No association between mutations in the human androgen receptor GGN repeat and inter-sex conditions

K.B.Lundin¹,²,³, A.Giwercman¹,², J.Richthoff¹,², P-A.Abrahamsson² and Y.L.Giwercman²

¹Fertility Centre, Scanian Andrology Centre and ²Department of Urology, Malmö University Hospital, Lund University, SE 205 02 Malmö, Sweden
³To whom correspondence should be addressed: Wallenberg Laboratory, entrance 46, Malmö University Hospital, SE – 205 02 Malmö, Sweden. E-mail: Kristina.Lundin@kir.mas.lu.se

The functional role of the GGN repeat in the human androgen receptor gene is unknown, although mutations in this region have been found in patients with inter-sex conditions. We have investigated the prevalence of GGN mutations in the androgen receptor in the Swedish population and their relation to male reproductive function. A physical examination and semen analysis was carried out in 223 men under medical examination before military service and in 94 men referred due to infertility and having sperm concentrations <5×10⁶/ml. The GGN and CAG repeats in the androgen receptor gene were directly sequenced. Both populations contained two predominant alleles of 23 and 24 GGN repeats, 83.8 and 90.5% respectively. Four mutations, three in the conscripts and one among the infertile men, were found, resulting in three GGC to GGT substitutions and one GGT to GGC substitution. None of the men presented with genital abnormalities, but two conscripts had low ejaculate volumes (0.3 and 0.9 ml). All men carrying a mutation also had GGN lengths ≥24. Three subjects with GGN >24, with no mutations and with normal seminal volumes, were also found. Our findings indicate that point mutations in the GGN repeat are frequently found in the general male population (1.3%; 95% CI: 0.3–3.9%), but are usually not associated with profound changes in the male phenotype.

Key words: androgen insensitivity syndrome/androgen receptor/male reproduction/polymorphism

Introduction

The human androgen receptor (AR) is the main regulator of androgen signalling in the cell (Eder et al., 2001). A functional AR is essential for differentiation of the male external and internal genitalia during embryogenesis, development of secondary sex characteristics at puberty, and maintenance of spermatogenesis thereafter. The hormones regulating these processes are mainly testosterone and 5α-dihydrotestosterone, which both bind to the AR. A failure in normal androgen production or sensitivity can result in a spectrum of abnormalities, ranging from a complete female phenotype to male infertility or minor degrees of undervirilization without genital malformations (Quigley et al., 1995; Hiort et al., 1996; Giwercman et al., 2000; 2001).

The AR gene, which is located on Xq11–12, is composed of three different functional domains: an N-terminal transactivating domain, a central DNA-binding domain and a C-terminal ligand-binding domain (Lubahn et al., 1988).

The N-terminal transactivating domain is highly variable, due to two polymorphic amino acid stretches (Chang et al., 1988; Lubahn et al., 1988; Faber et al., 1989). The most amino terminal of these is a polyglutamine stretch, encoded by (CAG)ₙCAA. Further downstream, a polyglycine stretch encoded by (GGT)ₙGGG(GGT)ₙGGCₙ is present. This repeat is generally designated the GGN repeat.

Abnormal expansion of the CAG repeat is associated with spinal and bulbar muscular atrophy (SBMA), also known as Kennedy’s disease (La Spada et al., 1991). As affected males also present with signs of decreased androgen sensitivity and impaired sperm output, this segment has been extensively investigated regarding other male disorders such as infertility and prostate cancer, with inconclusive results (Tut et al., 1997; Giwercman et al., 1998b; Kittles et al., 2001; Mononen et al., 2002).

Whole cell transfection experiments using androgen receptor constructs harbouring 15, 20 and 31 CAG repeats respectively, have shown an inverse relationship between the length of the CAG repeat and the transactivating capacity (Tut et al., 1997).

There is a considerable amount of data linking the length of the CAG repeat to androgen sensitivity in vivo as well as in vitro; however, there are only a few studies considering the functional aspects of the variations in the GGN repeat. The number of repeats is reported to vary between eight and 26 in a Caucasian population (Kittles et al., 2001). The most frequent alleles in Caucasians consist of 22 and 23 or 23 and 24 GGN repeats, depending on the population studied (Macke et al., 1993; Sleddens et al., 1993; Correa-Cerro et al., 1999). Virtually all studies on the GGN segment have, however, been done using sequence intra-allelic size markers to estimate the length of the repeat, thereby missing possible nucleotide substitutions. This may partly be due to technical difficulties in sequencing the GC-rich GGN tract. Therefore, there is a lack of information considering the biomedical importance of the GGN repeat in general and of nucleotide variations in the initial 6 codons of the GGN repeat in particular.

In a study of 10 controls and 51 patients with the androgen insensitivity syndrome (AIS), three cases with a GGC to GGT transversion in GGN repeat number 7 were found. In addition, one case with GCC to GGT transversions in number 7 and 8 was also found (Lumbroso et al., 1997). Based on these findings it was suggested that nucleotide alterations in the initial part of the GGN repeat might be associated with inter-sex conditions. However, it has not been shown in a non-patient population that such mutations imply...
alterations in the normal male phenotype. This information is important for better understanding of the role of the AR gene in the regulation of male reproductive function.

The aims of our study were to sequence the GGN repeat of the AR gene in the general Swedish population, to analyse the GGN allele distribution, to determine the prevalence of nucleotide alterations in the initial part of the GGN repeat and possibly relate such alterations to specific phenotypes.

### Materials and methods

#### Subjects

**Military conscripts**

A total of 305 men under medical examination for military service were enrolled in a study of reproductive function in young Swedish men (Richthoff et al., 2002). As the AR gene is located on the X chromosome, in order to exclude any impact of ethnic variation, genotyping of the GGN repeats was performed in all men (n = 223) with Swedish mothers. These men were born during the period 1979–1982 and their mean age at the time of the study was 18.1 years. All participants underwent a physical examination including the genital organs. Every participant also delivered a semen sample for which he received £55. All men participated with written informed consent according to protocols approved by the ethical review board of Lund University.

**Infertile men**

Since data from the conscripts indicated that mutations in the GGN repeat might have an impact on ejaculate volume, 94 consecutive Swedish men from infertile couples, presenting with sperm concentrations <5 × 10^6/ml in at least two ejaculates, were included in the study. Men with known genetic causes of infertility, e.g. Klinefelter syndrome or Y chromosome microdeletions, were excluded.

#### Analysis of semen

All conscripts delivered one semen sample whereas the infertile men provided at least two ejaculates, which were obtained by masturbation after ≥48 h of sexual abstinence. The assessment of volume, concentration and motility was performed as recommended by the World Health Organization (WHO, 1999). In order to assess the ejaculate volume, the weight of the empty plastic cup was subtracted from the total weight of cup and semen. Concentration was assessed using a modified Neubauer chamber and positive displacement pipettes for proper dilution of the ejaculate. According to the WHO recommendations for determination of motility, 200 sperms were scored in categories A, B, C and D, with A corresponding to rapid progressive motility, B to slow progressive motility, C to non-progressive motility, and D to immotile sperm.

The laboratory participates in the external Quality Control Programme organized by Nordic Association of Andrology and European Society of Human Reproduction and Embryology.

Semen data for the whole cohort of the general population have been published elsewhere (Richthoff et al., 2003).

#### PCR amplification and sequencing

Genomic DNA was prepared from peripheral leukocytes and the GGN repeat was amplified in a 25 μl PCR reaction containing 10 ng DNA, 0.3 μmol/l of each of the primers: forward, CCAGAGTCGGAGACTACTACAICTTCC, and reverse, CCAGAACACAGTGCTACTGCC (Invitrogen, UK). 2.5 μmol/l MgCl₂, 200 μmol/l of dATP, dCTP and dTTP each, 100 μmol/l of dGTP and 7-deaza-dGTP (Roche Diagnostics, Sweden), 45 mmol/l KCl, 10 mmol/l Tris–HCl (pH 8.4 at 70°C) and 0.5 IU of Dynazyme DNA polymerase (Finnzymes Oy, Finland). Amplification was performed for 40 cycles. Each cycle included denaturation at 96°C for 1 min, primer annealing at 56°C for 1 min and primer extension at 72°C for 3 min with an initial denaturation step at 96°C for 3 min and a final extension step at 72°C for 7 min. PCR reactions were carried out in an Eppendorf Mastercycler® (Eppendorf, Germany), with PCR mix without DNA sample as a negative control.

One μl of each PCR product was used for subsequent nested amplification by incubating 30 cycles at 96°C for 1 min, 52°C for 20 s, and 72°C for 3 min, with an initial denaturation step at 96°C for 3 min and a final extension step at 72°C for 7 min. A 50 μl reaction was prepared, containing 1.2 IU of Dynazyme DNA polymerase (Finnzymes Oy), 0.25 μmol/l of each of the primers: forward, ATCCCACGCTGACTCATCAAA, and reverse, CCAGAACACAGTGCTACTGCC (Invitrogen), 200 μmol/l of each dNTP and 0.65 mol/l Betain® (Sigma–Aldrich Chemie GmbH, Germany). Pure PCR mix, as well as mix with 1 μl of the negative control from the primary reaction, served as negative controls.

Nest PCR products were purified using JetPure PCR purification kit (Genomed GmbH, Germany) according to the protocol provided by the manufacturer. Approximately 30 ng of the purified products were submitted to a 20 μl sequencing reaction with the CEQ Quickstart kit (Beckman Coulter, Sweden) and the forward primer AGCCGCCGCTTCCTCATCCT and the reverse, CCAGAACAACAGAGTGCTACTGCTTTCC (Invitrogen), as described by the manufacturer. Sequencing products were ethanol-precipitated according to standard procedures, and resuspended in sample loading solution provided in the kit. The samples were analysed externally on a Beckman Coulter CEQ 2000XL sequencer.

The CAG repeat was amplified and sequenced according to similar principles as for the GGN repeat, but without 7-deaza-dGTP and Betain.

In the subjects in whom a mutation in the GGN repeat was found, the whole AR gene was sequenced, according to methods described previously (Giwercman et al., 1998a).

#### Statistical analysis

Statistical calculations were performed by use of SPSS 11.0 software (SPSS Inc., USA). Spearman’s ρ was calculated for defining the bivariate

---

**Table I. Seminal parameters in subjects with GGN lengths ≥24, with (A–D) or without (E–G) mutations in the GGN repeat**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Subject</th>
<th>Population</th>
<th>No. GGN</th>
<th>No. CAG</th>
<th>Ejaculate vol. (ml)</th>
<th>Sperm concentration (×10^6/ml)</th>
<th>Progressively motile sperm (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₂₀</td>
<td>A</td>
<td>Conscript</td>
<td>27</td>
<td>24</td>
<td>0.3</td>
<td>27.7</td>
<td>4</td>
</tr>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₁₉</td>
<td>B</td>
<td>Conscript</td>
<td>24</td>
<td>18</td>
<td>0.9</td>
<td>53.5</td>
<td>65</td>
</tr>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₁₇</td>
<td>C</td>
<td>Conscript</td>
<td>26</td>
<td>22</td>
<td>3.6</td>
<td>8.5</td>
<td>61</td>
</tr>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₁₅</td>
<td>D</td>
<td>Infertile</td>
<td>27</td>
<td>24</td>
<td>2.6</td>
<td>1.2</td>
<td>23</td>
</tr>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₁₃</td>
<td>E</td>
<td>Conscript</td>
<td>25</td>
<td>22</td>
<td>5.1</td>
<td>46.2</td>
<td>81</td>
</tr>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₁₁</td>
<td>F</td>
<td>Conscript</td>
<td>25</td>
<td>19</td>
<td>3.7</td>
<td>38.2</td>
<td>71</td>
</tr>
<tr>
<td>(GGT)₇(GGG(GGT)₉(GGC)₈)₉</td>
<td>G</td>
<td>Infertile</td>
<td>25</td>
<td>23</td>
<td>3.0</td>
<td>0.2</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

Median (95% CI) values in the general population

<table>
<thead>
<tr>
<th>Median (95% CI)</th>
<th>World Health Organization (1999) guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 (16–24)</td>
<td>&gt;2.0</td>
</tr>
<tr>
<td>21 (17–29)</td>
<td>≥20.0</td>
</tr>
<tr>
<td>3.2 (1.0–5.8)</td>
<td>≥50</td>
</tr>
<tr>
<td>3.0 (1.6–307.3)</td>
<td>≥50</td>
</tr>
<tr>
<td>1.2</td>
<td>≥50</td>
</tr>
</tbody>
</table>
correlation between the lengths of the GGN and the CAG repeat. Binomial distribution was applied for calculation of 95% confidence interval for frequency of GGN mutations.

Results

Sequencing of the AR gene

Military conscripts
Among 223 men, three subjects (1.3%; 95% CI: 0.3–3.9%) had alterations in the GGN repeat. Two GGC to GGT transversions at number 7 and 9 (A and C) and one GGT to GGC transversion in GGN number 6 (B) were found (Table I). Affected subjects presented with the following genotypes: (GGT)$_2$GGG(GGT)$_3$(GGC)$_{20}$; (GGT)$_3$GGG(GGT)$_3$(GGC)$_{19}$; (GGT)$_3$GGG(GGT)$_3$(GGC)$_{19}$; (GGT)$_3$(GGC)$_{20}$GGT(GGC)$_{17}$. None of these three mutations implied any alterations in the amino acid sequence of the AR protein. No mutations in the GGN repeat were found in the remaining 220 men.

Subjects A, B and C presented with CAG lengths of 24, 18 and 22 respectively. No additional mutations in the remaining part of the AR gene were found in these subjects.

The number of GGN triplets in the samples studied varied between 10 and 27. Fourteen different GGN allele classes were found, including the alleles with a mutation (Figure 1). Of the 223 X chromosomes studied, by far the most common were 23 and 24 GGN repeats, 52.0 and 31.8% respectively. Only four subjects presented with GGN lengths >24 (Table I, subjects A, C, E and F).

A statistically significant negative correlation between the lengths of the GGN and CAG repeats was found (Spearman’s rho = -0.154 and P = 0.02). Subjects A, B and C were not included in this analysis.

Infertile men
In a study population of 94 men, one had the (GGT)$_3$GGG(GGT)$_3$(GGC)$_{20}$ genotype (1.1%; 95% CI: 0.03–5.8%), with a total number of 27 GGN and a CAG length of 24 (Table I, subject D). The GGN distribution in this group of men was approximately the same as for the military conscripts, with 23 and 24 GGN repeats being the most common alleles, 54.3 and 36.2% respectively. Two men had GGN >24 (Table I, subjects D and G).

Reproductive parameters and GGN mutations

No genital malformations were recorded in any of the four subjects with mutations in the GGN repeat. However, they all had abnormalities in their semen parameters (Table I). Subject A as well as subject B presented with low seminal volumes of 0.3 and 0.9 ml, respectively. Subject C had oligozoospermia with a sperm concentration of 3.3 × 10$^6$/ml. Subject D had a sperm concentration of 1.2 × 10$^6$/ml and normal semen volume. None of the subjects with low ejaculate volume reported spillage at the semen collection procedure. Among the remaining 220 men from the general population, we found two individuals with long (= 25) GGN repeats, but without any AR mutations. These men had normal semen parameters (Table I, subject E and F). Among the infertile men, in addition to subject D, one man without a mutation, but with a GGN length of 25, was found (Table I, subject G). He also presented with normal semen volume.

Discussion

In a population-based study we found three point mutations, two GGC to GGT transversions (A and C) and one GGT to GGC transversion (B) in the first nine GGN repeats of the AR gene. Two of these mutations resulted in an alteration of the number of codons preceding the GGC stretch; whereas the third was located in the initial part of the polymorphic GGC repeat. One GGC to GGT alteration was also found in an infertile man (D). No mutations in the remaining domains of the AR gene were found in any of the subjects.

The (GGT)$_3$GGG(GGT)$_3$(GGC)$_{20}$ genotype found in subjects A and D has previously been described in three family members with AIS or suspected AIS, as well as in one of the first AR to be cloned (Chang et al., 1988; Lumbreroso et al., 1997). Notably, consistent with our findings, each of the four GGN alleles having the internal (GGT)$_3$ variant were followed by a (GGC)$_{20}$ segment.

We did not find any indications of AIS in any of the men with a point mutation. The AIS is considered the most common identifiable cause of male pseudohermaphroditism. The phenotype encompasses a wide range of genital ambiguities, which classically have been divided into two forms: complete AIS (CAIS) and partial AIS (PAIS), according to the severity. CAIS is a relatively rare disorder with a prevalence of 1:20 000 male births, according to a large Danish study that utilized a nationwide patient registry (Bangsboell et al., 1992). Because PAIS comprises a great variability of clinical manifestations, which sometimes are difficult to diagnose, the prevalence is unknown, but may be as rare as the complete form. However, all cases of AIS taken together, the prevalence of ~1:10 000 is far below our finding of 3.223 (1.3%) in the general Swedish population. Thus, these calculations, based on epidemiological observations, indicate that only an extremely low proportion, if any, of the type of mutations reported by us, is associated with AIS.

It cannot be excluded that these genotype alterations might have some impact on reproductive function. Two men (A and B) presented with an ejaculate volume below the 5th percentile of the remaining group of military conscripts, whereas subject (C) had a sperm concentration in the lower range of this reference interval. These men also had relatively long GGC stretches. A central question is whether the long GGN segments or the mutations per se might cause the seminal abnormalities seen in these men. Against length as a causative factor may be cited the presence of two additional conscripts and one infertile man with long GGN stretches, but without any mutations. These men presented with normal ejaculate volumes, which might indicate that the extended number of GGN repeats does not affect the function of the accessory sex glands. One could also speculate that the sequence variants found by us, although apparently not associated with changes in the amino acid sequence of the AR protein, could, however, still affect the function of the receptor. Others have previously suggested such an effect of a silent mutation on the
function of a gene. It has been reported that changes in the secondary mRNA structure could make it more unstable than the wild type mRNA (Chen et al., 1999). Another suggested mechanism for a functional role of silent mutations is an alteration in kinetics of protein translation, which can affect the in-vivo protein folding, leading to increased levels of protein misfolding (Komar et al., 1999). It could also be a combination of the two phenomena, long GGN segments together with a mutation that plays a role for AR function and subsequent androgen-dependent events in the body, such as spermatogenesis. Furthermore, we cannot rule out that our observations could be chance findings. In-vitro studies on the mutations, in combination with different GGN lengths, are needed in order to elucidate their specific roles for receptor function.

The distribution of the GGN lengths among the infertile men did not differ from that in the general population. This finding is in accordance with the only study to date, presented on GGN repeats in infertile men (Tut et al., 1997). However, the subset of infertile men selected for our study presented with extremely low sperm concentrations, and it cannot be excluded that the GGN repeat might be involved in the pathogenesis of other types of male infertility.

We found a weak but statistically significant correlation between the length of the GGN repeat and the number of CAG triplets in exon 1 of the AR gene. Such a correlation was previously found in patients with prostate cancer, but not in men from the general population, possibly due to the mixed ethnicity of the study population (Irvine et al., 1995). Although the lengths of the two repeats were correlated with each other, the magnitude of the correlation coefficient indicates that <2.5% of the variation in the length of one repeat can be explained by variation in the length of the other repeat. Thus, we conclude that their variation, at least in our population, is largely mutually independent.

In summary, in a study population of 223 unselected young men, we have found three divergences from the common GGN genotypes, of which two have not been described before. All men presented with normal male phenotype, although with slightly impaired seminal parameters. A GGN alteration was also found in an infertile man. The first codons of the GGN repeat of the AR gene are generally believed to be a constant (GGTG&G&). This study is to our knowledge the first report showing that variants are frequently found (>1%) in men without genital malformations. Although men presenting with inter-sex conditions are probably not presenting for the medical board before military service, based on the extremely low incidence of AIS (<1:10 000), we conclude that the mutations found by us are usually not associated with profound changes in the male phenotype. Further studies are needed in order to elucidate the relevance of the androgen receptor GGN repeat to male reproductive function.

Acknowledgements

We thank Camilla Anderberg for excellent technical assistance. This work was supported by grants from Swedish governmental funding for clinical research, the Swedish Cancer Society (Grant No: 4423), The Swedish Research Council (Grant No. 521-2002-3907), the Gunnar Nilsson Cancer Fund, the Crafoordska Fund, and Ove Tulefjords Fund and Foundation for Urological Research.

References


reproductive function in young males from the general Swedish population. *Environ. Health Perspect.*, in press.


Submitted on February 17, 2003; accepted on March 11, 2003