Longer CAG repeat length in the androgen receptor gene is associated with premature ovarian failure

Suparna Chatterjee¹, Rajender Singh², Seema Kadam¹, Anurupa Maitra³, Kumasamy Thangaraj⁴, Pervin Meherji¹, and Deepak Modi⁵,⁶

¹Department of Infertility and Reproductive Endocrinology, National Institute for Research in Reproductive Health (ICMR), J.M. Street, Parel, Mumbai 400 012, India ²Division of Endocrinology, Central Drug Research Institute, Lucknow, India ³Molecular Endocrinology, National Institute for Research in Reproductive Health (ICMR), J.M. Street, Parel, Mumbai 400 012, India ⁴Center for Cellular and Molecular Biology, Habsiguda, Hyderabad, India ⁵Molecular and Cellular Biology Laboratory, National Institute for Research in Reproductive Health (ICMR), J.M. Street, Parel, Mumbai 400 012, India

¹Correspondence address. Tel: +91-22-24192034; Fax: +91-22-24139412; Email: modid@nirrh.res.in or deepaknmodi@yahoo.com

BACKGROUND: Premature ovarian failure (POF) is a disorder characterized by lack of ovulation and elevated serum gonadotrophin levels before the age of 40 years. The cause of POF in most cases is unknown. As mice lacking the Androgen receptor (Ar) gene reportedly have a POF-like phenotype, we hypothesize that, variations in the AR gene maybe one of the causative factors for POF in humans. Thus the objective of the study is to evaluate the number of CAG repeats in exon 1 of the AR gene in non-familial, non-syndromic cases of POF.

METHODS: A clinic-based case–control study. Seventy-eight patients with non-familial, non-syndromic POF, and 90 controls were recruited to investigate the CAG repeat numbers in exon 1 of the AR gene by PCR and Gene Scan analysis.

RESULTS: The mean CAG repeat length in exon 1 of the AR gene of women with POF was 23.6 ± 3.8, which was significantly higher than controls (20.08 ± 3.45) (P < 0.001). The biallelic mean CAG repeat ranged from 11 to 32 in the control women, compared to 16 to 30 in the POF patients. The 22 CAG repeat allele followed by the 24 CAG repeat allele was found to be at highest frequency (15.38 and 12.8%) in POF cases, although the 19 CAG repeat allele was observed at highest frequency (12.2%) in controls.

CONCLUSIONS: The observation suggests that the CAG repeat length is increased in women with POF as compared with controls, and may be pathogenic for POF, at least in a subset of Indian women.

Key words: androgen receptor gene / CAG repeats length / premature ovarian failure

Introduction

Premature ovarian failure (POF) is a disorder characterized by lack of ovulation and elevated levels of serum gonadotrophins before the age of 40 years, considerably before the physiological age of menopause (~48 years). The incidence of POF is estimated to be 1% in the general population and increases to approximately 10% in non-ovulating infertile women (Goswami and Conway, 2005).

The known etiologies of POF include infections, radiation, chemotherapy and autoimmunity (Goswami and Conway, 2005). However, a significant proportion of POF cases still remain idiopathic. The latter is often defined as non-syndromic POF and genetic factors are thought to be a cause of ovarian failure in this group of patients (Goswami and Conway, 2005; Simpson, 2008). On the basis of the analysis of POF cases that have a structural defect of the X chromosome, the region Xq13–26 has been implicated to be associated with POF (Therman et al., 1990; Rizzolio et al., 2006). However, an in-depth analysis of this locus indicated that this region may not have any candidate genes that can be implicated in POF and it has been postulated that alterations in the genes flanking this locus may be associated with POF (Rizzolio et al., 2006). This observation prompted us to explore Xq11–12 region that includes the Androgen Receptor (AR) gene.

AR, a steroid hormone receptor, is indispensable for male sexual differentiation and maintenance of normal reproductive functions (Hiort, 2000). In females, AR is expressed in the ovary, mainly the granulosa cells, suggesting involvement in folliculogenesis (reviewed in Kimura et al., 2007). Indeed, female mice null for the Ar gene, although fertile, display a POF-like phenotype (Shiina et al., 2006). This observation led us to speculate that variations in the AR gene may be associated with POF.
Large numbers of mutations and polymorphisms have been identified in the AR gene that are associated with various disorders ranging from androgen insensitivity to infertility (ftp://www.ebi.ac.uk/pub/databases/androgen) (Gottlieb et al., 1998). Exon 1 of the AR gene consists of two polymorphic repetitive DNA motifs, of which the CAG triplet repeat begins at codon 58 and varies between 8 and 35 repeats (Singh et al., 2007). Inverse correlation has been observed between the CAG repeat length and AR function in vitro (Choong et al., 1996; Buchanan et al., 2004). The results of in vitro experiments are also supported by clinical observations, where reduced CAG repeat lengths are observed in diseases associated with hyperandrogenicity (Singh et al., 2007; Lappalainen et al., 2008; Shah et al., 2008).

The association of AR gene defects and POF has been recently investigated. In a group of women with secondary amenorrhea, no correlation was found with CAG repeat length and POF. In another study, shorter CAG repeat length was found to be associated with POF (Bretherick et al., 2008; Sugawa et al., 2009). The observations that AR null mice display a POF-like phenotype (Shiina et al., 2006), increased CAG repeats are associated with hypoandrogenicity in men (Tut et al., 1997; Singh et al., 2007) and altered CAG repeats in the AR gene observed in a small group POF women (Sugawa et al., 2009) led us to hypothesize that the CAG repeat length in exon 1 of the AR gene may be altered in women with POF.

**Materials and Methods**

**Study subjects**

A total of 80 women with clinically confirmed POF visiting the Reproductive, Endocrine and Infertility Clinic of the National Institute for Research in Reproductive Health (NIRRH), Mumbai, were recruited into the study. POF was defined as cessation of menstruation for 6 months or more before attaining the age of 40 years, and serum FSH levels of more than 40 mIU/ml (estimated on two different occasions at least 2 months apart). The clinical characteristics of the patients included are presented in Table I.

Cases with family history of POF or fragile X syndrome were excluded; none of the recruited patients had phenotypic features or family history of blepharophimosis ptosis and epicanthus inversus syndrome. Patients with a recent history of mumps, measles, surgery, chemotherapy or radiation to pelvic region were excluded. Patients with autoimmune disorders or presence of antiovarian antibodies in the serum, chromosomal aberrations and/or sex chromosome mosaicism detected by standard cytogenetic analysis were also excluded from this cohort.

Thus the group included for the present study is defined as non-syndromic non-familial POF.

A total of 90 regularly cycling women served as controls (Table I). The control samples were collected from women visiting the same clinic for reasons other than infertility. All women had experienced normal puberty and menarche and had not undergone any kind of gynecological procedures, including Assisted Reproductive Technology. All women were phenotypically normal, with no family history of POF. Patients and controls enrolled belonged to a cosmopolitan population, residing in Mumbai, India. Blood was drawn from the control women during the follicular phase of the cycle (Day 3–Day 5).

All POF patients and controls enrolled were subjected to sequence analyses of coding regions, along with exon–intron boundaries of FOXL2, FMR1 and Inhibin (Inhibin alpha, Inhibin beta A and Inhibin beta B) genes. No mutations or premutations were detected in the Inhibin genes and FMR1 gene (data not shown). A FOXL2 mutation (Chatterjee et al., 2007) was detected in two of the POF patients, hence they were excluded from this study. Remaining 78 POF patients and 90 control subjects were evaluated for AR-CAG repeat length variations.

This study was approved by the Institutional Ethical Committee. Informed written consent was obtained from all the women included in the study.

**Table I** Characteristics of the patients and controls included in the present study

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Patients (N = 80)</th>
<th>Controls (N = 90)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>32.0 ± 3.4 (26–38)</td>
<td>32.2 ± 3.9 (22–40)</td>
</tr>
<tr>
<td>Age at menarche (years)</td>
<td>15.10 ± 2.29 (11–22)</td>
<td>14.67 ± 1.9 (11–17)</td>
</tr>
<tr>
<td>Normal cycle (%)</td>
<td>52.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Menstrual cycle length (days)</td>
<td>27.7 ± 2.5 (23–30)</td>
<td>27.1 ± 1.8 (23–30)</td>
</tr>
<tr>
<td>Irregular cycle (%)</td>
<td>47.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Amenorrhoeic since (age, years)</td>
<td>25.26 ± 24.26</td>
<td>NA</td>
</tr>
<tr>
<td>Normal ovaries (%)</td>
<td>62.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Streak ovaries (%)</td>
<td>37.5</td>
<td>0.0</td>
</tr>
<tr>
<td>FSH (mIU/ml)</td>
<td>115.42 ± 17b</td>
<td>6.16 ± 1.72b</td>
</tr>
<tr>
<td>LH (mIU / ml)</td>
<td>37.17 ± 3.12b</td>
<td>5.15 ± 1.00b</td>
</tr>
<tr>
<td>Proven fertile (numbers)</td>
<td>9</td>
<td>43</td>
</tr>
<tr>
<td>Spontaneous miscarriages (numbers)</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Failed to conceive spontaneously (numbers)</td>
<td>66</td>
<td>0</td>
</tr>
<tr>
<td>Number of individuals with CAG repeats &lt;20</td>
<td>21 (13.46%)</td>
<td>83 (46.1%)</td>
</tr>
<tr>
<td>Number of individuals with CAG repeats &gt;20</td>
<td>135 (86.5%)b</td>
<td>97 (53.8%)</td>
</tr>
</tbody>
</table>

*In POF patients, prior to onset of amenorrhea.

*Values with the same superscript letter are significantly different (P < 0.001, Student’s t-test).
FSH and serum Inhibin B estimations were done by radioimmunoassay and Enzyme Linked Immunosorbent assay, respectively, using commercial kits from Diagnostic Systems Laboratories (Webster, USA). The inter and intra-assay variations for both the assays was <10%.

Amplification of the CAG repeat
Genomic DNA was extracted by the conventional salting out or by phenol-chloroform method and quantified spectrophotometrically. DNA was stored in aliquots at −80°C until use. CAG repeat motif of exon 1 in the AR gene was amplified from the genomic DNA isolated from controls, patients and the reference population using primers flanking the CAG repeat region (Cram et al., 2000). The forward primer was labeled with 5’ FAM (carboxy fluorescein). PCRs were performed under the following conditions: initial denaturation at 94°C for 12 min, followed by 30 cycles of denaturation at 94°C, annealing at 60°C for 1 min and polymerization at 72°C deg for 1 min with a final extension at 72°C for 20 min.

For genotyping. 3.0 µl of PCR product was mixed with 0.3 µl of LIZ500™ and 6.5 µl of Hi-Di formamide and analyzed on 3730 DNA analyzer using Gene Mapper software to ascertain the size of AR alleles. PCR and Gene Scan analysis were repeated for all samples to confirm the number of repeats.

Allele distribution profiles
We previously analyzed CAG repeat for more than 3000 samples from different parts of India (Singh and Thangaraj, unpublished data). Therefore, the distribution of CAG repeat for our control samples were first compared with the CAG repository for the general population of this region. Total alleles were plotted, whereby each woman contributed two independent values that represented both CAG repeat alleles. The presence of two alleles in females and the phenomenon of X chromosome inactivation renders the analyses more complex. Therefore, we sought analyses in three different modes of allele representation; (i) the mean value of the two alleles (biallelic mean), (ii) the shorter allele alone, (iii) the longer allele alone; 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. In addition the cases and controls were segregated based on the median value of the CAG repeat length of the control group.

Statistical analysis
An independent t-test was performed to test the significance in difference in mean CAG repeats between POF and control populations, and a two-tailed χ² test was used to compare allele frequency between POF and control groups using GraphPad software (http://www.graphpad.com/). Statistical Power of the test was calculated using software from DSS research (http://www.dssresearch.com/toolkit/default.asp). The effect size was calculated using software from HyperStat online (http://davidmlane.com/hyperstat/).

Results
Clinical characteristics
The clinical and molecular characteristics of the cases and controls included in the study are detailed in Table I. Of the 80 POF patients studied, a FOXL2 mutation (Chatterjee et al., 2007) was found in two and hence they were excluded. Remaining 78 POF patients and 90 control subjects were evaluated for AR-CAG repeat length variations. No significant differences were observed in the mean age, age at pubarche and menstrual cycle length in the POF patients compared with controls. Upon pelvic ultrasound, ovaries were either undetectable or appeared as streak in approximately 38% of POF patients. Mean serum FSH and LH levels were significantly elevated in POF women as compared with controls.

CAG repeat length
The pattern of CAG repeat distribution in our control sample did not differ significantly from the CAG repeat distribution of general population (Table II). Figure 1A compares the mean CAG repeat length in control and POF women. The average CAG repeat length was significantly greater in POF women (23.6 ± 3.8) as compared with control women (20.08 ± 3.45). As females have two AR alleles (for the two X chromosomes), analysis was performed on the short and long alleles separately. The mean short and long allele lengths were significantly higher in patients as compared with the controls (Fig. 1B). The effect size of the test was found to be 3.03 and the statistical power of the test was calculated to be 100%.

The frequency distribution of the two CAG repeat lengths and the biallelic mean for the control and POF women is represented in Fig. 2. Figure 2A gives the total alleles, whereby each woman contributed two independent values that represented both CAG repeat alleles. The frequency distribution of CAG repeats in POF patients are shifted to the right, compared with controls. In the 90 control women, the biallelic mean CAG repeat (Fig. 2B) ranged from 11 to 32, whereas among the POF patients the CAG repeats ranged from 16 to 36. In contrast to controls, POF patients did not show any repeats in the range of 11–16. The short repeat length (Fig. 2C) ranged from 11 to 32 in control women, whereas it ranged from 16 to 33 in POF women. In the longer allele (Fig. 2D), repeat length ranged from 12 to 32 in the control women, whereas it was in the range of 17–36 in POF women. The 22 and 24 CAG repeat alleles were found to be at highest frequency (15.38% and 12.8%) in POF cases, although the 19 CAG repeat allele was observed at highest frequency (12.2%) in controls.

The median CAG repeat in exon 1 of AR gene of the control population was found to be 20. Nearly 87% of POF cases had CAG repeat length above the median of controls, only 13% of POF cases had a repeat equal to or less than controls (Table I).

Table II Comparison of CAG repeat length for case, controls and reference population

<table>
<thead>
<tr>
<th>Category</th>
<th>Minimum CAG length</th>
<th>Maximum CAG length</th>
<th>Mean Bi-allelic CAG length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>16</td>
<td>36</td>
<td>23.6 ± 3.8*</td>
</tr>
<tr>
<td>Controls</td>
<td>11</td>
<td>32</td>
<td>20.08 ± 3.45**</td>
</tr>
<tr>
<td>Reference population</td>
<td>11</td>
<td>31</td>
<td>19.5 ± 2.56</td>
</tr>
</tbody>
</table>

*p < 0.001 compared with controls and reference population.

**p = 0.19 compared with reference population.
repeat length ranged from 11 to 32 whereas in the POF group it ranged from 17 to 36. Intriguingly, more than 80% of POF women had CAG repeat length above the median repeat size of 20 observed for the control population. These evidences suggest that the CAG repeat length in women with POF is greater than controls.

Contrary to our findings, a study reported no significant difference in allele distribution of the CAG repeat between control and POF patients of Caucasian origin (Bretherick et al., 2008). Differences in the ethnicity of the population reported here versus that reported by Bretherick et al., maybe a cause of the observed differences. Furthermore, it is important to note that the POF group included in the Bretherick et al. study consisted of women with idiopathic secondary amenorrhea, whereas in the present study, we have rigorously scrutinized all patients and only non-familial, non-syndromic cases have been included.

In another recent study conducted in Japanese women, shorter CAG repeat were found to be associated with POF (Sugawa et al., 2009). These findings are surprising as shorter CAG repeats would lead to hyperandrogenicity and is reported to be reduced in conditions like PCOS (Shah et al., 2008; Lappalainen et al., 2008 and references therein). To the best of our knowledge POF is not associated with hyperandrogenicity, in fact, diminished AR functions are associated with ovarian failure (Hu et al., 2004; Shina et al., 2006; Walters et al., 2007, 2009). Since long CAG repeats in exon 1 is associated with hypoandrogenicity, it is likely that the increased CAG repeats in women with POF might reduce the androgen receptor activity. Furthermore, as Ar knockout mice have POF like phenotype (Shina et al., 2006), it is tempting to speculate that increased CAG repeat lengths may contribute to POF in these women.

The AR gene is X-linked and is known to undergo X-inactivation. Although this process is essentially random in normal women, the pattern of X-inactivation is found to be skewed in some women with POF (Sato et al., 2004; Bretherick et al., 2007). Although in this study, we have not investigated the pattern of X-inactivation, it is plausible that effects of repeat length at the functional level may be further amplified in the event the shorter allele is preferentially inactivated. In this context, it is of interest to note that mice carrying functionally deficient Ar alleles (Walters et al., 2007) do not display a POF-like phenotype as observed in complete Ar nulls (Shina et al., 2006), but does compromise ovarian functions and demonstrate a hormonal phenotype resembling hypogonadism as observed in POF women (Walters et al., 2007, 2009). It would, therefore, be of interest to correlate how the AR polymorphisms associate with the pattern of X-inactivation and affect at the ovarian functions.

In the present study, we observed that the CAG repeat length is higher in women with POF as compared with control and this increased CAG repeat may contribute to hypoandrogenicity. However, how reduced activity of AR gene leads to POF is a matter of speculation. Although the reduced levels of androgens have been observed in serum of POF women (Bachelot et al., 2005; Benetti-Pinto et al., 2005), increased theca cell steroidogenic activity also have been reported in the ovaries from these patients (Bachelot et al., 2005). AR transcripts and protein have been localized to the granulosa cells of the human ovary (reviewed in Kimura et al., 2007) suggesting that the granulosa cells are sites of direct androgen action, and these cell types are responsive to AR activity. Interestingly, several genes known to be involved in the oocyte–granulosa cell

**Discussion**

The results of the present study demonstrate that the CAG repeat length in exon 1 of the AR gene in women with non-syndromic non-familial POF is significantly longer than that observed in normal women. To the best of our knowledge this is the first report investigating and implicating the involvement of AR gene variations with idiopathic non-familial, non-syndromic POF.

Analysis of 78 cases with POF revealed that the mean CAG repeat length in exon 1 of AR gene was found to be significantly increased than that observed in the normal women. Interestingly, when the two X chromosome alleles were segregated based on the repeat length, the mean CAG repeat length of both the longer and shorter allele individually was found to be significantly higher in women with POF as compared with controls. This observation further prompted us to compare the frequency distribution of the repeats in both groups. As evident, in the control group the mean biallelic CAG

![Figure 1](image-url) **Figure 1** CAG repeat length in AR gene of control and POF women. (A) Biallelic mean CAG repeat numbers, *significantly different as compared with controls (P < 0.001). (B) Mean CAG repeat numbers segregated based on short and long alleles, *significantly different (P < 0.001) as compared with short control, **significantly different (P < 0.001) as compared with long control.
regulatory loop such as KIT ligand (Kitl), bone morphogenetic protein 15 (Bmp15) and growth differentiation factor 9 (Gdf-9) have been found to be down-regulated in the ovaries of Ar null mice (Shiina et al., 2006). Interestingly, BMP 15 and GDF 9 are preferentially expressed in oocytes and their role in ovarian function and POF has been well documented (Carabatsos et al., 1998; Joyce et al., 2000; Yan et al., 2001; Juengel et al., 2002; Otsuka and Shimasaki, 2002; Mazerbourg and Hsueh, 2003; McNatty et al., 2005). It will be of potential interest to identify the downstream factors responsive to AR in the human ovary to understand the role of androgens and AR in pathophysiology of POF.

Acknowledgements

Sincere thanks to Mr A.G. Reddy, Ms Priyamvada Singh and Mr Arun Khatri (CCMB) for technical help. We express our thanks to Director, CCMB, Hyderabad for the support and encouragement.

Funding

S.C. is thankful to the Lady Tata Memorial Trust for providing Junior and Senior Research Fellowship and to Indian Council of Medical Research for Senior Research Fellowship. A portion of this work included in this publication (NIIRRH/MS/33/2008) has been supported financially by grants from the Indian Council of Medical Research (ICMR), New Delhi, India.

References


Submitted on January 14, 2009; resubmitted on July 16, 2009; accepted on July 21, 2009.