# Polymorphic CAG repeat in the androgen receptor gene in polycystic ovary syndrome patients

YANJIE XIA<sup>1</sup>, YENA CHE<sup>1</sup>, XINLIN ZHANG<sup>1</sup>, CHENGWEI ZHANG<sup>1</sup>, YUNXIA CAO<sup>2</sup>, WENJUN WANG<sup>3</sup>, PEI XU<sup>1</sup>, XIAOKE WU<sup>4</sup>, LONG YI<sup>1</sup>, QIAN GAO<sup>1</sup> and YONG WANG<sup>1</sup>

<sup>1</sup>Center for Translational Medicine and Jiangsu Key Laboratory of Molecular Medicine, Medical School of Nanjing University,

Nanjing 210093; <sup>2</sup>Department of Obstetrics and Gynecology, Anhui Medical University, Hefei 230022;

<sup>3</sup>Centre of Reproduction, Department of Obstetrics and Gynecology, Memorial Hospital of Sun Yat-Sen University,

Guangzhou 510120; <sup>4</sup>Department of Obstetrics and Gynecology, The First Affiliated Hospital,

Heilongjiang University of Chinese Medicine, Harbin 150040, P.R. China

Received November 24, 2011; Accepted February 2, 2012

DOI: 10.3892/mmr.2012.789

Abstract. Human androgen receptor (AR) contains a highly polymorphic polyglutamine tract encoded by CAG repeats [(CAG)<sub>n</sub>] in exon 1 of the AR gene. The CAG repeats, ranging from 11 to 38, have been reported to be inversely correlated with AR activity. A case-control study involving 261 polycystic ovary syndrome (PCOS) patients and 278 healthy controls was conducted. Fluorescently labeled DNA fragments containing (CAG)<sub>n</sub> were obtained by PCR and genotyped via capillary electrophoresis. AR (CAG)<sub>n</sub> ranges were 6, 12-28 in PCOS cases and 9, 10, 12-32 in controls. In the PCOS group, a higher frequency of short (CAG)<sub>n</sub> alleles was found compared with that of controls (P=0.007). Similarly, CAG biallelic mean distributions also showed statistical difference between the two groups (P=0.025). In conclusion, shorter alleles of the (CAG)<sub>n</sub> in exon 1 of the AR gene enhanced the susceptibility to PCOS, either by upregulating AR activity or by causing hyperandrogenism.

#### Introduction

Polycystic ovary syndrome (PCOS), characterized by hyperandrogenism, chronic anovulation and oligo-ovulation, affects an estimated 7.4% of women of reproductive age in China (1-3). The pathogenesis of PCOS remains unknown, but there are increasing implications for a strong genetic basis (4,5). Familial aggregation of hyperandrogenic symptoms (hirsutism and oligomenorrhea) was found in a series of family-based studies in patients presenting with hyperandrogenism or PCOS (6). Hyperandrogenism is regarded as one of the causal factors of PCOS, and common variants of androgen-related genes may influence the activity of androgen. Androgen exerts its action primarily through the androgen receptor (AR), a liganddependent nuclear transcription factor. In vitro studies suggest that androgen promotes the development of follicles (7,8), which is mainly governed by ARs via the upregulation of genes encoding follicle-stimulating hormone receptor, insulin-like growth factor (IGF-1) and IGF-1 receptor (9,10). Numerous studies have concentrated on a various number tandem repeat (VNTR) polymorphism consisting of CAG repeats [(CAG)<sub>n</sub>] in exon 1 of the AR gene located at Xq11-12. The (CAG)<sub>n</sub> encodes for a polyglutamine tract located in the transactivation domain of AR. Early studies suggest that the number of CAG repeats is inversely related to AR activity (11); thus a shorter CAG track is more common in patients with ovarian hyperandrogenism (11,12) and hirsutism (12). Higher activity of AR induces the development of follicles, and thus may cause the formation of PCOS; meanwhile, AR hyperactivity may explain some of the PCOS patients exhibiting normal serum androgen levels but hyperandrogenism symptoms (13). In this study, we hypothesized that shorter CAG tracks are associated with PCOS, and designed a case-control study to confirm.

#### Materials and methods

Subjects and methods. The study population included 261 patients with PCOS [age,  $28.06\pm6.48$  (SD) years] and 278 healthy control women (age,  $33.70\pm5.57$  years) All of these women were of Han Chinese origin, recruited either from the Memorial Hospital of Sun Yat-Sen University or the First Affiliated Hospital, Heilongjiang University of Chinese Medicine. PCOS was diagnosed according to the 2003 Rotterdam criteria (14), as described elsewhere (15), simply with the existence of two of the following three criteria: i) oligo and/or anovulation; ii) clinical and/or biochemical signs of hyperandrogenism; iii) polycystic ovaries, and excluding other reasons of hyperandrogenism such as congenital adrenal hyperplasia, androgen-secreting tumors, or Cushing's syndrome. All controls had normal menstrual cycles with no

*Correspondence to:* Dr Yong Wang, Center for Translational Medicine and Jiangsu Key Laboratory of Molecular Medicine, Medical School of Nanjing University, Nanjing 210093, P.R. China E-mail: yongwang@nju.edu.cn

Key words: polycystic ovary syndrome, androgen receptor, CAG repeat

signs of hirsutism, acne or endocrine dysfunction. The study was approved by the Medical School of Nanjing University, and informed consent was obtained from all the women prior to inclusion.

*Clinical and biochemical determination.* Age was acquired through inquiry. Body weight (BW) and height (H) were directly measured and body mass index (BMI) was calculated as BW/H<sup>2</sup> (kg/m<sup>2</sup>). After a 12-h overnight fast, all venous blood samples were obtained between 8 and 9 a.m. For those who had menstruation, the sample-collecting time was further restricted to the 3rd-5th day of the menstrual cycle. Estradiol (E2), testosterone (T), follicle-stimulating hormone (FSH), and luteinizing hormone (LH) levels in the sera were detected by radioimmunoassay (RIA, Beijing North Institute of Biological Technology of China and the CIS Company of France). All assays had intra- and inter-assay coefficients of variation <10%.

Genotyping of the (CAG)<sub>n</sub>. Genomic DNA was extracted from blood using the SBS UltraPure<sup>™</sup> Genome DNA kit (SBS Genetech, Shanghai, China) according to the manufacturer's instructions. PCR was carried out in a total volume of 25  $\mu$ l consisting of 50 ng DNA, 7.5 pmol of each primer,  $2.5 \mu$ l STR (short tandem repeat) 10X buffer (Promega, Madison, WI, USA) and 0.75 unit of GoTaqDNA polymerase (Promega). The forward primer was 5'-ACC GAG GAG CTT TCC AGA AT-3', the reverse primer was 5'-CTC ATC CAG GAC CAG GTA GC-3' and labeled at the 5' with 6-FAM, a laser-activated fluorescent dye. All PCR reactions were performed with a MultiGene Gradient Thermal Cycler (Labnet, Edison, NJ, USA) under the following conditions: an initial denaturating step at 96°C for 2 min, followed by 10 cycles of denaturation at 94°C for 1 min, annealing at 60°C for 1 min and extension at 70°C for 1.5 min, followed by another 25 cycles of denaturation at 90°C for 1 min, annealing at 60°C for 1 min and extension at 70°C for 1.5 min, and a final extension at 72°C for 30 min. PCR products were then 5-fold diluted with deionized water, and 1 µl was mixed with 8.7 µl HI-DI formamide and 0.3 µl Genescan-500 (LIZ) (Applied Biosystems, Warrington, UK). After denaturing at 95°C for 5 min, each sample was analyzed by capillary electrophoresis on an ABI 3130 Genetic Analyzer (Applied Biosystems). The results were analyzed in GeneMapper version 3.7 (Applied Biosystems) (Fig. 1), and PCR product sizes were exported and used for calculating the repeat number of CAG in the AR gene. Finally, a subset was directly sequenced to confirm the exact repeat size.

Setting categories. We divided AR  $(CAG)_n$  into different categories according to CAG repeat number. By setting 20, the mean CAG repeat of controls, as the cut-off value, we obtained three alleles, <20 (short allele), 20 (mean allele) and >20 (long allele), represented by the letter 'S' 'M' 'L', respectively. To further analyze the CAG distribution in the PCOS population, we divided PCOS women into 'obese case' and 'lean case' categories by taking BMI 23 as the borderline (16), different from the World Health Organization (WHO) criteria.

Statistical methods. The results of serum hormone levels, age, BMI are reported as means  $\pm$  SD. The two-sample inde-

Table I. Clinical and endocrine characteristics of the PCOS patients and controls.

	Control	PCOS
Number	278	261
Age (years) <sup>a</sup>	33.7±5.57	28.06±6.48
BMI (kg/m <sup>2</sup> ) <sup>a</sup>	22.17±3.21	23.06±4.04
FSH (IU/l) <sup>a</sup>	7.18±2.47	5.94±2.41
LH (IU/l) <sup>a</sup>	5.87±5.58	$10.69 \pm 5.49$
E2 (pg/ml) <sup>a</sup>	44.59±31.12	59.27±29.77
T (nmol/l) <sup>a</sup>	2.17±2.98	$3.00 \pm 2.75$

<sup>a</sup>P<0.05 compared with control groups. All data are reported as mean ± SD. BMI, body mass index; FSH, follicle-stimulating hormone; LH, luteinizing hormone; E2, estradiol; T, testosterone.



Figure 2. Distribution of CAG alleles in PCOS cases and controls. A total of 261 PCOS patients and 278 controls were genotyped in this study. Each subject contributed two alleles (paternal and maternal) to the histogram. (P=0.283437). PCOS, polycystic ovary syndrome.

pendent t-test was used to compare clinical and biochemical characteristics between the women with PCOS and controls. The CAG allelic distributions between PCOS and controls were compared by  $\chi^2$  test. Genotypic distribution among three newly formed alleles (S, M and L) was analyzed by using Fisher's exact test. Biochemical steroid levels among different genotypes were tested by analysis of covariance (ANCOVA) to correct for age and BMI. P<0.05 was considered statistically significant. All statistic analyses were performed using SPSS version 17.0 (SPSS Inc., Chicago, IL, USA).

#### Results

*Clinical and endocrine characteristics*. Clinical and endocrine characteristics of the PCOS patients were consistent with Rotterdam criteria (Table I) (Fig. 1).

*Distribution of CAG alleles in the samples.* Androgen receptor CAG repeats ranged from 12-28 in PCOS cases and 12-32 in



Figure 1. The AR gene  $(CAG)_n$  size obtained after capillary electrophoresis. The GeneScan-500-Liz size standard is indicated by arrows, and target  $(CAG)_n$  DNA fragments are indicated by arrowheads. AR, and rogen receptor.

Table II. Distribution of PCOS cases and controls according to the qualitative categories of CAG repeats.

(CAG) <sub>n</sub> length	Total number of CAG alleles			CAG biallelic mean		
	PCOS (2n=522)	Controls (2n=556) n (%)	P-value <sup>a</sup>	Cases (n=261) n (%)	Controls (n=278) n (%)	P-value <sup>b</sup>
<20	195 (37.4)	160 (28.8)		105 (40.2)	81 (29.1)	
20	102 (19.5)	111 (20.0)	0.007	26 (10.0)	31 (11.2)	0.025
>20	225 (43.1)	285 (51.3)		130 (49.8)	166 (59.7)	

<sup>a</sup>Based on allelic frequencies vs. controls; <sup>b</sup>Based on biallelic mean frequencies vs. control. <sup>a,b</sup>From Fisher's exact test. All data are reported as mean ± SD.

the controls (Fig. 2). In addition, few extremely short alleles were detected in our study, with one  $(CAG)_6$  allele noted in PCOS patient, one  $(CAG)_9$  allele and three  $(CAG)_{10}$  alleles in the control group; no homozygous genotype of these alleles was observed (Fig. 2).

Distribution of cases and controls via qualitative categories of CAG repeats. The newly formed S-M-L allelic distribution was significantly different between the PCOS group and controls (P=0.007) (Table II), exhibiting a higher frequency of the 'S' allele and a lower frequency of the 'L' allele in the PCOS group. Similarly, CAG biallelic mean distributions also showed significant statistical difference between the two groups (P=0.025) (Table II). *Clinical and endocrine characteristics among different categories of samples.* Clinical and biochemical hormone levels of the three 'alleles' are shown in Table III, no statistical difference was found both in PCOS patients and controls.

*Mean CAG biallelic distribution among lean and obese PCOS cases.* When the PCOS group was divided into lean cases and obese cases, no significant difference was found in terms of allelic distribution (P=0.244) (Table IV).

### Discussion

In this study, we found a strong association between a shorter AR-CAG repeat length and a higher presence of PCOS,

	PCOS				Control			
	<20	20	>20	P-value	<20	20	>20	P-value
Age (years)	28.32±5.43	29.56±6.20	27.63±7.23	0.38298	33.84±5.24	34.69±6.46	33.46±5.60	0.70759
AAM (years)	13.85±1.23	13.84±1.34	14.23±1.85	0.51947	14.31±2.83	14.37±1.84	14.41±1.51	0.96918
BMI (kg/m <sup>2</sup> )	23.10±4.62	23.01±3.51	22.48±3.88	0.55877	22.01±3.06	22.23±2.81	21.55±2.81	0.56222
FSH (IU/l)	$5.88 \pm 2.52$	$5.58 \pm 2.20$	6.12±3.23	0.6676	$7.96 \pm 2.93$	6.67±1.88	7.51±2.62	0.27225
LH (IU/I)	10.36±5.90	10.04±4.71	11.64±9.56	0.45829	5.18±3.04	$5.55 \pm 2.40$	6.32±6.71	0.53864
LH/FSH	2.22±1.82	2.61±2.71	2.07±1.72	0.47533	0.71±0.48	0.86±0.50	0.84±0.67	0.50658
T (nmol/l)	3.25±5.25	2.32±1.31	3.39±7.96	0.79572	$2.40 \pm 2.04$	1.94±1.74	2.40±1.60	0.79397
E2 (pg/ml)	55.13±30.82	57.65±27.86	65.81±35.45	0.12018	57.01±30.81	41.49±22.41	49.87±27.91	0.20022

Table III. Anthropometric characteristics and serum hormone concentrations of PCOS cases and controls according to the qualitative categories of the x-CAG biallelic mean.

AAM, age at menarche; BMI, body mass index; FSH, follicle-stimulating hormone; LH, luteinizing hormone; T, testosterone; E2, estradiol.

Table IV. Qualitative distribution of the CAG biallelic mean among lean and obese PCOS cases.

	Total number of CA		
Mean (CAG) <sub>n</sub> length	Obese (2n=188) n (%)	Lean (2n=258) n (%)	P-value
<18 18-22 >22	11 (11.7) 62 (66.0) 21 (22.3)	12 (9.3) 98 (76.0) 19 (14.7)	0.244

Body mass index (BMI) equal to 23 was set as the cut-off value to define 'obese PCOS' and 'lean PCOS' cases. A total of 38 PCOS patients lacked clinical variables, BMI included.

providing a convincing evidence for the role AR (CAG)<sub>n</sub> plays in the pathogenesis of PCOS in Han Chinese women. This result indicates that short CAG allele repeats may be a potential risk for PCOS. Our results were consistent with previous functional studies demonstrating that a short CAG repeat resulted in high intrinsic activity of AR (17-19). Ibanez et al (20) found that the serum testosterone level was markedly reduced after adminstration of flutamide, an inhibitor of AR. Therefore, we inferred that increased AR activity enhances ovarian androgen production, resulting in a high level of testosterone in PCOS women. Similarly, other studies strongly support the hypothesis that PCOS phenotypes may result from enhanced sensitivity to androgens caused by a more active AR (21-23). The most striking evidence demonstrating AR (CAG)<sub>n</sub> as a leading causal factor of ovarian hyperandrogenism was demonstrated in a follow-up study on Spanish girls. Ibanez et al found that girls with premature pubarche and subsequent ovarian hyperandrogenism exhibited a shorter AR (CAG)<sub>n</sub> repeat length compared with healthy controls (12), while another investigator insisted that shorter CAG repeats heighten androgen levels only when combined with longer alleles of SHBG  $(TAAAA)_n$  pentanucleotide polymorphism (24).

On the contrary, other studies yielded divergent or even extremely opposing results. While most studies revealed no difference in the distribution of allelic frequencies or biallelic mean between PCOS cases and controls (23,25-27), Hickey *et al* (22) demonstrated that longer alleles were more common in PCOS patients. However, additional studies suggest that the AR (CAG)<sub>n</sub> polymorphism may be a modulator of androgenmediated phenotypes in certain individuals instead of a determinant factor of PCOS (28).

Several interacting factors may account for the discrepancies among studies. The first and most significant reason is due to the ethnic composition. The AR (CAG)<sub>n</sub> allele frequencies were different. Ackerman et al (29) revealed that potential ethnic differences might exist in androgenic pathway activity and androgen sensitivity after a large-sample study in four ethnic populations (29). Secondly, patients chosen were also different: i) a globally accepted diagnostic scheme for PCOS is not available. The 2003 Rotterdam criteria was used in this study, while other definitions of PCOS were also applied in other studies (21,22,28). ii) Different diseases with symptoms of hyperandrogenism were also studied, i.e. PCOS and precocious pubarche, and other factors specifically affecting one particular disease may affect the outcome of studies. Thirdly, lack of a large enough sample size could, to a certain extent, make some studies less credible.

In summary, short AR  $(CAG)_n$  variants may be a potential risk factor for PCOS. More systematic studies with large sample sizes are needed in the future to clarify the possible mechanisms involved.

#### Acknowledgements

We are extremely grateful to all the women who participated in this study. This study was supported by the National Natural Science Foundation of China (81170541), the Natural Basic Research Program of China (973 program 2010CB945103) and the Natural Science Foundation of Jiangsu Province (BK2009241).

## References

- 1. Ehrmann DA: Polycystic ovary syndrome. N Engl J Med 352: 1223-1236, 2005.
- 2. Franks S: Polycystic ovary syndrome. N Engl J Med 333: 853-861, 1995.
- 3. Qiao J: Prevalence of Polycystic Ovary Syndrome in China. Second International Symposium on PCOS: 174-175, 2008.
- 4. Xita N, Georgiou I and Tsatsoulis A: The genetic basis of polycystic ovary syndrome. Eur J Endocrinol 147: 717-725, 2002.
- Franks S, Gharani N, Waterworth D, et al: The genetic basis of polycystic ovary syndrome. Hum Reprod 12: 2641-2648, 1997.
- Givens JR: Familial polycystic ovarian disease. Endocrinol Metab Clin North Am 17: 771-783, 1988.
- Murray AA, Gosden RG, Allison V and Spears N: Effect of androgens on the development of mouse follicles growing in vitro. J Reprod Fertil 113: 27-33, 1998.
- 8. Wang H, Andoh K, Hagiwara H, *et al*: Effect of adrenal and ovarian androgens on type 4 follicles unresponsive to FSH in immature mice. Endocrinology 142: 4930-4936, 2001.
- 9. Weil S, Vendola K, Zhou J and Bondy CA: Androgen and folliclestimulating hormone interactions in primate ovarian follicle development. J Clin Endocrinol Metab 84: 2951-2956, 1999.
- Vendola K, Zhou J, Wang J and Bondy CA: Androgens promote insulin-like growth factor-I and insulin-like growth factor-I receptor gene expression in the primate ovary. Hum Reprod 14: 2328-2332, 1999.
- Carson-Jurica MA, Schrader WT and O'Malley BW: Steroid receptor family: structure and functions. Endocr Rev 11: 201-220, 1990.
- Ibanez L, Ong KK, Mongan N, *et al*: Androgen receptor gene CAG repeat polymorphism in the development of ovarian hyperandrogenism. J Clin Endocrinol Metab 88: 3333-3338, 2003.
- Legro RS, Shahbahrami B, Lobo RA and Kovacs BW: Size polymorphisms of the androgen receptor among female Hispanics and correlation with androgenic characteristics. Obstet Gynecol 83: 701-706, 1994.
- 14. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group: Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril 81: 19-25, 2004.
- 15. Zhang XL, Zhang CW, Xu P, Liang FJ, Che YN, Xia YJ, Cao YX, Wu XK, Wang WJ, Yi L, *et al*: SNP rs2470152 in CYP19 is correlated to aromatase activity in Chinese polycystic ovary syndrome patients. Mol Med Rep 5: 245-249, 2012.
- Steering Committee: The Asia-Pacific perspective: Redefining obesity and its treatment. International Diabetes Institute, Melbourne, 2000.

- Hsiao PW, Lin DL, Nakao R and Chang C: The linkage of Kennedy's neuron disease to ARA24, the first identified androgen receptor polyglutamine region-associated coactivator. J Biol Chem 274: 20229-20234, 1999.
- 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.
- Beilin J, Ball EM, Favaloro JM and Zajac JD: 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.
- 20. Ibanez L, Potau N, Marcos MV and de Zegher F: Treatment of hirsutism, hyperandrogenism, oligomenorrhea, dyslipidemia, and hyperinsulinism in nonobese, adolescent girls: effect of flutamide. J Clin Endocrinol Metab 85: 3251-3255, 2000.
- 21. Mifsud A, Ramirez S and Yong EL: Androgen receptor gene CAG trinucleotide repeats in anovulatory infertility and polycystic ovaries. J Clin Endocrinol Metab 85: 3484-3488, 2000.
- 22. Hickey T, Chandy A and Norman RJ: The androgen receptor CAG repeat polymorphism and X-chromosome inactivation in Australian Caucasian women with infertility related to polycystic ovary syndrome. J Clin Endocrinol Metab 87: 161-165, 2002.
- 23. Van Nieuwerburgh F, Stoop D, Cabri P, Dhont M, Deforce D and De Sutter P: Shorter CAG repeats in the androgen receptor gene may enhance hyperandrogenicity in polycystic ovary syndrome. Gynecol Endocrinol 24: 669-673, 2008.
- 24. Xita N, Georgiou I, Lazaros L, Psofaki V, Kolios G and Tsatsoulis A: The role of sex hormone-binding globulin and androgen receptor gene variants in the development of polycystic ovary syndrome. Hum Reprod 23: 693-698, 2008.
- 25. Ferk P, Perme MP, Teran N and Gersak K: Androgen receptor gene (CAG)<sub>n</sub> polymorphism in patients with polycystic ovary syndrome. Fertil Steril 90: 860-863, 2008.
- 26. Liu Q, Hong J, Cui B, *et al*: Androgen receptor gene CAG(n) trinucleotide repeats polymorphism in Chinese women with polycystic ovary syndrome. Endocrine 33: 165-170, 2008.
- Kim JJ, Choung SH, Choi YM, Yoon SH, Kim SH and Moon SY: Androgen receptor gene CAG repeat polymorphism in women with polycystic ovary syndrome. Fertil Steril 90: 2318-2323, 2008.
- Calvo RM, Asuncion M, Sancho J, San Millan JL and Escobar-Morreale HF: The role of the CAG repeat polymorphism in the androgen receptor gene and of skewed X-chromosome inactivation, in the pathogenesis of hirsutism. J Clin Endocrinol Metab 85: 1735-1740, 2000.
- Ackerman CM, Lowe LP, Lee H, *et al*: Ethnic variation in allele distribution of the androgen receptor (AR) (CAG)<sub>n</sub> repeat. J Androl: May 19, 2011 (E-pub ahead of print).