors thank Professors Qinjing Pan, Jun Qi and Zhihua Liu for helpful discussions, Drs Lin Yang and Wenting Huang for statistic assistance, Guolian Wei, Xinhua Xue, Lei Guo, Fang Ding and Yongqiang Xie for their technical assistance. The project was supported by the National Natural Science Foundation of China, Grants 30171052, 30572125 and 30772508 and the CAMS Cancer Hospital Clinical Research Project LC2007A21.

## References

1. Parkin DM, Bray F, Ferlay J and Pisani P: Global cancer statistics, 2002. *CA Cancer J Clin* 55: 74-108, 2005.
2. Pisani P, Parkin D, Bray F and Ferlay J: Estimates of the worldwide incidence of 25 major cancers in 1990. *Int J Cancer* 80: 18-29, 1999.
3. Gabbert HE, Shimoda T and Hainaut P: Squamous cell carcinoma. In: *World Health Organization Classification. Pathology and Genetics of Tumors of The Digestive System*. Hamilton SR and Aaltonen LA (eds). IARC Press, Lyon, pp11-19, 2000.
4. Fairlamb D and Boesen E: Gynaecomastia associated with gonadotrophin-secreting carcinoma of the lung. *Br Med J* 53: 269-271, 1977.
5. Kaiser U, Hofmann J, Schilli M, Wegmann B, Klotz U, Wedel S, Virmani AK, Wollmer E, Branscheid D, Gazdar AF and Havemann K: Steroid-hormone receptors in cell lines and tumor biopsies of human lung cancer. *Int J Cancer* 67: 357-364, 1996.
6. Tihan T, Harmon JW, Wan X, Younes Z, Nass P, Duncan KL and Duncan MD: Evidence of androgen receptor expression in squamous and adenocarcinoma of the esophagus. *Anticancer Res* 21: 3107-3114, 2001.
7. Miyamoto H, Yang Z, Chen YT, Ishiguro H, Uemura H, Kubota Y, Nagashima Y, Chang YJ, Hu YC, Tsai MY, Yeh S, Messing EM and Chang C: Promotion of bladder cancer development and progression by androgen receptor signals. *J Natl Cancer Inst* 99: 558-568, 2007.
8. Lubahn DB, Joseph DR, Sullivan PM, Willard HF, French FS and Wilson EM: Cloning of human androgen receptor complementary DNA and localization to the X chromosome. *Science* 240: 327-330, 1988.
9. Chang CS, Kokontis J and Liao ST: Molecular cloning of human and rat complementary DNA encoding androgen receptors. *Science* 240: 324-326, 1988.
10. Chamberlain NL, Driver ED and Miesfeld RL: The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Res* 22: 3181-3186, 1994.
11. Claessens F, Denayer S, Van Tilborgh N, Kerkhofs S, Helsen C and Haelens A: Diverse roles of androgen receptor (AR) domains in AR-mediated signaling. *Nucl Recept Signal* 6: e008, 2008.
12. Ferro P, Catalano MG, Dell'Eva R, Fortunati N and Pfeffer U: The androgen receptor CAG repeat: a modifier of carcinogenesis? *Mol Cell Endocrinol* 193: 109-120, 2002.
13. Alevizaki M, Cimponeriu AT, Garofallaki M, Sarika HL, Alevizaki CC, Papamichael C, Philippou G, Anastasiou EA, Lekakis JP and Mavrikakis M: The androgen receptor gene CAG polymorphism is associated with the severity of coronary artery disease in men. *Clin Endocrinol* 59: 749-755, 2003.
14. Hardy DO, Scher HI, Bogenreider T, Sabbatini P, Zhang ZF, Nanus DM and Catterall JF: Androgen receptor CAG repeat lengths in prostate cancer: correlation with age of onset. *J Clin Endocrinol Metab* 81: 4400-4405, 1996.
15. Giovannucci E, Stampfer MJ, Krithivas K, Brown M, Dahl D, Brufsky A, Talcott J, Hennekens CH and Kantoff PW: The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci USA* 94: 3320-3323, 1997.
16. Balic I, Graham ST, Troyer DA, Higgins BA, Pollock BH, Johnson-Pais TL, Thompson IM and Leach RJ: Androgen receptor length polymorphism associated with prostate cancer risk in Hispanic men. *J Urol* 168: 2245-2248, 2002.
17. Santarosa M, Bidoli E, Gallo A, Steffan A, Boiocchi M and Viel A: Polymorphic CAG repeat length within the androgen receptor gene: identification of a subgroup of patients with increased risk of ovarian cancer. *Oncol Rep* 9: 639-644, 2002.



Baldwin RL and Karlan BY: Short androgen receptor length is a poor prognostic factor in epithelial ovarian carcinoma. *Clin Cancer Res* 9: 3667-3673, 2003.

19. Schildkraut JM, Murphy SK, Palmieri RT, Iversen E, Moorman PG, Huang Z, Halabi S, Calingaert B, Gusberg A, Marks JR and Berchuck A: Trinucleotide repeat polymorphisms in the androgen receptor gene and risk of ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 16: 473-480, 2007.
20. Yu MW, Cheng SW, Lin MW, Yang SY, Liaw YF, Chang HC, Hsiao TJ, Lin SM, Lee SD, Chen PJ, Liu CJ and Chen CJ: Androgen-receptor gene CAG repeats, plasma testosterone levels, and risk of hepatitis B-related hepatocellular carcinoma. *J Natl Cancer Inst* 92: 2023-2028, 2000.
21. Suter NM, Malone KE, Daling JR, Doody DR and Ostrander EA: Androgen receptor (CAG)<sub>n</sub> and (GGC)<sub>n</sub> polymorphisms and breast cancer risk in a population-based case-control study of young women. *Cancer Epidemiol Biomarkers Prev* 12: 127-135, 2003.
22. Sasaki M, Dahiya R, Fujimoto S, Ishikawa M and Oshimura M: The expansion of the CAG repeat in exon 1 of the human androgen receptor gene is associated with uterine endometrial carcinoma. *Mol Carcinog* 27: 237-244, 2000.
23. Houston KD, Hunter DS, Hodges LC and Walker CL: Uterine leiomyomas: mechanisms of tumorigenesis. *Toxicol Pathol* 29: 100-104, 2001.
24. Cai Y-R, Diao X-L, Wang S-F, Zhang W, Zhang H-T and Su Q: X-chromosomal inactivation analysis of uterine leiomyomas reveals a common clonal origin of different tumor nodules in some multiple leiomyomas. *Int J Oncol* 31: 1379-1389, 2007.
25. Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM and Belmont JW: Methylation of *Hpa* II and *Hha* I sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet* 51: 1229-1239, 1992.
26. Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM and French FS: Androgen receptor defects: historical, clinical, and molecular perspectives. *Endocr Rev* 16: 271-321, 1995.
27. Bain DL, Heneghan AF, Connaghan-Jones KD and Miura MT: Nuclear receptor structure: implications for function. *Annu Rev Physiol* 69: 201-220, 2006.
28. Dehm SM and Tindall DJ: Androgen receptor structural and functional elements: role and regulation in prostate cancer. *Mol Endocrinol* 21: 2855-2863, 2007.
29. Palazzolo I, Gliozzi A, Rusmini P, Sau D, Crippa V, Simonini F, Onesto E, Bolzoni E and Poletti A: The role of the polyglutamine tract in androgen receptor. *J Steroid Biochem Mol Biol* 108: 245-253, 2008.
30. Mhatre AN, Trifiro MA, Kaufman M, Kazemi-Esfarjani P, Figlewicz D, Rouleau G and Pinsky L: Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy. *Nat Genet* 5: 184-188, 1993.
31. Kazemi-Esfarjani P, Trifiro MA and Pinsky L: Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenetic relevance for the (CAG)<sub>n</sub>-expanded neuropathies. *Hum Mol Genet* 4: 523-527, 1995.
32. Tut TG, Ghadessy FJ, Trifiro MA, Pinsky L and Yong EL: Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. *J Clin Endocrinol Metab* 82: 3777-3782, 1997.
33. Zitzmann M and Nieschlag E: The CAG repeat polymorphism within the androgen receptor gene and maleness. *Int J Androl* 26: 76-83, 2003.
34. Rajender S, Singh L and Thangaraj K: Phenotypic heterogeneity of mutations in androgen receptor gene. *Asian J Androl* 9: 147-179, 2007.
35. Patrizio P, Leonard DG, Chen KL, Hernandez-Ayup S and Trounson AO: Larger trinucleotide repeat size in the androgen receptor gene of infertile men with extremely severe oligozoospermia. *J Androl* 22: 444-448, 2001.
36. Katagiri Y, Neri QV, Takeuchi T, Moy F, Sills ES and Palermo GD: Androgen receptor CAG polymorphism (Xq11-12) status and human spermatogenesis: a prospective analysis of infertile males and their offspring conceived by intracytoplasmic sperm injection. *Int J Mol Med* 18: 405-413, 2006.
37. La Spada AR, Roling DB, Harding AE, Warner CL, Spiegel R, Hausmanowa-Petrusewicz I, Yee WC and Fischbeck KH: Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Nat Genet* 2: 301-304, 1992.
38. Stanford JL, Just JJ, Gibbs M, Wicklund KG, Neal CL, Blumenstein BA and Ostrander EA: Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res* 57: 1194-1198, 1997.
39. Rebbeck TR, Kantoff PW, Krithivas K, Neuhausen S, Blackwood MA, Godwin AK, Daly MB, Narod SA, Garber JE, Lynch HT, Weber BL and Brown M: Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. *Am J Hum Genet* 64: 1371-1377, 1999.
40. Giguere Y, Dewailly E, Brisson J, Ayotte P, Laflamme N, Demers A, Forest VI, Dodin S, Robert J and Rousseau F: Short polyglutamine tracts in the androgen receptor are protective against breast cancer in the general population. *Cancer Res* 61: 5869-5874, 2001.
41. Liede A, Zhang W, De Leon Matsuda ML, Tan A and Narod SA: Androgen receptor gene polymorphism and breast cancer susceptibility in The Philippines. *Cancer Epidemiol Biomarkers Prev* 12: 848-852, 2003.
42. González A, Javier DF, Rodríguez G, Brito B, Rodríguez MA, Cabrera A, Díaz-Chico JC, Reyes R, Aguirre-Jaime A and Nicolás Díaz-Chico B: Increased risk of breast cancer in women bearing a combination of large CAG and GGN repeats in the exon 1 of the androgen receptor gene. *Eur J Cancer* 43: 2373-2380, 2007.
43. Yaron M, Levy T, Chetrit A, Levavi H, Sabah G, Schneider D, Halperin R, Ben-Rafael Z and Friedman E: The polymorphic CAG repeat in the androgen receptor gene in Jewish Israeli women with endometrial carcinoma. *Cancer* 92: 1190-1194, 2001.
44. Sasaki M, Sakuragi N and Dahiya R: The CAG repeats in exon 1 of the androgen receptor gene are significantly longer in endometrial cancer patients. *Biochem Biophys Res Commun* 305: 1105-1108, 2003.
45. Momas I, Daures J, Festy B, Bontoux J and Gremy F: Bladder cancer and black tobacco cigarette smoking. Some results from a French case-control study. *Eur J Epidemiol* 10: 599-604, 1994.
46. Castela JE, Yuan JM, Skipper PL, Tannenbaum SR, Gago-Dominguez M, Crowder JS, Ross RK and Yu MC: Gender- and smoking-related bladder cancer risk. *J Natl Cancer Inst* 93: 538-545, 2001.
47. Boffetta P: Tobacco smoking and risk of bladder cancer. *Scand J Urol Nephrol (Suppl)*: 45-54, 2008.
48. Pelucchi C, Gallus S, Garavello W, Bosetti C and La Vecchia C: Alcohol and tobacco use, and cancer risk for upper aerodigestive tract and liver. *Eur J Cancer Prev* 17: 340-344, 2008.
49. Dietzsch E, Laubscher R and Parker MI: Esophageal cancer risk in relation to GGC and CAG trinucleotide repeat lengths in the androgen receptor gene. *Int J Cancer* 107: 38-45, 2003.
50. Liu CH, Huang JD, Huang SW, Hour TC, Huang YK, Hsueh YM, Chiou HY, Lee TC, Jan KY, Chen CJ and Pu YS: Androgen receptor gene polymorphism may affect the risk of urothelial carcinoma. *J Biomed Sci* 15: 261-269, 2008.
51. Okolo S: Incidence, aetiology and epidemiology of uterine fibroids. *Best Pract Res Clin Obstet Gynaecol* 22: 571-588, 2008.
52. Blake RE: Leiomyomata uteri: hormonal and molecular determinants of growth. *J Natl Med Assoc* 99: 1170-1184, 2007.
53. Frigo P, Eppel W, Asseryanis E, Sator M, Golaszewski T, Gruber D, Lang C and Huber J: The effects of hormone substitution in depot form on the uterus in a group of 50 perimenopausal women, a vaginosonographic study. *Maturitas* 21: 221-225, 1995.
54. Sener AB, Seckin NC, Ozmen S, Gokmen O, Dogu N and Ekici E: The effects of hormone replacement therapy on uterine fibroids in postmenopausal women. *Fertil Steril* 65: 354-357, 1996.
55. Englund K, Blanck A, Gustavsson I, Lundkvist U, Sjöblom P, Norgren A and Lindblom B: Sex steroid receptors in human myometrium and fibroids: changes during the menstrual cycle and gonadotropin-releasing hormone treatment. *J Clin Endocrinol Metab* 83: 4092-4096, 1998.
56. Hoekstra AV, Sefton EC, Berry E, Lu Z, Hardt J, Marsh E, Yin P, Clardy J, Chakravarti D, Bulun S and Kim JJ: Progesterins activate the AKT pathway in leiomyoma cells and promote survival. *J Clin Endocrinol Metab* 94: 1768-1774, 2009.
57. Ying Z and Weiyuan Z: Dual actions of progesterone on uterine leiomyoma correlate with the ratio of progesterone receptor A:B. *Gynecol Endocrinol* 25: 520-523, 2009.
58. Tamaya T, Fujimoto J and Okada H: Comparison of cellular levels of steroid receptors in uterine leiomyoma and myometrium. *Acta Obstet Gynecol Scand* 64: 307-309, 1985.

59. Fujimoto J, Nishigaki M, Hori M, Ichigo S, Itoh T and Tamaya T: The effect of estrogen and androgen on androgen receptors and mRNA levels in uterine leiomyoma, myometrium and endometrium of human subjects. *J Steroid Biochem Mol Biol* 50: 137-143, 1994.
60. Leitao MM, Soslow RA, Nonaka D, Olshen AB, Aghajanian C, Sabbatini P, Dupont J, Hensley M, Sonoda Y, Barakat RR and Anderson S: Tissue microarray immunohistochemical expression of estrogen, progesterone, and androgen receptors in uterine leiomyomata and leiomyosarcoma. *Cancer* 101: 1455-1462, 2004.
61. Van Voorhis BJ, Romitti PA and Jones MP: Family history as a risk factor for development of uterine leiomyomas. Results of a pilot study. *J Reprod Med* 47: 663-669, 2002.
62. Sato F, Mori M, Nishi M, Kudo R and Miyake H: Familial aggregation of uterine myomas in Japanese women. *J Epidemiol* 12: 249-253, 2002.
63. Okolo SO, Gentry CC, Perrett CW and Maclean AB: Familial prevalence of uterine fibroids is associated with distinct clinical and molecular features. *Hum Reprod* 20: 2321-2324, 2005.
64. Uimari O, Suomalainen-König S, Sakkinen N, Santala M, Nieminen P and Ryyanen M: Natural history of familial myomas. *Eur J Obstet Gynecol Reprod Biol* 125: 255-258, 2006.
65. Villanova FE, Andrade PM, Otsuka AY, Gomes MT, Leal ES, Castro RA, Girão MJ, Nishimura E, Barakat EC and Silva ID: Estrogen receptor alpha polymorphism and susceptibility to uterine leiomyoma. *Steroids* 7: 11-12, 2006.
66. Hsieh YY, Wang YK, Chang CC and Lin CS: Estrogen receptor alpha-351 *Xba* I\*G and -397 *Pvu* II\*C-related genotypes and alleles are associated with higher susceptibilities of endometriosis and leiomyoma. *Mol Hum Reprod* 13: 117-122, 2007.
67. Gomes MT, Castro Rde A, Villanova FE, da Silva ID, Barakat EC, de Lima GR and Girão MJ: The progesterone receptor gene polymorphism, PROGINS, may be a factor related to the development of uterine fibroids. *Fertil Steril* 87: 1116-1121, 2007.
68. Renner SP, Strick R, Fasching PA, Oeser S, Oppelt P, Mueller A, Beckmann MW and Strissel PL: Single nucleotide polymorphisms in the progesterone receptor gene and association with uterine leiomyoma tumor characteristics and disease risk. *Am J Obstet Gynecol* 199: 648.e1-648.e9, 2008.
69. Hsieh YY, Chang CC, Tsai FJ, Lin CC, Yeh LS and Peng CT: Androgen receptor trinucleotide polymorphism in leiomyoma. *J Assist Reprod Genet* 21: 453-457, 2004.
70. Rosa FE, Canevari Rde A, Ambrosio EP, Ramos Cirilo PD, Pontes A, Rainho CA and Rogatto SR: Polymorphisms of CYP17A1, CYP19, and androgen in Brazilian women with uterine leiomyomas. *Clin Chem Lab Med* 46: 814-823, 2008.