WILEY-BLACKWELL

Androgen insensitivity syndrome
DOI: 10.1111/j.1745-7262.2007.00337.x

AJA

·*Original Article*·

Androgen insensitivity syndrome: do trinucleotide repeats in androgen receptor gene have any role?

Singh Rajender[1,3], Nalini J. Gupta[2], Baidyanath Chakravarty[2], Lalji Singh[1], Kumarasamy Thangaraj[1]

[1]*Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India*
[2]*Institute of Reproductive Medicine, Salt Lake, Kolkata PN 700064, India*
[3]*Division of Endocrinology, Central Drug Research Institute, Lucknow, India*

Abstract

Aim: To investigate the role of CAG and GGN repeats as genetic background affecting androgen insensitivity syndrome (AIS) phenotype. **Methods:** We analyzed lengths of androgen receptor (*AR*)-CAG and GGN repeats in 69 AIS cases, along with 136 unrelated normal male individuals. The lengths of repeats were analyzed using polymerase chain reaction (PCR) amplification followed by allelic genotyping to determine allele length. **Results:** Our study revealed significantly shorter mean lengths of CAG repeats in patients (mean 18.25 repeats, range 14–26 repeats) in comparison to the controls (mean 22.57 repeats, range 12–39 repeats) (two-tailed $P < 0.0001$). GGN repeats, however, did not differ significantly between patients (mean 21.48 repeats) and controls (mean 21.21 repeats) (two-tailed $P = 0.474$). Among patients' groups, the mean number of CAG repeats in partial androgen insensitivity cases (mean 15.83 repeats) was significantly less than in complete androgen insensitivity cases (mean 19.46 repeats) (two-tailed $P < 0.0001$). **Conclusion:** The findings suggest that shorter lengths of repeats in the AR gene might act as low penetrance genetic background in varying manifestation of androgen insensitivity. *(Asian J Androl 2008 Jul; 10: 616–624)*

Keywords: androgen receptor; CAG repeat; GGN repeat; androgen insensitivity

1 Introduction

Androgens (testosterone and dihydrotestosterone) mediate their functions through androgen receptors (AR). The end organ resistance to androgens manifest in a spectrum of androgen insensitivity syndromes (AIS), with mild form AIS (MAIS) represented by male infertility or undermasculinization, partial form AIS (PAIS) represented by ambiguous genitalia and complete form AIS (CAIS) represented by female phenotype in genetically male individuals [1]. Mutations in the AR gene are a frequent cause of AIS; however, the origin of certain other cases remains elusive [2, 3]. The genotype–phenotype correlations in AIS had been very enigmatic. Same nucleotide substitutions in the *AR* gene have been reported in different grades of AIS [4] (see www.androgendb.mcgill.ca/). G2445A substitution is known to cause both PAIS [5] and CAIS [6]. C2296A substitution, with proven

Correspondence to: Dr Kumarasamy Thangaraj, Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India.
Tel: +91-40-2719-2636 Fax: +91-40-2716-0591
E-mail: thangs@ccmb.res.in
Received 2006-10-10 Accepted 2007-07-23

pathogenicity in PAIS [7] and CAIS [8] patients, was also observed in an absolutely normal individual [9]. Another interesting example is G995A mutation, which is known to cause MAIS [10], PAIS [11] and CAIS [11], and has also been reported to exist in 8% of the normal population [12].

Some studies propose the role of 'genetic background' or 'modifiers', in varying manifestation of AIS [13–15]. The mutations/polymorphisms in the genes involved in androgen action might affect overall response to androgens. The somatic mutations in androgen responsive organs have been reported to account for phenotypic variations in some cases [14, 15]. The exposure of the fetus to androgens at embryonic stage has also been suggested to count for the variations [16]. Holterhus *et al.* [16] report a family with four affected individuals sharing the same mutation in the *AR* gene but having different phenotypes. On the basis of the response to testosterone, it was concluded that variation in the phenotype was the result of variation in ligand concentration in early fetal life. In a similar study, Boehmer *et al.* [17] reported that variation in the phenotype in affected individuals might vary depending on 5-α reductase 2 activities. However, the above factors have been able to explain only few cases out of more than 20 known cases of phenotypic variations.

The AR gene has two polymorphic (CAG and GGN) repeats in exon 1, encoding variable lengths of polyglutamine and polyglycine stretches, respectively, in the N-terminal region of the protein. CAG, a simple repeat, varies in length from 8 to 35 repeats, whereas GGN, a complex repeat, represented by $(GGT)_3GGG(GGT)_2(GGC)_n$, varies in length from 10 to 30 repeats. Variations in lengths of CAG repeats have been associated with diverse clinical conditions [18]. These studies along with *in vitro* assays have shown an inverse correlation between CAG repeat length and AR transactivation function [19, 20]. However, GGN repeat length variation has been relatively less studied, therefore, more studies are required to reach conclusions about the association of repeat length variation with AR function. Taking into consideration the functional proof of inverse correlation of lengths of CAG repeats with AR function, we undertook the present study on *AR*-CAG and GGN repeats to determine whether variations in the length of these repeats was associated with phenotypic variations in AIS.

2 Materials and methods

2.1 Subjects and clinical evaluation

A total of 69 AIS cases were recruited through the Institute of Reproductive Medicine (IRM), Kolkata, India. The patients belonged to Indo-European and Austro-Asiatic linguistic affiliations and were inhabitants of four states (West Bengal, Orissa, Bihar and Jharkhand) of India. All the patients were subjected to physical and clinical evaluations, and family histories of all the patients were recorded. Upon complete clinical examination, phenotypes were diagnosed as PAIS in 23 patients and as CAIS in the remaining 46 (Table 1).

The assignment of androgen insensitivity was based upon the presence of 46, XY karyotpye, abdominal gonads and testicular tissue in gonads (histology done wherever possible). Patients were further classified as PAIS and CAIS categories on the basis of individual phenotype. Serum levels of testosterone, leutinizing hormone (LH) and follicle stimulating hormone (FSH) were measured by radioimmunoassay. Absolute values of testosterone and LH were multiplied to obtain androgen sensitivity index (ASI) values. Out of a total of 69 patients, 52 underwent surgical extraction of gonads, biopsy from which was used for histopathological analyses. At least 57 patients had a first-degree or second-degree relative with AIS phenotype. Although we could not collect samples of the siblings for all cases, we had samples of two siblings from three families and three siblings from one family. However, phenotype was the same in two or more affected siblings from these two families. A total

Table 1. Hormone profile of the patients. The percentage of the patients in each category is followed by absolute numbers in bracket. The percentages were calculated separately for complete form androgen insensitivity syndrome (CAIS) and partial form AIS (PAIS) categories. LH, leutinizing hormone; FSH, follicle stimulating hormone; ASI, androgen sensitivity index.

Category of	Testosterone		LH		FSH		ASI	
patients	Elevated	Normal	Elevated	Normal	Elevated	Normal	Elevated	Normal
CAIS	47.83 (22)	52.17 (24)	43.48 (20)	56.52 (26)	8.69 (4)	91.3 (42)	47.83 (22)	52.17 (24)
PAIS	39.13 (9)	60.87 (14)	43.48 (10)	56.52 (13)	13.04 (3)	86.96 (20)	47.83 (11)	52.17 (12)

of 136 healthy, unrelated and ethnically matched fertile men with no symptoms of undervirilization were recruited in the study as controls. Blood samples from the patients and controls were collected with their fully informed written consent.

In addition to the above patients and controls, we also analyzed repeat length data available on AR mutation database (www.androgendb.mcgill.ca/). Information regarding CAG and GGN repeat length is available in the database for a significant number of cases. These cases represent patients from different populations of the world, which therefore serves as a stringent dataset to validate our results. We selected this as a third dataset to cross check our findings of differences between PAIS and CAIS cases. For this purpose, all the mutations resulting in amino acid substitutions were taken into account but not the mutations resulting in reduction of *AR* expression or truncation of AR protein.

2.2 CAG and GGN repeat length analysis

DNA was isolated from peripheral blood lymphocytes using protocol described in our earlier study [21]. CAG repeat region of the *AR* gene was amplified with a pair of primers, forward: 5'-FAM-CAGAATCTGTTCCAGAGCGTGC-3', reverse: 5'-AAGGTTGCTGTTCCTCATCCAG-3' flanking the repeat. Polymerase chain reaction (PCR) mixture consisted of 1.0 μL PCR buffer (10 ×), 1.0 μL $MgCl_2$ (25 mmol/L), 1.0 μL dNTPs (10 mmol/L), 1.0 pmol/L of each primer, 0.5 units Amplitaq Gold DNA polymerase and 20 ng genomic DNA. PCR was performed under the following conditions: initial denaturation at 94ºC for 12 min followed by 30 cycles of 94ºC for 1 min, 55ºC for 1 min and 72ºC for 1 min, with a final extension at 72ºC for 30 min. GGN repeat was amplified with a pair of primers: forward 5'-FAM-CCGCTTCCTCATCCTGGCACAC-3' and reverse 5'-GCCGCCAGGGTACCACACATC-3' flanking the repeat region. PCR reaction mixture included 1.0 μL PCR buffer (10 ×), 1.0 μL $MgCl_2$ (25 mmol/L), 1.0 μL dNTPs (10 mmol/L), 1.0 μL DMSO (100%), 1.0 μL glycerol (100%), 1.0 pmol/L of each primer, 0.5 units Amplitaq Gold DNA polymerase and 20 ng genomic DNA. PCR conditions consisted of denaturation at 96ºC for 15 minutes, followed by 40 cycles at 96ºC for 1.5 min, 65ºC for 1 min and 72ºC for 3 min and a final extension at 72ºC for 20 min. For GeneScan, 3.0 μL of PCR product was mixed with 0.2 μL of LIZ500 and 6.8 μL Hi-Di formamide. Upon denaturation for 5 min at 95ºC and cooling for 5 min on ice, samples were run on 3730 DNA analyzer (Applied Biosystems, Foster City, CA, USA). PCR and genotyping were repeated for all samples to confirm the number of the repeats. The raw data were further analyzed using GeneMapper software (Applied Biosystems).

2.3 Statistical analysis

Mean, median and mode were calculated as descriptive statistics using SPSS software (version 11, SPSS, Chicago, IL, USA). Initially, the difference in mean repeat length for cases and controls was estimated to obtain overall difference between the two. Later the patients were categorized in two groups (CAIS and PAIS) and repeat data was compared between these groups. The significance of the differences in mean repeat length was tested by independent samples *t*-test using SPSS software. Only two-tailed *P*-values were used for acceptance or rejection of null hypotheses and $P < 0.05$ were considered significant. For joint analysis, the numbers of CAG and GGN repeats were cross-classified in three groups of less than average, average and more than average repeat length to observe the differences between cases and controls.

To understand the relationship between various clinical and genetic factors, we analyzed the data for Pearson correlation. To explore the possibility of various clinical and genetic parameters to be used as diagnostic factors for discrimination between CAIS and PAIS phenotype, we created receiver's operating characteristic (ROC) curves. For ROC curves, the phenotype (CAIS or PAIS) was taken as the state variable while CAG and GGN repeats length; testosterone, LH and FSH level were taken as predicting parameters. To understand the relationship between the phenotype and clinical parameters, we further analyzed the data by regression. For regression analyses, phenotype (CAIS or PAIS) was taken as the dependent parameter and various clinical and genetic factors as independent parameters.

3 Results

3.1 Clinical evaluation

Measurements of serum hormone levels showed that testosterone, LH and FSH levels were in the upper normal male range in the majority of the cases, and LH and FSH values were very high in some cases (Table 1). Testosterone and LH were elevated in almost 50% of CAIS

cases and 40% of PAIS cases. In total, testosterone and LH levels were elevated in approximately 90% of the cases. In contrast, FSH levels were elevated in approximately 10% of the cases. Among all the cases, ASI was elevated in approximately 50%; however, it was in the upper normal male range in more than 90% of cases. Histopathology of testicular biopsies indicated testicular cancers or hyperplasia of interstitial or Sertoli cells in 52 (75.36%) cases [22].

3.2 Genetic analyses

We observed narrowed CAG repeat length range in patients (14–26 repeats) as compared to controls (12–39 repeats) (Figure 1). Additionally, the mean number of the repeats was statistically significantly less in patients (18.25 repeats) in comparison to the controls (22.57 repeats) at the 95% confidence level ($P < 0.0001$). The modal value of CAG repeats was lower for the patients (15 and 19 repeats) than for the controls (23 repeats). Among patients groups, PAIS cases had significantly smaller mean CAG repeat length (15.83 repeats) in comparison to CAIS cases (19.46 repeats) at the 95% level of confidence ($P < 0.0001$) (Table 2). All PAIS cases had ≤ 23 CAG repeats (assuming 22 and 23 repeats as maximum frequent or average number of repeats), whereas CAIS cases had a normal distribution of CAG repeats. In familial cases, all affected siblings had a similar phenotype, however the CAG repeat length did not differ significantly among these individuals.

In contrast, GGN repeat was less polymorphic and mean repeats length did not differ significantly between patients (mean length 21.48 repeats) and control groups (mean length 21.21 repeats) (Table 3). Alleles with 21 and 22 repeats were most frequent with a very low frequency of other alleles on either side of the average (Figure 2). *AR* alleles with 17 GGN repeats were ob-

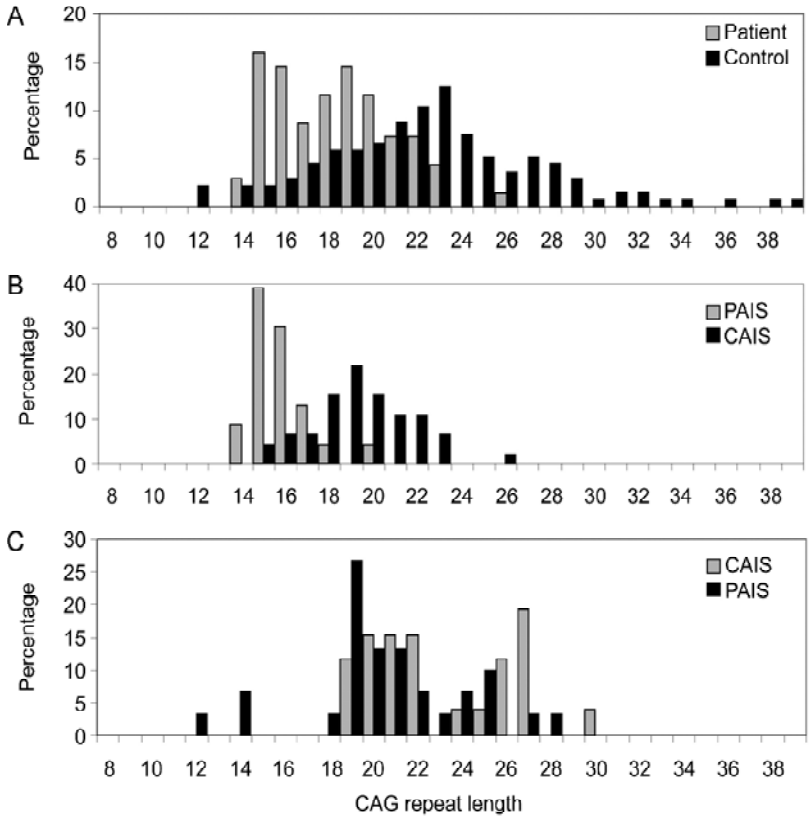

Figure 1. The distribution of CAG repeats in androgen receptor (*AR*) gene. (A): Androgen insensitivity syndrome (AIS) patients versus controls of the present study. (B): Complete form AIS (CAIS) versus partial form AIS (PAIS) cases of the present study. (C): CAIS versus PAIS cases from AR mutation database (http://www.androgendb.mcgill.ca/). The percentage was calculated separately for PAIS, CAIS and control samples.

served only in the patients. Mean number of GGN repeats length did not differ significantly between PAIS (21.43 repeats) and CAIS (21.50 repeats) categories. Modal value of this repeat was the same (22 repeats) in PAIS, CAIS and control individuals (Table 3). Joint distribution analysis of CAG/GGN repeats length showed significant differences in the distribution of these repeats in various combinations between cases and controls (Table 4).

Comparison of repeat length variation for the data from AR mutation database revealed that 77% of PAIS cases had ≤ 23 repeats, whereas only 23% had more than 23 repeats and only 2 (6.67%) out of a total 30 cases had more than 25 repeats (Figure 1). Among both CAIS and MAIS cases, almost an equal number of individuals had repeats below and above average. The mean number of repeats in PAIS cases (20.63 repeats) was statistically significantly less than for CAIS cases (23.12 repeats) at the 95% level of confidence ($P = 0.01$) (Table 2). CAG repeat in CAIS did not show the expected normal distribution, probably because this data was constituted by samples from different ethnic populations. Unfortunately, CAG repeat information was not available for all the cases annotated in the AR mutation database, which would have given a much clearer picture about the relationship between number of repeats and AIS phenotype. This shows the importance of analyzing CAG repeats length in studies involving the AR gene.

Significant correlations were observed between phe-

Table 2. CAG repeat length distribution in patients and controls. For a cross comparison, the data from androgen receptor (AR) mutation database (http://www.androgendb.mcgill.ca/) is also included. AIS, androgen insensitivity syndrome; CAIS, complete form androgen insensitivity syndrome; PAIS, partial form androgen insensitivity syndrome. SD, standard deviation.

Category	Total samples	Range	No. of alleles	Mean ± SD	Mode
Present study					
All AIS patients	69	14–26	11	18.25 ± 2.65	15, 19
PAIS	23	14–20	06	15.83 ± 1.34	15
CAIS	46	15–26	10	19.46 ± 2.30	19
Controls	136	12–39	25	22.57 ± 4.99	23
AR-database					
PAIS	30	12–28	12	20.63 ± 3.61	19
CAIS	26	19–30	09	23.12 ± 3.25	27

Table 3. GGN repeat length distribution in patients and controls. AIS, androgen insensitivity syndrome; CAIS, complete form androgen insensitivity syndrome; PAIS, partial form androgen insensitivity syndrome.

Category	Total samples	Range	No. of alleles	Mean ± SD	Mode
All AIS patients	69	15–26	8	21.42 ± 1.92	22
PAIS	23	17–26	5	21.52 ± 1.95	22
CAIS	46	15–26	7	21.37 ± 1.92	22
Controls	136	15–26	7	21.21 ± 1.97	22

Table 4. Joint distribution of CAG and GGN repeat in cases and controls. The percentage of cases or controls in each category is followed by absolute numbers in bracket.

CAG repeat length	GGN repeat length					
	Cases			Controls		
	Less than 21	Average (21, 22)	More than 22	Less than 21	Average (21, 22)	More than 22
Less than 22	11.59 (08)	68.11 (47)	7.25 (05)	8.82 (12)	28.67 (39)	3.67 (05)
Average (22, 23)	1.45 (01)	10.14 (07)	0 (0)	3.67 (05)	18.38 (25)	0.74 (01)
More than 23	0 (0)	1.45 (01)	0 (0)	4.41 (06)	30.15 (41)	1.47 (02)

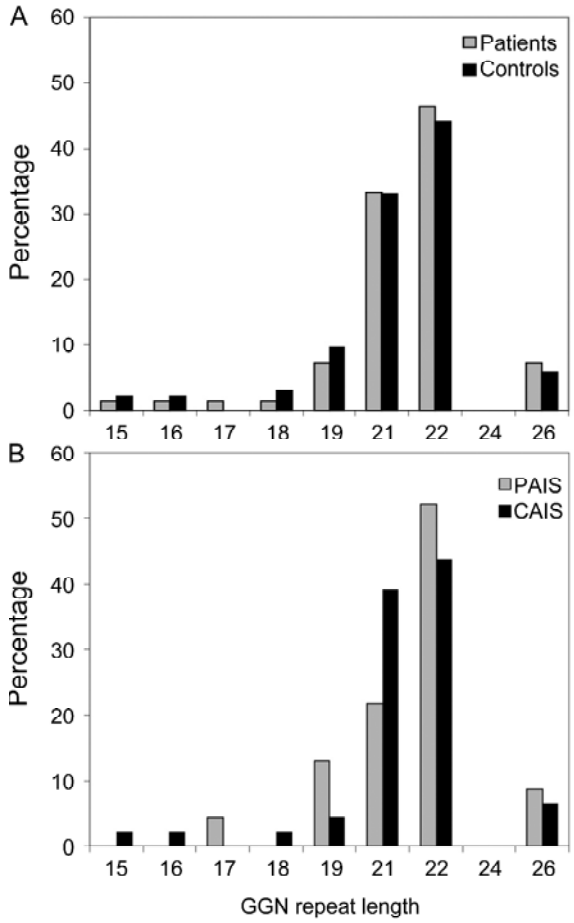

Figure 2. The distribution of GGN repeats in androgen receptor (*AR*) gene. (A): Androgen insensitivity syndrome (AIS) patients versus controls of the present study. (B): Complete form AIS (CAIS) versus partial form AIS (PAIS) cases of the present study.

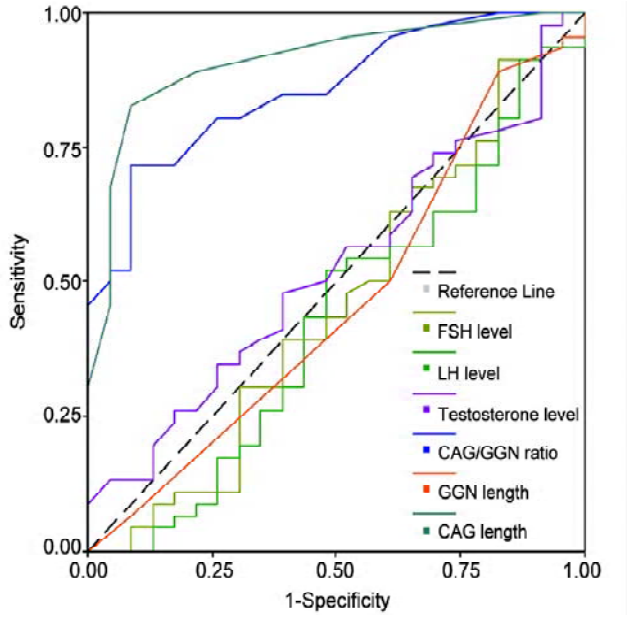

Figure 3. Receiver's operating characteristic curves (ROC) for predicting the best diagnostic factor. The phenotype (CAIS or PAIS) was taken as state variable and various genetic and clinical parameters as test or predicting variables. LH, leutinizing hormone; FSH, follicle stimulating hormone.

notypes (CAIS or PAIS), CAG repeat length (r = 0.661, P = 0.000) and FSH and LH levels (r = 0.27, P = 0.025) (Table 5). However, CAG and GGN repeat lengths did not correlate significantly each other. ROC curve analysis indicates that CAG repeat length is the best diagnostic factor to differentiate between CAIS and PAIS (Figure 3) (Table 6). Length variation of CAG repeats could account for 91.2% variability in AIS phenotype (area under the curve = 0.912, P = 0.000). Although ratio of CAG and GGN repeat length could also be used for diagnosis (area under the curve = 0.857, P = 0.000), it could be attributed to CAG repeat only because the GGN repeat did not correlate significantly with the AIS phenotype (Table 6). For diagnostic purposes aimed at discrimination between CAIS and PAIS, the best cut-off value for CAG repeat was 17.5 repeats, whereas the best cut-off for CAG/GGN ratio was 0.85. The remaining parameters were not significant for diagnostic purposes. Regression analysis revealed that the best predicting factor for AIS phenotype was CAG repeat length (R^2 = 0.455). The variation in AIS phenotype could be largely attributed to CAG repeat length polymorphisms (standardized regression coefficient, β = 0.656) (Table 7), whereas other factors did not account significantly for variability in AIS phenotype.

4 Discussion

In the present study, we have analyzed CAG and GGN repeat length polymorphisms in the AR gene, to observe the role of these repeats as genetic modifiers of androgen insensitivity. Considering the differences observed between PAIS and CAIS cases, we hypothesize that one major component of genetic background in manifestation of AIS is constituted by CAG repeat of the AR gene. Molecular defects that disrupt androgen binding completely are likely to result in CAIS irrespective of

number of CAG repeats in the *AR* gene. However, eventual level of androgen insensitivity in the cases with molecular defects resulting in partial loss of androgen binding or transactivation potential might be affected by the number of CAG repeats in the *AR* gene. The presence of shorter/less than average number of CAG repeats might compensate partially for loss of androgen sensitivity, whereas the presence of more than average number of

Table 5. Pearson correlation between various clinical and genetic parameters of the patients. FSH, follicle stimulating hormone; LH, leutinizing hormone. **Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).

		Diagnosis	CAG length	GGN length	Testosterone level	LH level	FSH level
Diagnosis	Pearson correlation	1.000	0.661**	–0.038	0.085	–0.187	–0.150
	Significance (2-tailed)	—	0.000	0.759	0.485	0.124	0.217
CAG length	Pearson correlation	0.661**	1.000	–0.063	0.221	–0.170	–0.054
	Significance (2-tailed)	0.000	—	0.605	0.068	0.162	0.657
GGN length	Pearson correlation	–0.038	–0.063	1.000	–0.071	–0.010	–0.203
	Significance (2-tailed)	0.759	0.605	—	0.564	0.933	0.094
Testosterone	Pearson correlation	0.085	0.221	–0.071	1.000	–0.102	0.172
level	Significance (2-tailed)	0.485	0.068	0.564	—	0.406	0.158
LH level	Pearson correlation	–0.187	–0.170	–0.010	–0.102	1.000	0.270*
	Significance (2-tailed)	0.124	0.162	0.933	0.406	—	0.025
FSH level	Pearson correlation	–0.150	–0.054	–0.203	0.172	0.270*	1.000
	Significance (2-tailed)	0.217	0.657	0.094	0.158	0.025	—

Table 6. Area under ROC curve for different clinical and genetic parameters of the patients. For ROC curves, phenotype (CAIS or PAIS) was taken as state variable and various clinical and genetic parameters as predicting parameters. [a]Under nonparametric assumption; [b]Null hypothesis: true area = 0.5. FSH, follicle stimulating hormone; LH, leutinizing hormone. ROC, receiver's operating characteristic.

Test Result Variable(s)	Area	Standard error[a]	Asymptotic significance[b]	Asymptotic 95% confidence interval	
				Lower bound	Upper bound
CAG length	0.912	0.037	0.000	0.840	0.984
GGN length	0.462	0.076	0.611	0.314	0.611
CAG/GGN ratio	0.857	0.044	0.000	0.771	0.944
Testosterone level	0.525	0.072	0.741	0.383	0.666
LH level	0.429	0.076	0.340	0.280	0.578
FSH level	0.454	0.077	0.537	0.304	0.605

Table 7. Regression coefficients for various clinical parameters. For regression analysis, the diagnosis (CAIS or PAIS) was taken as dependent variable and clinical parameters as independent variables. CAIS, complete form androgen insensitivity syndrome; FSH, follicle stimulating hormone; LH, leutinizing hormone; PAIS, partial form androgen insensitivity syndrome; B, values for the regression equation for predicting the dependent variable from the independent variable.

Model	Unstandardized coefficients		Standardized coefficients (Beta)	*t*	Sig.	95% confidence interval for B		Correlations		
	B	SE				Lower bound	Upper bound	Zero-order	Partial	Part
Constant	–0.242	0.648		–0.374	0.710	–1.537	1.052			
CAG length	0.121	0.018	0.656	6.774	0.000	0.085	0.157	0.661	0.649	0.630
GGN length	–4.846E-03	0.024	–0.020	–0.206	0.838	–0.052	0.042	–0.038	–0.026	–0.019
Testosterone level	–8.664E-05	0.000	–0.050	–0.510	0.612	0.000	0.000	0.085	–0.064	–0.047
LH level	–2.437E-03	0.004	–0.055	–0.553	0.582	–0.011	0.006	–0.187	–0.070	–0.051
FSH level	–5.014E-03	0.005	–0.095	–0.944	0.349	–0.016	0.006	–0.150	–0.118	–0.088

repeats might enunciate the extent of androgen insensitivity. Therefore, molecular defects resulting in partial loss of sensitivity to androgens might manifest as PAIS when presented in combination with equal to or less than average number of CAG repeats, but as CAIS when presented with more than average number of this repeat. It is possible that a longer CAG repeat might give rise to MAIS with or without other contributing factors. Molecular defects having mild effects on androgen sensitivity, when presented in combination with very few CAG repeats (much less than average) in the AR gene might not be pathogenic at all. The finding of relatively shorter repeats in PAIS in comparison to CAIS cases from the AR mutation database validated out hypothesis.

More than 70 different proteins have been identified as interacting with AR for its downstream action (see www.androgendb.mcgill.ca/). The polyglutamine tract is located in the region of the AR protein known to interact with some AR coregulators. Transfection assays have demonstrated that interaction of AR with coactivator ARA24 decreases with increasing *AR*-CAG repeat length, resulting in decreased AR transactivation potential. Similarly, longer lengths of CAG repeats result in decreased ability of AR to be activated by members of the steroid receptor coactivators (SRC) family of coregulators (SRC-1, SRC-3 and transcriptional intermediary factor (TIF-2) [23]. The above factors indicate that length of CAG repeats is crucial for AR action and might affect androgen action. AR is also known to interact with many tumor suppressor genes [23]; however, the influence of CAG repeat on the interaction of AR with these genes is yet to be deciphered. Therefore, in addition to polymorphic variations in *AR* gene, polymorphisms in AR interacting genes or in promoter regions of AR target genes might influence overall transactivation potential of AR gene and, hence, sensitivity to androgens. In contrast, GGN repeat does not seem to affect the level of androgen insensitivity. Although joint analyses of the two repeats (haplotypes) showed certain differences between cases and controls, they were not statistically significant and could be attributed mainly to CAG repeat distribution and the smaller sample size of the cases in comparison to controls. ROC curves and regression analyses also showed that the majority of variations in AIS phenotype could be attributed to CAG repeat length with minor contribution from other factors.

The levels of testosterone, LH and their multiplication product (ASI) were elevated in approximately 50% of the cases; however, more than 90% of the cases had ASI levels in the upper male range. This indicates that ASI values in the upper normal male range or higher than that might indicate AIS in an individual. The analysis of CAG repeats length among siblings sharing the molecular defect but displaying different phenotypes might further help in understanding the role of this triplet repeat in varying manifestation of androgen insensitivity. The residual AR function in AR knockout mice might depend upon the type of mutation and also the number of CAG repeats in the background of the mutation. Therefore, knockout studies might further assist in understanding the effect of coding triplet repeats on AR function.

To conclude, CAG repeat probably functions as a low penetrance allele for androgen sensitivity, whereas GGN repeat length does not seem to affect the same. In molecular defects resulting in a partial loss of androgen sensitivity, the presence of a less than average number of CAG repeats in background might partially compensate for the phenotypic effect, while CAG repeat lengths above average might enunciate the extent of androgen insensitivity. However, the number of CAG repeats does not seem to have any effect in combination with molecular defects resulting in complete loss of androgen sensitivity. The penetrance of CAG repeats might further vary between populations depending upon the combinations from other polymorphisms in *AR* and its interacting genes, somatic mosaicism in *AR* gene in androgen target tissues, ethnic origin and normal CAG repeat range in the population. Undoubtedly, the level of androgen insensitivity will be grossly determined by type and site of molecular defect, but it will be fine-tuned by CAG repeat allele and other polymorphisms in *AR* and its interacting genes. Taking into consideration the importance of CAG repeats, we recommend CAG repeat length analysis in all the studies on *AR* gene in human subjects. Analysis of *AR* gene for CAG repeat length, *AR* somatic mosaicism, and pubertal hormone levels to calculate overall ASI will help in appropriate management of androgen insensitivity cases.

Acknowledgment

Financial support of Council of Scientific and Industrial Research, and Indian Council of Medical Research, Government of India, New Delhi is gratefully acknowledged.

References

1 Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS. Androgen receptor defects: historical, clinical, and molecular perspectives. Endocr Rev 1995; 16: 271–321.

2 Mebarki F, Forest MG, Josso N, Bertrand AM, David M, Chatelain P, *et al.* Complete direct sequencing of the entire AR gene in 45 unrelated patients with androgen insensitivity syndrome: mutations identified in 32 patients (18 novel mutations), no mutation detected in 13 other patients (29%). Am J Hum Genet 1994; 55: 231.

3 Batch JA, Williams DM, Davies HR, Brown BD, Evans BA, Hughes IA, *et al.* Androgen receptor gene mutations identified by SSCP in fourteen subjects with androgen insensitivity syndrome. Hum Mol Genet 1992; 1: 497–503.

4 Gottlieb B, Beitel LK, Wu JH, Trifiro M. The androgen receptor gene mutations database (ARDB): 2004 update. Hum Mutat 2004; 23: 527–33.

5 Hiort O, Sinnecker GH, Holterhus PM, Nitsche EM, Kruse K. Inherited and *de novo* androgen receptor gene mutations: investigation of single-case families. J Pediatr 1998; 132: 939–43.

6 Ris-Stalpers C, Trifiro MA, Kuiper GG, Jenster G, Romalo G, Sai T, *et al.* Substitution of aspartic acid-686 by histidine or asparagine in the human androgen receptor leads to a functionally inactive protein with altered hormone-binding characteristics. Mol Endocrinol 1991; 5: 1562–9.

7 Hiort O, Sinnecker GH, Holterhus PM, Nitsche EM, Kruse K. The clinical and molecular spectrum of androgen insensitivity syndromes. Am J Med Genet 1996; 63: 218–22.

8 MacLean HE, Ball EM, Rekaris G, Warne GL, Zajac JD. Novel androgen receptor gene mutations in Australian patients with complete androgen insensitivity syndrome. Hum Mutat 2004; 23: 287.

9 Nordenskjold A, Soderhall S. An androgen receptor gene mutation (A645D) in a boy with a normal phenotype. Hum Mutat 1998; 11: 339.

10 Wang Q, Ghadessy FJ, Yong EL. Analysis of the transactivation domain of the androgen receptor in patients with male infertility. Clin Genet 1998; 54: 185–92.

11 Gottlieb B, Vasiliou DM, Lumbroso R, Beitel LK, Pinsky L, Trifiro MA. Analysis of exon 1 mutations in the androgen receptor gene. Hum Mutat 1999; 14: 527–39.

12 Batch JA, Williams DM, Davies HR, Brown BD, Evans BA, Hughes IA, *et al.* Androgen receptor gene mutations identified by SSCP in fourteen subjects with androgen insensitivity syndrome. Hum Mol Genet 1992; 1: 497–503.

13 Yong EL, Tut TG, Ghadessy FJ, Prins G, Ratnam SS. Partial androgen insensitivity and correlations with the predicted three dimensional structure of the androgen receptor ligand-binding domain. Mol Cell Endocrinol 1998; 137: 41–50.

14 Gottlieb B, Beitel LK, Trifiro MA. Somatic mosaicism and variable expressivity. Trends Genet 2001; 17: 79–82.

15 Gottlieb B, Beitel LK, Trifiro MA. Variable expressivity and mutation databases: The androgen receptor gene mutations database. Hum Mutat 2001; 17: 382–8.

16 Holterhus PM, Sinnecker GH, Hiort O. Phenotypic diversity and testosterone-induced normalization of mutant L712F androgen receptor function in a kindred with androgen insensitivity. J Clin Endocrinol Metab 2000; 85: 3245–50.

17 Boehmer AL, Brinkmann AO, Nijman RM, Verleun-Mooijman MC, de Ruiter P, Niermeijer MF, *et al.* Phenotypic variation in a family with partial androgen insensitivity syndrome explained by differences in 5alpha dihydrotestosterone availability. J Clin Endocrinol Metab 2001; 86: 1240–6.

18 Rajender S, Singh L, Thangaraj K. Phenotypic heterogeneity of mutations in androgen receptor gene. Asian J Androl 2007; 9: 147–79.

19 Chamberlain NL, Driver ED, Miesfeld RL. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acids Res 1994; 22: 3181–6.

20 Kazemi-Esfarjani P, Trifiro MA, Pinsky L. Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenetic relevance for the (CAG) n-expanded neuronopathies. Hum Mol Genet 1995; 4: 523–7.

21 Thangaraj K, Joshi MB, Reddy AG, Gupta NJ, Chakravarty B, Singh L. CAG repeat expansion in the androgen receptor gene is not associated with male infertility in Indian populations. J Androl 2002; 23: 813–6.

22 Singh R, Shastry PK, Rasalkar AA, Singh L, Thangaraj K. A novel androgen receptor mutation resulting in complete androgen insensitivity syndrome and bilateral Leydig cell hyperplasia. J Androl 2006; 27: 510–6.

23 Heinlein CA, Chang C. Androgen receptor in prostate cancer. Endocr Rev 2004; 25: 276–308.

Edited by Dr Aleksander Giwercman