Associations between androgen receptor CAG & GGN repeat polymorphism & recurrent spontaneous abortions in Chinese women

Zhang Chuan¹,², Dang Jie¹,², Xu Hao³, Bao Junhua⁴, Guo Mengjing¹, Pei Liguó¹, Yan Yousheng¹,⁵, Lu Hong¹,² & Huo Zhenghao¹,²

¹Key Laboratory of Fertility Preservation & Maintenance, Ministry of Education, Key Laboratory of Reproduction & Heredity of Ningxia Hui Autonomous Region, Ningxia Medical University, Yinchuan, Ningxia, ²Department of Medical Genetics & Cell Biology, Ningxia Medical University, Yinchuan, Ningxia, ³School of Stomatology, Fourth Military Medical University, Xi’an, ⁴Yinchuan Maternal Child Health Care Hospital, Yinchuan, Ningxia & ⁵Gansu Provincial Maternity Child-Care Hospital, Gansu Province Medical Genetics Center, Lanzhou Gansu, PR China

Received December 18, 2012

Background & objectives: Recurrent spontaneous abortion (RSA) is a reproductive problem that occurs in women in reproductive age with a frequency of 1-3 per cent. Previous studies have reported high levels of serum androgens to be associated with RSAs. At the molecular level, the effect of androgens is mediated through the activation of the androgen receptor (AR). The CAG and GGN repeat polymorphisms of the AR gene are associated with the AR activity. We hypothesize that the AR CAG/GGN repeat polymorphism may be associated with levels of serum androgens. Thus, this study as undertaken to evaluate the relationship between CAG/GGN repeats in exon 1 of the AR gene in women with RSAs.

Methods: This case-control study was performed in Ningxia, PR China, including 149 women with RSAs and 210 controls. The CAG and GGN repeats of the AR gene were genotyped using a PCR-based assay and were analyzed using Peak Scanner Software v1.0 to determine the CAG/GGN repeat length.

Results: CAG repeats ranged from 15 to 29 in the RSA patients, compared to 14 to 35 in the control group. The median value of CAG repeats was 22 for the RSA group and 24 for control group. The total AR CAG alleles (≤22 repeats), shorter AR CAG alleles (≤22 repeats), and biallelic means (≤22.5 repeats) were significantly different in the RSA group in comparison to the control group (P<0.001, P<0.01). The median value of the GGN repeats was 23 for the cases and 22 for controls. The total number of AR GGN alleles (≤23 repeats) was significantly different in the RSA group compared to the control group (P<0.5). There was no difference between the RSA group and the control groups in regards to shorter alleles, longer alleles, and biallelic means.

Interpretation & conclusions: Our observation suggests that the CAG and GGN repeat length is shorter in women with RSAs as compared with controls and that shorter CAG and GGN repeats may be pathogenic for RSAs in Chinese women. Further studies need to be done in different ethnic populations.

Key words Androgen - androgen receptor gene - CAG repeats - GGN repeats - recurrent spontaneous abortion
Recurrent spontaneous abortions (RSA) is defined as repeated occurrence of three or more miscarriages before 24th week of gestation. Epidemiological studies have suggested that RSAs comprise a multifactorial condition enhanced by a variety of genetic and environmental factors. Some known aetiologic factors of RSAs include parental chromosome abnormalities, uterine abnormalities, endocrinologic diseases, infections, thrombophilia, immunologic factors, nutritional, and environmental factors. However, the true cause of RSAs cannot be determined in approximately 50 per cent of cases.

Hormonal factors have been proposed to contribute to RSAs. Hormonal disorders may result from problems with certain endocrine glands, such as the pituitary, thyroid, adrenal, as well as the ovaries. An increased miscarriage rate has been observed in women with polycystic ovarian syndrome (PCOS), a disorder which is characterized by amenorrhoea, hirsutism, and an increase in androgen concentrations. The prevalence of PCOS in women with RSAs has been reported to be between 44 and 63 per cent. The presence of PCOS has not been shown to predict miscarriage, however, those women who miscarried have been shown to demonstrate higher levels of total testosterone, may be associated with RSAs.

Androgens are steroid hormones with several physiologic effects in both sexes, and their production is associated with androgen insensitivity syndromes of different degrees. These affect the metabolism of lipoprotein, the vascular endothelium and smooth muscle, and blood coagulation. At the molecular level, the effect of androgens is mediated by the activation of the androgen receptor (AR). The androgen receptor (AR) gene is a transcription factor mediating the physiologic effects of androgens. AR is essential for male sexual differentiation and maintenance of normal reproductive functions. In females, the AR is expressed in the ovary, mainly on the granulosa cells, suggesting some involvement in folliculogenesis. It is located on chromosome Xq11-12 and has a variable NH2-terminal domain that contains two functionally polymorphic microsatellites, a polyglutamine tract encoded by CAG repeats, and a polyglycine tract (GGN) encoded by (GGT)3 GGG (GGT)2 (GGC)n repeat.

Studies have shown that longer CAG repeats are associated with reduced AR transactivation activity and weaker transcriptional potential, while androgen responsiveness is increased in cell cultures with shorter GGN repeat. Increased androgen bioactivity may result from elevated circulating androgen concentrations. Sugawa and co-workers found that shorter CAG repeat lengths would increase serum levels of androgens.

In a study that has examined the AR exon 1 CAG repeat variation in RSAs, Aruna and associates have reported that longer CAG repeat lengths are associated with increased odds for RSAs in Indian women with statistical power estimated to be 90 per cent. To our knowledge, there is no study that has examined the GGN repeat variation in the setting of RSAs. For comprehensive assessment of the functional competence of the AR as a transcriptional regulator, both the CAG repeat and the GGN repeat need to be considered. Therefore, we undertook this study to investigate the possible role of the CAG and GGN repeat polymorphisms in the manifestation of RSAs in Chinese women.

Material & Methods

Venous blood samples (5 ml) were collected from 149 randomly selected women (27.97±4.49 yr, 19-42 yr) who had been diagnosed with RSAs during 2006-2011 for which no aetiology was known at the General Hospital of Ningxia Medical University, Ningxia, PR China. For the control group, blood samples from 210 women (27.67±3.67 yr, 20-36 yr), representing similar age, ethnicity, and socio-economic status as compared to the RSA group, were recruited from Yincuan Maternal Child Health Care Hospital in Ningxia. Women who had the spontaneous loss of three or more consecutive pregnancies before 20 weeks of gestation were included in the RSA group. The patients were investigated to exclude established causes of RSA, such as chromosomal abnormalities, uterine abnormalities, hereditary thrombophilias, endocrinologic disorders, immunologic factors and infections. Women in the control group were included if they had no history of abortions or fertility treatments and had normal menstrual cycles. This study was approved by the Institutional Ethical Committee and informed written consent was obtained from all women included in the study.

Genomic DNA was extracted by the TIANamp Blood DNA Kit (TianGen Biotech, Beijing, China) and quantified spectrophotometrically. The CAG and GGN repeats were genotyped using a PCR-based assay. Primer sequences were F:5'-FAM-

CHUAN et al: AR CAG & GGN REPEAT POLYMORPHISM & RSA 731
TCCAGAATCTGGTCCAGAGCGTC-3’ and R:5’ -GCTGTGAAGGTTGCTGGTTCCTCAT-3’ for the CAG
repeat, and primer sequences were F:5’-FAM-TCCTGGCACACTCTCTTCAC-3’ and R:5’-GCCAGGTTACCAACAGA-TCAGGT-3’ for the
gGN
repeat. PCR conditions for CAG amplification
were: initial denaturation of 94°C for 5 min, followed
by 30 cycles of 30 sec at 94°C, 30 sec at 65°C and 45 sec
at 72°C and finally 7 min at 72°C. PCR conditions for
GGN amplification were: initial denaturation at 94°C
for 5 min, followed by 30 cycles of 30 sec at 94°C,
30 sec at 60°C and 45 sec at 72°C and finally 7 min at
72°C. All PCR reactions were performed in a 12.5 µl
volume of PCR MIX 6.25 µl, Taq DNA polymerase
0.125 µl (5U/µl), 0.25 µl of each primer, DNA 0.5 µl,
double distilled water 5.125 µl.

For genotyping, 0.5 µl of PCR product was mixed
with 0.1 µl of LIZ500™ (Applied Biosystems, Foster
City, CA) and 7.4 µl of Hi-Di formamide (Applied
Biosystems, Foster City, CA) and analyzed on ABI
3730 DNA analyzer (Applied Biosystems, Foster City,
CA), using Peak Scanner software V1.0 to ascertain
the size of AR alleles. PCR and Gene Scan analysis
were repeated for all samples to confirm the number of
repeats.

As women have two AR alleles (for the two
X chromosomes), similar to the earlier method analyses were done in three different modes of allele
representation; (i) the mean value of the two alleles
(biallelic mean), (ii) the shorter allele alone (shorter),
and (iii) the longer allele alone (longer). The use of
the terms shorter and longer is relative for each individual
and does not represent an absolute CAG repeat number
or range of numbers.

Statistical analysis: An independent t-test (α=0.05)
was performed to test the significance in difference
in mean CAG repeats between RSA and control
populations. Chi-squared analysis was done to test for
the homogeneity of allele frequencies between the case
and control groups. Logistic regression analysis was
performed and odds ratios were obtained for perceived
risk alleles to determine if a significant association to
RSAs could be identified. Data were analyzed using
SPSS 11.5 statistical software (SPSS, Chicago, IL).

Results

CAG repeats ranged from 15 to 29 in the RSA
patients, compared to 14 to 35 in the control group
(Fig. 1). The average CAG repeat length was
significantly greater in the control group as compared
with RSA group (P=0.041, P=0.038; Table I). The
mean number of longer CAG repeats for the RSA
and control groups was similar (Table I). The
mean value of CAG repeats was 22 for the RSA group
and 24 for control group. For qualitative comparisons,
the median value of 22 repeats was considered, which
bisects the entire polymorphic range providing three
categories viz. <22, 22, and >22 repeats. The total AR
CAG alleles (≤22 repeats) and shorter AR CAG alleles
(≤22 repeats), were significantly (P<0.001, P<0.01)
different in the RSA group in comparison to the control
group. There was no significant difference between
the two groups for longer AR CAG alleles (Table II).
When the biallelic mean was categorized into >22.5
and ≤22.5 repeats, a significant difference was seen in
the biallelic mean distribution pattern (Fig. 1) between
the RSA and control groups (P<0.01, Table II).

The androgen receptor GGN repeats ranged from
18-28 and 16-28, respectively (Fig. 2). The median
value of the GGN repeats was 23 for the cases and 22
for controls. The mean number of shorter GGN repeat
length was significantly greater in control women as
compared with RSA women (Table III). For qualitative

<table>
<thead>
<tr>
<th>Category</th>
<th>Minimum CAG length</th>
<th>Maximum CAG length</th>
<th>Mean Biallelic CAG length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shorter</td>
<td>Cases</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td>14</td>
<td>32</td>
</tr>
<tr>
<td>Longer</td>
<td>Cases</td>
<td>15</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td>15</td>
<td>35</td>
</tr>
<tr>
<td>Biallelic mean</td>
<td>Cases</td>
<td>15</td>
<td>26.5</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td>15</td>
<td>33.5</td>
</tr>
</tbody>
</table>

*P<0.05 compared to controls
comparisons, the median value of 23 repeats was considered, which bisects the entire polymorphic range providing three categories viz. ≤23, 23 and >23 repeats. The total AR GGN alleles were significantly different in the RSA group in comparison to the control group (P<0.05, Table IV). No significant difference was found in the shorter and longer alleles distribution pattern (Fig. 2) between the RSA cases and controls.

Owing to the significant heterogeneity in the biallelic mean frequency between the RSA cases and the controls, the odds ratio was also computed to examine if there was any effect of repeat numbers in the upper polymorphic range of the allele spectrum in the manifestation of RSAs. Odds were computed between the RSA cases and the controls for the CAG repeats ≤22 vs. >22 both for the total and shorter alleles and ≤22.5 vs. >22.5 for the biallelic mean, and for the GGN repeats ≤23 vs. >23 for the total alleles (Table V). Highly significant odds were obtained for each of the categories suggesting strong association of extreme CAG repeat lengths with RSA. The result of logistic regression analysis illustrated that shorter CAG and shorter GGN repeat lengths might be associated with increased odds of RSA.

<table>
<thead>
<tr>
<th>Repeat size</th>
<th>Cases N (%)</th>
<th>Controls N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of CAG alleles</td>
<td>≤22</td>
<td>175 (58.7)**</td>
</tr>
<tr>
<td>Shorter</td>
<td>&gt;22</td>
<td>123 (41.2)**</td>
</tr>
<tr>
<td>Longer</td>
<td>≤23</td>
<td>76 (51)</td>
</tr>
<tr>
<td>&gt;23</td>
<td>73 (49)</td>
<td>115 (54.8)</td>
</tr>
<tr>
<td>Biallelic Mean</td>
<td>≤22.5</td>
<td>92 (61.7)**</td>
</tr>
<tr>
<td>&gt;22.5</td>
<td>57 (38.3)**</td>
<td>109 (51.9)</td>
</tr>
</tbody>
</table>

For total number of CAG alleles, Cases: 2N = 298 Controls: 2N = 420; For Shorter, Longer, Biallelic Mean, Cases: N = 149 Controls: N = 210 P<0.01, **<0.001, compared to controls

Discussion

The outcome of pregnancy relies on the success rate of various early events, such as implantation, establishment of foeto-maternal circulation and maintenance of increased blood flow to the implantation

![Fig. 1](image-url). Distribution of CAG repeat numbers in exon 1 of the AR gene in RSA cases and controls. (A) Total alleles, (B) short CAG allele, (C) long CAG alleles, (D) biallelic mean.
The development of RSA is complex and is regulated by multiple genetic pathways. Different genes encoding for proteins involved in various biologic pathways have been reported to be associated with recurrent spontaneous abortion. Several studies were performed to investigate whether polymorphisms of different candidate genes could be risk factors for RSA. However, the results are often controversial, leading to uncertain conclusions about the role of these polymorphisms in RSA pathogenesis.

There is evidence that androgens and their receptor are involved in uterine cell proliferation in rats. Several groups have investigated the relationship between AR gene polymorphisms and RSA. So, we hypothesized that polymorphisms in this gene might affect normal gene function and could be associated with RSA.
The effect of androgens is mediated through the activation of the androgen receptor. Ackerman et al\(^2\) have recently reported that there may be potential ethnic differences in androgenic pathway activity and androgen sensitivity. In this study, the relative AR (CAG)(n) repeat length in the different groups was: Afro-Caribbean (shortest repeat lengths and greatest predicted AR activity) < Caucasian < Hispanic < Thai (longest repeat length and lowest predicted AR activity). Lundin et al\(^2\) have shown that AR alleles with 23 GGN repeats have the strongest trans-activational activity and each change in the GGN tract length causes a slight impairment of receptor function. The length of the GGN repeat (number of glycine residues) was linearly and inversely associated with AR protein content in cell cultures, and longer GGN tracts resulted in a linearly reduced AR activity per cell\(^1\).

Our findings demonstrated that women in the RSA group had a significantly greater frequency of total AR CAG alleles (≤ 22 repeat), shorter AR CAG alleles (≤ 22 repeat), and biallelic means (≤ 22.5 repeat) than women in the control group. The shorter CAG repeat lengths were associated with an increased risk of RSAs. In contrast to our study, Aruna et al\(^8\) reported that women with RSAs had a significantly greater frequency of longer AR CAG alleles than the control women, and that these longer CAG repeat lengths were associated with an increased risk of RSAs. Our study provides evidence of the association of shorter CAG alleles with RSAs, which is in contrast to the findings by Aruna et al\(^8\). We found a significant difference in the CAG allele distribution in the RSA group compared with the control group. We also found a significant difference in the GGN total allele distribution in the RSA group as compared to the control group, thus it could perhaps be used as a genetic marker for the assessment of risk for RSA in women. However, there may be potential ethnic differences in the androgenic pathway activity and androgen sensitivity. So it is necessary to do more studies to discover the relationship between AR CAG / GGN repeat polymorphisms and RSA in other populations of the region.

In conclusion, there is a paucity of information in the literature regarding the relationship AR CAG and GGN polymorphisms in relation to RSAs. Thus, our study provides the necessary foundation so that other investigators may contribute to the development of relevant endocrine therapeutic strategies for the prevention of recurrent miscarriages.

**Acknowledgment**

The study was funded by the National Natural Science Foundation of China (Grant No. 31060147), PR China.

**References**


Reprint requests: Dr Huo Zhenghao, Department of Medical Genetics & Cell Biology, Ningxia Medical University
1160 Shengli Street, Yinchuan, Ningxia, PR China
E-mail: huozhh@163.com