

ORIGINAL ARTICLE

Androgen receptor gene haplotype is associated with male infertility

M. Saare,* A. Belousova,† M. Punab,‡ M. Peters,* K. Haller,*§ K. Ausmees,‡ O. Poolamets,‡ H. Karro,* A. Metspalu† and A. Salumets*†

*Department of Obstetrics and Gynaecology, University of Tartu, Tartu, Estonia, †Institute of Molecular and Cell Biology, Department of Biotechnology, University of Tartu, Tartu, Estonia, ‡Andrology Unit of Tartu University Clinicum, Tartu, Estonia, and § Institute of General and Molecular Pathology, Department of Immunology, University of Tartu, Tartu, Estonia

Summary

Keywords:

androgen receptor gene, gene haplotypes, male infertility, microsatellite analysis

Correspondence:

Andres Salumets, Department of Obstetrics and Gynaecology, University of Tartu, Lossi 36, Tartu 51003, Estonia.
E-mail: asalumets@novavita.ee

Received 11 December 2006;
revised 24 January 2007, 18 February 2007;
accepted 22 February 2007

doi:10.1111/j.1365-2605.2007.00782.x

The purpose of the current study was to evaluate the importance of androgen receptor (*AR*) gene haplotypes and polymorphic CAG/GGN microsatellites in the aetiology of male infertility. We genotyped six haplotype-tagging single nucleotide polymorphisms and CAG/GGN microsatellites of the *AR* gene in 112 infertile and 212 control Estonian men. A total of 13 *AR* haplotypes (HAP1–13) were identified, among which HAP4 was found to confer increased risk for male infertility (OR = 5.15, 95% CI = 1.75–15.15, $p = 0.003$). However, infertile patients and controls had similar lengths and distributions of both *AR* CAG (mean \pm SD number of repeats 21.1 ± 2.5 vs. 21.2 ± 2.3 , respectively) and GGN (mean \pm SD number of repeats 22.5 ± 1.5 vs. 22.4 ± 1.9 , respectively) repeats. In addition, HAP2 was associated with more CAG repeats ($r = 1.17$, $p = 0.033$) and HAP3 with fewer CAG repeats ($r = -2.93$, $p < 0.001$) than the major haplotype HAP1. HAP3 and HAP4 were associated with more GGN repeats ($r = 1.35$, $p = 0.001$ and $r = 1.36$, $p = 0.002$, respectively) than HAP1. In conclusion, our results implicated the *AR*-HAP4 gene haplotype in increased risk for male infertility, while no association was found between *AR* CAG/GGN microsatellites and impaired spermatogenesis.

Introduction

Androgens, mainly testosterone and 5- α -dihydrotestosterone, play a key role in stimulating and maintaining spermatogenesis through specific binding to the androgen receptor (*AR*), which initiates transcription of androgen-responsive genes. The *AR* gene is located on the X chromosome at Xq11–q12, and consists of eight exons coding for four domains: an N-terminal transcription activation domain, a DNA-binding domain, a hinge region and a C-terminal ligand-binding domain (Lubahn *et al.*, 1988). *AR* gene exon 1 harbours CAG and GGN variable repeat motifs that are indispensable for *AR* transcription activation function. The CAG tract, coding for a polyglutamine stretch, typically ranges from 11 to 35 repeats (Giwerzman *et al.*, 1998; Kittles *et al.*, 2001). In vitro studies have suggested that progressive expansion of *AR* CAG repeats results in reduced *AR* transcriptional activity

(Chamberlain *et al.*, 1994). The less variable polyglycine-coding GGN trinucleotide region typically has 22–24 repeats (Lundin *et al.*, 2003), with the *AR* protein containing 23 glycines exhibiting the highest transactivating capacity (Lundin *et al.*, 2007). Several single nucleotide polymorphisms (SNPs) contribute additional variability to the *AR* gene, with six haplotype-tagging SNPs (htSNPs) responsible for the determination of all the common haplotypes of the *AR* gene (Terry *et al.*, 2005).

The presence of genetic heterogeneity and the possible physiological importance of the allelic variations in the *AR* gene have lead many researchers to explore the relationship between *AR* gene variations and male infertility. Previous studies in different ethnic populations have been contradictory, with some studies reporting substantially longer CAG tracts in infertile males than in controls (Tut *et al.*, 1997; Dowsing *et al.*, 1999; Mifsud *et al.*, 2001; Patrizio *et al.*, 2001; Wallerand *et al.*, 2001; Mengual *et al.*,

2003), while others found no correlation between the number of AR-CAG repeats and impaired spermatogenesis (Dadze *et al.*, 2000; Thangaraj *et al.*, 2002; Van Golde *et al.*, 2002; Lund *et al.*, 2003; Ferlin *et al.*, 2004; Ruhayel *et al.*, 2004). Research elucidating the importance of GGN repeat numbers in the aetiology of male infertility has, so far, consistently demonstrated a lack of causal association between GGN repeat number and failure of spermatogenesis (Tut *et al.*, 1997; Ferlin *et al.*, 2004; Ruhayel *et al.*, 2004; Rajender *et al.*, 2006). However, the length of the GGN segment, in combination with CAG repeat size might, in some ethnic groups, modify the AR protein function, leading to varying degrees of impaired spermatogenesis (Ferlin *et al.*, 2004; Ruhayel *et al.*, 2004). So far, no studies have investigated the possible link between the AR gene haplotypes and diminished sperm production. Therefore, the purpose of the current study was to investigate the relationship between AR gene haplotypes, CAG/GGN microsatellite polymorphisms and the incidence of male infertility.

Materials and methods

Subjects

A total of 112 male patients from the Andrology Unit of Tartu University Clinicum were recruited in 2005. All study subjects reported an infertility period that exceeded 12 months, and underwent a detailed medical history, general physical examination and a comprehensive evaluation of fertility status, which included assessment of the scrotal content by palpation, determination of testicular size using a Prader orchidometer, measurements of testosterone (T), follicle-stimulating hormone (FSH) and luteinizing hormone (LH) levels by chemiluminescence immunoassay (Immulite 2000; DPC, Los Angeles, CA, USA), and analysis of semen samples according to WHO guidelines (WHO, 1999). Infertile males with non-obstructive azoospermia and oligozoospermia ($\leq 20 \times 10^6$ sperm cells per mL of ejaculate) were eligible for the study (WHO, 1999), while patients with known causes of

male infertility, such as chromosomal abnormalities, Y chromosome microdeletions and pathologies of the epididymis and/or vas deferens were excluded. The majority (92.9%, 104/112) of recruited patients exhibited either non-obstructive azoospermia (34.8%, 39/112) or severe oligozoospermia with sperm concentration $\leq 5 \times 10^6$ /mL (58.1%, 65/112). Only eight men (7.1%) had a sperm count between 6 and 14×10^6 /mL. Two hundred and twelve controls were selected to achieve twice the number of controls compared with the cases. Study controls were selected based on the best sperm concentration from the young men who underwent compulsory medical examination for possible military service (Jensen *et al.*, 2004). The lower limit of sperm count for controls was set $>40 \times 10^6$ /mL (Bonde *et al.*, 1998). Both the study and control groups of the current study consisted of mixed Estonian-Russian population. The study was approved by the Tallinn Medical Research Ethics Committee, and informed consent was obtained from all participants after the nature of the study was explained to them.

SNP genotyping

A total of six htSNPs (SNP1 – rs962458 T/C SNP in the upstream region of exon 1; SNP2 – rs6152 A/G synonymous SNP in exon 1; SNP3 – rs1204038 C/T SNP in intron 1; SNP4 – rs2361634 A/G SNP in intron 1; SNP5 – rs1337080 G/A SNP in intron 2 and SNP6 – rs1337082 C/T SNP in the downstream region of exon 8) were chosen for AR gene haplotype determination (Fig. 1), as previously described (Terry *et al.*, 2005). Restriction length polymorphism analysis and allele-specific polymerase chain reaction (PCR) were utilized to establish four (SNP1, SNP2, SNP4 and SNP5) and two (SNP3 and SNP6) htSNPs, respectively. Genomic DNA was isolated from the peripheral EDTA-blood using a salting-out method (Aljanabi & Martinez, 1997). PCR was carried out in a total volume of 15 μ L containing 50 ng DNA, 0.2 mM dNTPs, 2.5 mM MgCl₂, 1 \times PCR buffer (Solis BioDyne, Tartu, Estonia), 10 pmol primers, and 0.8 U

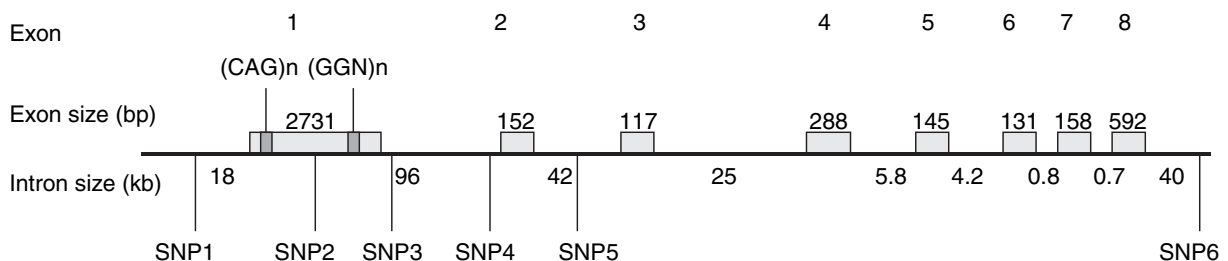


Figure 1 The structure of AR gene, with polymorphic CAG and GGN microsatellite repeats and htSNPs indicated. The sizes of exons and introns are indicated in basepairs and kilobases, respectively.

Table 1 The primer sequences and restriction enzymes used for androgen receptor (*AR*) haplotype determination

Polymorphisms	Primer sequences (5'–3') ^a	Restriction enzymes	Alleles	DNA fragment sizes (bp)
SNP1 (rs962458) ^b	F-CCGTAGCCTTCTGGAAAACATC	<i>Bsh</i> 1236I	T	149
	R-GTCACAATTTACTTAAAAATGCGC		C	24 + 125
SNP2 (rs6152) ^b	F-CTCCGGACGAGGATGACTCA	<i>Eco</i> 147I	A	255
	R-TGGCGTTGTCAGAAATGGTCGAA		G	186 + 69
SNP3 (rs1204038) ^c	F1-GAACTTGGTCTAATTCCTTC	–	C	160
	R-CTTCGCTAGACACGAGTTCA		T	
SNP4 (rs2361634) ^b	F-ATGAGGTAAAGTTACAAACCTGG	<i>Xce</i> I	A	344 + 54
	R-AACATGGTCCCTGGCAGTCTC		G	224 + 120 + 54
SNP5 (rs1337080) ^b	F-GAATTGGAGGGAGATAGGTG	<i>Nde</i> I	G	268
	R-CTCAATGAGGACAAGGAATC		A	137 + 131
SNP6 (rs1337082) ^c	F-GGAGGCATCTTGAGACCAAT	–	C	217
	R1-ATCTTGAACAATTATTTAACCTTC		T	
	R2-ATCTTGAACAATTATTTAACCTTC			

^aF and R indicate, respectively, forward and reverse primers;

^bSingle nucleotide polymorphisms (SNPs) determined using restriction fragment polymorphism analysis;

^cSNPs determined using allele-specific PCR analysis.

Hot Start thermostable DNA polymerase HOT FIREPol (Solis BioDyne). Reaction mixtures were preheated (95 °C, 15 min) and were followed by PCR with 35 cycles of amplification (95 °C, 30 sec; 56–63 °C, 30 sec and 72 °C, 30 sec). The PCR primer sequences and restriction enzymes (MBI Fermentas, Vilnius, Lithuania) used are listed in Table 1. The PCR products, as well as the restriction fragments, were detected on 2% agarose gel containing 10 µg/mL ethidium bromide, and were visualized by UV transillumination.

Microsatellite genotyping

The number of trinucleotides was used to characterize *AR* gene CAG and GGN variations (Fig. 1). The *AR* gene CAG region was amplified using the forward 6-FAM-5'-GTCTA CCCTCGGCCCGCCGTC-3' and reverse 5'-TAGCCTGTGG GGCTCTACG-3' primers, while TAMRA-5'-CCTGGCA CACTCTCTTACACA-3' and 5'-ACATCAGGTGCGGTGAA GTC-3' were, respectively, used as forward and reverse primers for GGN region amplification. CAG and GGN variable repeats were amplified and detected separately. The reaction mixture (15 µL) contained 50 ng DNA, 0.2 mM dNTPs, 2.5 mM MgCl₂, 1× PCR buffer, 10 pmol primers, and 0.8 U Hot Start thermostable DNA polymerase HOT FIREPol. Additionally, 2× solution S (Solis BioDyne) was used in amplification of GGN fragments. The PCR protocol utilized 34 amplification cycles with the denaturation, annealing and extension carried out at 95 °C, 30 sec; 60 °C, 30 sec and 72 °C, 30 sec, respectively. The size of amplification products was determined twice by electrophoresis under standard conditions on an ABI

Prism 377 DNA Sequencer, using Genescan 2.1 software (PE Applied Biosystems, Foster City, CA, USA). Rox 500 (PE Applied Biosystems) was used as an internal size markers standard. Size-marker genotyping of CAG and GGN repeats was confirmed by DNA sequencing of 40 patients and controls using BigDye Terminator v3.1 cycle sequencing kit (PE Applied Biosystems) and ABI 3730 XL DNA Sequencer.

Statistical analysis

Mann–Whitney *U* tests were conducted for the comparisons of clinical parameters. Differences in CAG and GGN repeat numbers were evaluated by *t*-tests. Linkage disequilibrium analysis of htSNPs utilized the Haploview program, version 3.2 (<http://www.broad.mit.edu/mpg/haploview>), using pairwise combination of all htSNPs and Lewontin's *D'* statistics (Gabriel *et al.*, 2002). Examination of the influence of *AR* gene haplotypes on male fertility potential utilized logistic and Poisson's regression analyses. All statistical tests were carried out using the R2.3.1 A Language and Environment (Free Software Foundation, Boston, MA, USA) with *p* < 0.05 considered statistically significant.

Results

The clinical parameters of the study participants are summarized in Table 2. As expected, infertile patients had significantly lower T values, smaller mean testicular volumes and lower sperm concentration and progressive motility compared with controls. Additionally, eight infertility

Table 2 The comparisons of fertility characteristics between the patients and controls, using Mann–Whitney *U* test

Characteristics	Patients (<i>n</i> = 112)			Controls (<i>n</i> = 212)			<i>p</i> value
	Median	Minimum	Maximum	Median	Minimum	Maximum	
Age (years)	31.0	22.0	49.0	18.0	17.0	25.0	–
Mean testicular volume (mL)	18.0	3.0	50.0	26.0	15.0	45.0	<0.001
Duration of abstinence (days)	4.0	1.0	8.0	4.6	0	13.0	<0.001
Ejaculate volume (mL)	3.5	0	10.2	3.2	0.9	8.9	NS
Sperm concentration (10 ⁶ /mL)	0.8	0	14.0	122.0	75.0	716.0	<0.001
Sperm progressive motility (%)	20.0	0	74.0	57.5	30	80	<0.001
Serum testosterone (nmol/L)	16.9	3.5	49.2	20.6	10.9	48.6	<0.001
Serum FSH (IU/L)	9.4	1.3	62.6	3.1	0.7	9.3	<0.001
Serum LH (IU/L)	4.7	0.6	23.8	2.9	0.9	10.2	<0.001

NS, Statistically not significant; FSH, follicle-stimulating hormone; LH, luteinizing hormone.

Haplotype number	Haplotype	Patients % (<i>n</i>)	Controls % (<i>n</i>)	OR (95% CI), <i>p</i> value
HAP1 ^a	TGCAAT	67.0 (75)	75.9 (161)	1
HAP2	TGCGAT	3.6 (4)	7.1 (15)	0.57 (0.18–1.78), 0.336
HAP3	CATAGC	7.1 (8)	5.2 (11)	1.56 (0.60–4.04), 0.359
HAP4	TATAAC	10.7 (12)	2.4 (5)	5.15 (1.75–15.15), 0.003
Others	–	11.6 (13)	9.4 (20)	1.40 (0.66–2.95), 0.384

Table 3 Prevalence and odds ratios (ORs) of androgen receptor (*AR*) gene haplotypes among infertile and control men

^aHAP1 was used as the baseline genotype and 95% confidential interval (CI).

patients (7.1%) had T levels below the lower limit of 9.9 nmol/L for normal adult men.

In total, six htSNPs were genotyped to define common *AR* haplotypes. The analysis demonstrated that all htSNPs analysed were characterized by strong linkage disequilibrium ($D' \geq 0.8$) forming a haploblock of approximately 238 kb in size. Haplotype analysis revealed 13 haplotypes (HAP1–13), with the frequencies for the four most common haplotypes (HAP1–4) given in Table 3. Our results suggested that *AR*-HAP4 conferred significantly increased risk for male infertility (OR = 5.15, 95% CI = 1.75–15.15, $p = 0.003$), if the main HAP1 haplotype was used as the baseline genotype. The results were the same, if HAP2, HAP3, other haplotypes (HAP5–13), or all but HAP4 were used as a baseline haplotypes. Our data also suggested that the A allele of SNP2 (OR = 2.15, 95% CI = 1.17–3.96, $p = 0.014$), the T allele of SNP3 (OR = 1.90, 95% CI = 1.00–3.60, $p = 0.051$) and the C allele of SNP6 (OR = 2.02, 95% CI = 1.14–3.60, $p = 0.017$) were risk alleles for male infertility. All these results were the same, if patients with moderate oligozoospermia ($n = 8$, 7.1%, sperm count between 6 and 14×10^6 /mL) were excluded from the analysis.

The HAP4 showed the association with the diminished sperm count, as the frequency of HAP4 increased along with the decrease of sperm concentration in the following order: controls with normozoospermia (2.4% of HAP4,

5/212) – patients with oligozoospermia (6.9%, 5/73) – patients with azoospermia (12.8%, 5/39) (OR = 2.60, 95% CI = 1.61–4.20, $p < 0.001$, Poisson regression analysis). The associations between the presence of HAP4 and the serum hormonal (T, FSH and LH) levels or mean testicular volume were assessed by utilizing Mann–Whitney *U* tests. HAP4 was neither correlated with serum hormonal levels nor mean testicular volume in both the patients and controls (data not shown).

No significant differences were noted in *AR*-CAG (mean \pm SD were 21.1 ± 2.5 and 21.2 ± 2.3 , respectively; *t*-test, $p = 0.806$) and GGN (mean \pm SD were 22.5 ± 1.5 and 22.4 ± 1.9 , respectively; *t*-test, $p = 0.672$) repeat numbers or distributions between infertile patients and controls (Fig. 2). *AR*-CAG alleles ranged from 16 to 34 repeats in the infertile patients and from 16 to 28 repeats in the controls. Meanwhile, *AR*-GGN sizes ranged from 17 to 27 repeats in the infertile patients and from 16 to 27 repeats in the controls (Fig. 2). Our data suggested that there was a positive correlation (Pearson's correlation $r = 0.147$, $p = 0.032$) between the CAG and GGN segment lengths in the *AR* gene among controls, but not in infertile men (Pearson's correlation $r = 0.093$, $p = 0.329$). The most prevalent *AR* gene CAG/GGN combinations were CAG = 21/GGN = 22 (7.1%, 8/112) and CAG = 22/GGN = 22 (6.6%, 14/212) in infertile and control men, respectively. No specific *AR*-CAG/GGN

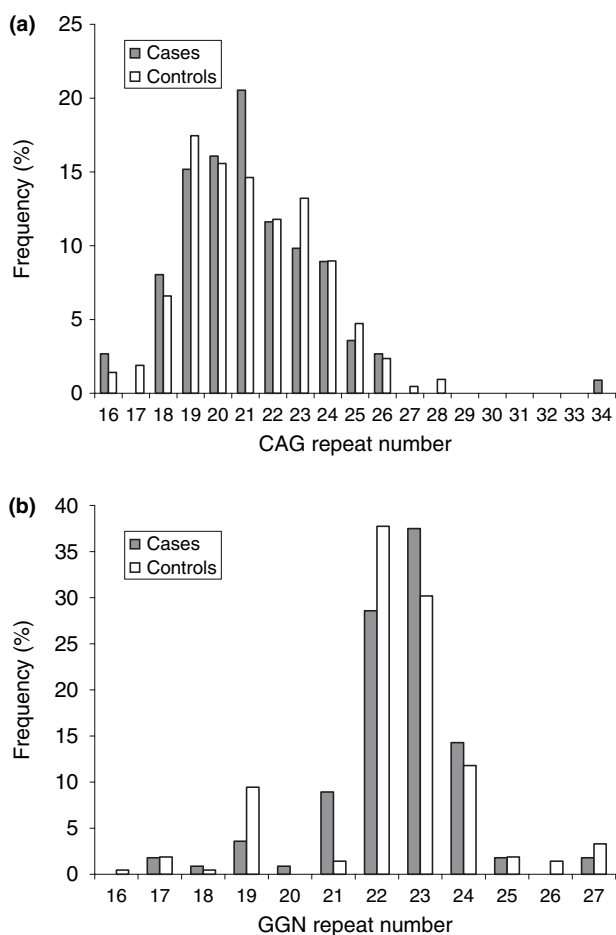


Figure 2 The distribution of CAG (A) and GGN (B) allele frequencies in infertile and control men.

Table 4 The associations between androgen receptor (AR) haplotypes and the lengths of CAG and GGN repeats among both cases and controls. Means of CAG and GGN repeats were compared between the different AR haplotypes by using linear regression models (r , regression coefficient)

AR haplotypes	Mean CAG repeat		Mean GGN repeat	
	no. \pm SD	r (p value)	no. \pm SD	r (p value)
HAP1 ^a	21.4 \pm 2.5	–	22.2 \pm 1.9	–
HAP2	22.5 \pm 1.9	1.17 (0.033)	22.4 \pm 0.5	0.14 (NS)
HAP3	18.4 \pm 0.7	–2.93 (<0.001)	23.6 \pm 0.6	1.35 (0.001)
HAP4	20.8 \pm 0.8	–0.45 (NS)	23.6 \pm 0.5	1.36 (0.002)

NS, Statistically not significant.

^aHAP1 was used as the baseline genotype.

combination was found to affect the risk for male infertility (data not shown).

Further analysis revealed that AR haplotypes predicted the lengths of CAG and GGN repeats both in cases and

controls utilizing linear regression models (Table 4). HAP2 was found to be associated with significantly longer mean CAG repeats ($r = 1.17$, $p = 0.033$) and HAP3 with shorter mean CAG repeats ($r = -2.93$, $p < 0.001$) compared with the reference HAP1 haplotype. In addition, HAP3 and HAP4 were associated with significantly longer GGN repeats ($r = 1.35$, $p = 0.001$ and $r = 1.36$, $p = 0.002$, respectively) compared with the AR-HAP1 haplotype.

Discussion

Androgens and normally functioning AR are critical for spermatogenesis as indicated by identification of numerous deleterious AR gene mutations in men (Yong *et al.*, 2003). However, the AR gene mutations are only very seldom responsible for low or absent sperm production in phenotypically normal men, leading for search of relationships between the AR gene variations only slightly altering the AR function and the incidence of male infertility. In the current study, we aimed to evaluate the predictive value of AR gene variants on the prevalence of male infertility.

A novel aspect of the current study was the examination of the correlation between AR gene haplotypes and impaired spermatogenesis. The studies examining AR haplotypes in ovarian (Terry *et al.*, 2005) and endometrial (McGrath *et al.*, 2006) cancers have found no significant haplotype effects. However, previous population-based case-control studies have been inconclusive regarding the relationship between AR gene haplotypes and the prostate cancer risk (Freedman *et al.*, 2005; Lindstrom *et al.*, 2006). In the current study, the genotyping of six htSNPs revealed the occurrence of 13 different AR haplotypes. The three major haplotypes for controls were HAP1 (75.9%), HAP2 (7.1%) and HAP3 (5.2%), comprising 88.2% of alleles. While the three most prevalent haplotypes (84.8%) in the infertile group were HAP1 (67.0%), HAP4 (10.7%) and HAP3 (7.1%). The three less common AR-HAP2–4 haplotypes have diverged from the most predominant AR-HAP1 haplotype, leading to HAP2 (SNP4 A to G change) in one direction, and HAP4 (SNP2 G to A, SNP3 C to T and SNP6 T to C changes) and eventually HAP3 (SNP1 T to C and SNP5 A to G changes) in the other direction (Lundin *et al.*, 2006).

We strongly suggest a four-fold increase in male infertility associated with the AR-HAP4 haplotype, as there was a substantially higher incidence of infertile men carrying AR-HAP4 (10.7%, 12/112) compared with controls (2.4%, 5/212). Recently, AR-HAP4 containing X chromosomes was found in 6.1% of general Swedish male population (Lundin *et al.*, 2006). The difference of AR-HAP4 prevalence in our control group (2.4%) and that of Lun-

din *et al.* (2006) (6.1%) can be because of the peculiarities of control group formation. Although in both studies controls were recruited among military conscripts, the controls in our study were selected based on the highest sperm count, while in the Swedish study, controls were invited to the study irrespective of the sperm parameters.

The current study also suggested that certain SNPs were associated with an elevated risk for male infertility, with the logistic regression analysis showing that the A allele of SNP2, the T allele of SNP3 and the C allele of SNP6 constituted risk alleles for male infertility. On the contrary, recent studies have demonstrated the AR gene SNP2 (rs6152) G allele associated with elevated prostate cancer risk (Hayes *et al.*, 2005; Lindstrom *et al.*, 2006). These reports along with our present data corroborate the putative opposite effect of AR SNP2 genotype on prostate cancer and infertility risk.

Our findings are predominantly restricted to severe male infertility patients of our mixed Estonian-Russian population, as majority of the patients (92.9%, 104/112) were either azoospermic or severely oligozoospermic (sperm count $\leq 5 \times 10^6$ /mL of ejaculate) and only a few patients had moderate oligozoospermia (WHO, 1999). The current study suggests that HAP4 associates with the diminished sperm count, in particular, as the prevalence of HAP4 increased in the order of normozoospermia - oligozoospermia - azoospermia. At the same time, the presence of HAP4 seemed not to be correlated with the levels of serum hormones or with the mean testicular volume. Therefore, our results could shed light on the putative reasons of idiopathic male infertility, as only patients lacking the known causes for infertility were analysed. However, the association between HAP4 and serum hormonal levels need further evaluation in a population study rather than in case-control one, where study subjects might not represent the normal range of hormonal values. In these future studies, AR-HAP4 can be discriminated from the remaining 12 neutral haplotypes by detection of only three SNP1, SNP2 and SNP3 loci. As all the SNPs determining the AR haplotypes are either non-coding or synonymous, additional research is also warranted to fully understand the impacts of AR haplotypes on AR gene/protein function and on the possible pathogenesis of male infertility of unknown cause.

The role of functional AR gene CAG/GGN microsatellites in the aetiology of male infertility has been the focus of much research. Previous studies exploring the modulating effect of the AR-CAG repeats in the impaired sperm production came up with conflicting results. The conclusions of the current study, showing similar mean values as well as the distributions of AR-CAG repeats among infertile patients and controls, are in agreement with previous studies of European (Dadze *et al.*, 2000; Van

Golde *et al.*, 2002; Lund *et al.*, 2003; Ferlin *et al.*, 2004; Ruhayel *et al.*, 2004) and Asian (Thangaraj *et al.*, 2002) populations. Other studies, however, have found longer polyglutamine tracts in the AR gene of male infertile patients compared with controls in European (Wallerand *et al.*, 2001; Mengual *et al.*, 2003) or multi-ethnic American, Asian and Australian (Tut *et al.*, 1997; Dowsing *et al.*, 1999; Mifsud *et al.*, 2001; Patrizio *et al.*, 2001) populations. Previous research has also suggested that 28 or more CAG repeats in the AR can result in greater than a four-fold higher risk of male infertility, while the probability of defective spermatogenesis is halved when the tract contains ≤ 23 CAG repeats (Tut *et al.*, 1997). However, in the current study, only one patient in the infertility group possessed an X chromosome with an AR gene containing more than 28 CAG repeats (34 CAG repeats), while in the control group, two men had AR-CAG alleles with 28 repeats. Additionally, our data demonstrated that 83.9% (94/112) of men in the infertility group and 82.5% (175/212) of controls had AR-CAG alleles with ≤ 23 repeats. Although the exact reason is still unknown, ethnic differences, patient recruitment biases and sample size restrictions can be the likely sources for the conflicting results of studies analysing the causal link between AR gene CAG polymorphisms and the occurrence of male infertility.

Few studies have focussed on the role of GGN polyglycine variability in impaired spermatogenesis (Tut *et al.*, 1997; Ferlin *et al.*, 2004; Rajender *et al.*, 2006). Our results are in agreement with previous studies, with no differences noted in the mean values nor in the distributions of GGN repeat lengths between infertile (mean of 22.5 repeats) and control (mean of 22.4 repeats) men. However, male infertility-related AR-HAP4 haplotype was associated with significantly longer GGN repeats compared with the ancestral AR-HAP1 haplotype, as also found in an earlier study (Lundin *et al.*, 2006).

Although it seems likely that AR gene GGN and CAG variations independently exert limited influence on AR activity, the combined GGN and CAG AR gene variants may still modify the susceptibility to male infertility (Ferlin *et al.*, 2004; Ruhayel *et al.*, 2004). The combination of GGN = 23 and < 21 CAG may confer a lower risk of male infertility, with four times more carriers of this GGN/CAG combination found among fertile controls (16%) than among infertile patients (4%), as demonstrated in a recent study of Swedish men (Ruhayel *et al.*, 2004). Our results, however, did not suggest any GGN/CAG AR gene variant to be more prevalent in infertile Estonian men compared with controls.

In conclusion, we have described the AR gene variations in infertile and control men in Estonian population, using six mutually associated htSNPs and

polymorphic CAG/GGN microsatellites. The AR gene haplotype analysis revealed, for the first time, the existence of an AR gene risk haplotype (HAP4) for impaired spermatogenesis and male infertility, while no effects of CAG/GGN polymorphisms on male fertility potential were observed. However, additional epidemiological case-control studies are needed to validate these findings for other ethnic groups.

Acknowledgements

This study was supported by the Estonian Science Foundation Grants 6498 and 6585; by the Estonian Ministry of Education and Science Core Grants 0182582Cs03, PBGMR07903 and 0182582s03; by the EU FP6 Grant LSHB-CT-2004-503243; and by the Artur Lind Gene Technology Grant Fond of Estonian Gene Centre.

References

- Aljanabi, S. M. & Martinez, I. (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* 25, 4692–4693.
- Bonde, J. P., Ernst, E., Jensen, T. K., Hjollund, N. H., Kolstad, H., Henriksen, T. B., Scheike, T., Giwercman, A., Olsen, J. & Skakkebaek, N. E. (1998) Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet* 352, 1172–1177.
- Chamberlain, N. L., Driver, E. D. & Miesfeld, R. L. (1994) The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Research* 11, 3181–3186.
- Dadze, S., Wieland, C., Jakubiczka, S., Funke, K., Schröder, E., Royer-Pokora, B., Willers, R. & Wieacker, P. F. (2000) The size of the CAG repeat in exon 1 of the androgen receptor gene shows no significant relationship to impaired spermatogenesis in an infertile Caucasoïd sample of German origin. *Molecular Human Reproduction* 6, 207–214.
- Dowsing, A. T., Yong, E. L., Clark, M., McLachlan, R. I., de Kretser, D. M. & Trounson, A. O. (1999) Linkage between male infertility and trinucleotide repeat expansion in the androgen-receptor gene. *Lancet* 21, 640–643.
- Ferlin, A., Garolla, A., Foresta, C., Bartoloni, L., Rizzo, G. & Roverato, A. (2004) Androgen receptor gene CAG and GGC repeat lengths in idiopathic male infertility. *Molecular Human Reproduction* 10, 417–421.
- Freedman, M. L., Pearce, C. L., Penney, K. L., Hirschhorn, J. N., Kolonel, L. N., Henderson, B. E. & Altshuler, D. (2005) Systematic evaluation of genetic variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study. *American Journal of Human Genetics* 1, 82–90.
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B. et al. (2002) The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229.
- Giwercman, Y. L., Xu, C., Arver, S., Pousette, A. & Reneland, R. (1998) No association between the androgen receptor gene CAG repeat and impaired sperm production in Swedish men. *Clinical Genetics* 54, 435–436.
- Hayes, V. M., Severi, G., Eggleton, S. A., Padilla, E. J., Southey, M. C., Sutherland, R. L., Hopper, J. L. & Giles, G. G. (2005) The E211 G>A androgen receptor polymorphism is associated with a decreased risk of metastatic prostate cancer and androgenetic alopecia. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology* 14, 993–996.
- Jensen, T. K., Jorgensen, N., Punab, M., Haugen, T. B., Suominen, J., Zilaitiene, B. et al. (2004) Association of in utero exposure to maternal smoking with reduced semen quality and testis size in adulthood: a cross-sectional study of 1,770 young men from the general population in five European countries. *American Journal of Epidemiology* 159, 49–58.
- Kittles, R. A., Young, D., Weinrich, D., Hudson, J., Argyropoulos, G., Ukoli, F., Adams-Campbell, L. & Dunston, G. M. (2001) Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: implications for prostate cancer risk. *Human Genetics* 109, 253–261.
- Lindstrom, S., Wiklund, F., Adami, H. O., Balter, K. A., Adolfsson, J. & Gronberg, H. (2006) Germ-line genetic variation in the key androgen-regulating genes androgen receptor, cytochrome P450, and steroid-5-alpha-reductase type 2 is important for prostate cancer development. *Cancer Research* 66, 11077–11083.
- Lubahn, D. B., Joseph, D. R., Sullivan, P. M., Willard, H. F., French, F. S. & Wilson, E. M. (1988) Cloning of human androgen receptor complementary DNA and localisation of the X chromosome. *Science* 240, 327–330.
- Lund, A., Tapanainen, J. S., Lahdetie, J., Savontaus, M. L. & Aittomaki, K. (2003) Long CAG repeats in the AR gene are not associated with infertility in Finnish males. *Acta Obstetrica et Gynecologica Scandinavica* 82, 162–166.
- Lundin, K. B., Giwercman, A., Richthoff, J., Abrahamsson, P. A. & Giwercman, Y. L. (2003) No association between mutations in the human androgen receptor GGN repeat and inter-sex conditions. *Molecular Human Reproduction* 9, 375–379.
- Lundin, K. B., Nordenskjold, A., Giwercman, A. & Giwercman, Y. L. (2006) Frequent finding of the androgen receptor A645D variant in normal population. *The Journal of Clinical Endocrinology and Metabolism* 91, 3228–3231.
- Lundin, K. B., Giwercman, A., Dizelyi, N. & Giwercman, Y. L. (2007) Functional in vitro characterisation of the androgen receptor GGN polymorphism. *Molecular and Cellular Endocrinology* 264, 184–187.
- McGrath, M., Lee, I. M., Hankinson, S. E., Kraft, P., Hunter, D. J., Buring, J. & De Vivo, I. (2006) Androgen receptor polymorphisms and endometrial cancer risk. *International Journal of Cancer* 1, 1261–1268.

- Mengual, L., Oriola, J., Ascaso, C., Balleca, J. L. & Olivia, R. (2003) An increased CAG repeat length in the androgen receptor gene in azoospermic ICSI candidates. *Journal of Andrology* 24, 279–284.
- Mifsud, A., Sim, C. K., Boettger-Tong, H., Moreira, S., Lamb, D. J., Lipshultz, L. I. & Yong, E. L. (2001) Trinucleotide (CAG) repeat polymorphisms in the androgen receptor gene: molecular markers of risk for male infertility. *Fertility and Sterility* 75, 275–281.
- Patrizio, P., Leonard, D. G., Chen, K. L., Hernandez-Ayup, S. & Trounson, A. O. (2001) Larger trinucleotide repeat size in the androgen receptor gene of infertile men with extremely severe oligozoospermia. *Journal of Andrology* 22, 444–448.
- Rajender, S., Rajani, V., Gupta, N. J., Chakravarty, B., Singh, L. & Thangaraj, K. (2006) No association of androgen receptor GGN repeat length polymorphism with infertility in Indian men. *Journal of Andrology* 27, 785–789.
- Ruhayel, Y., Lundin, K., Giwercman, Y., Hallden, C., Willen, M. & Giwercman, A. (2004) Androgen receptor gene GGN and CAG polymorphisms among severely oligozoospermic and azoospermic Swedish men. *Human Reproduction* 19, 2076–2083.
- Terry, K. L., De Vivo, I., Titus-Ernstoff, L., Shih, M. C. & Cramer, D. W. (2005) Androgen receptor cytosine, adenine, guanine repeats, and haplotypes in relation to ovarian cancer risk. *Cancer Research* 65, 5974–5981.
- Thangaraj, K., Joshi, M. B., Reddy, A. G., Gupta, N. J., Chakravarty, B. & Singh, L. (2002) CAG repeat expansion in the androgen receptor gene is not associated with male infertility in Indian populations. *Journal of Andrology* 23, 815–818.
- Tut, T. G., Ghadessy, F. J., Trifiro, M. A., Pinsky, L. & Yong, E. L. (1997) Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. *Journal of Clinical Endocrinology and Metabolism* 82, 3777–3782.
- Van Golde, R., Van Houwelingen, K., Kiemeny, L., Kremer, J., Tuerlings, J., Schalken, J. & Meuleman, E. (2002) Is increased CAG repeat length in the androgen receptor gene a risk factor for male subfertility? *Journal of Urology* 167, 621–623.
- Wallerand, H., Remy-Martin, A., Chabannes, E., Bermont, L., Adessi, G. L. & Bittard, H. (2001) Relationship between expansion of the CAG repeat in exon 1 of the androgen receptor gene and idiopathic male infertility. *Fertility and Sterility* 76, 769–774.
- WHO. (1999) WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction.
- Yong, E. L., Loy, C. J. & Sim, K. S. (2003) Androgen receptor gene and male infertility. *Human Reproduction Update* 9, 1–7.