Nuclear receptors and disease: androgen receptor

Bruce Gottlieb*†, Lenore K. Beitel*‡, Jianhui Wu*, Youssef A. Elhaji*§ and Mark Trifiro*‡§

*Lady Davis Institute for Medical Research, Sir Mortimer B. Davis-Jewish General Hospital, 3755 Chemin de la Côte-Ste-Catherine, Montréal, Québec, Canada H3T 2E1, †Department of Biology, John Abbott College, P.O. Box 2000, Ste Anne de Bellevue, Québec, Canada H9X 3L9, ‡Department of Medicine, McGill University, Montréal, Québec, Canada H3T 1E2, and §Department of Human Genetics, McGill University, Montréal, Québec, Canada H3T 1E2

Abstract

The androgen receptor (AR) protein regulates transcription of certain genes. Usually this depends upon a central DNA-binding domain that permits the binding of androgen–AR complexes to regulatory DNA sequences near or in a target gene. The AR also has a C-terminal ligand-binding domain and an N-terminal transcription modulatory domain. These N- and C-terminal domains interact directly, and with co-regulatory, non-receptor proteins, to exert precise control over a gene’s transcription rate. The precise roles of these proteins are active research areas. Severe X-linked AR gene (AR) mutations cause complete androgen insensitivity, mild ones impair virilization with or without infertility, and moderate ones yield a wide phenotypic spectrum sometimes among siblings. Different phenotype expressivity may reflect variability of AR-interactive proteins. Mutations occur throughout the AR but are concentrated in specific areas of the gene known as hot spots. A number of these mutations

1To whom correspondence should be addressed, at the Lady Davis Institute for Medical Research (e-mail bruce.gottlieb@mcgill.ca).
of somatic origin are associated with prostate cancer. N-terminal polyglutamine (polyGln) tract expansion reduces AR transactivation, and when there are more than 38 glutamine residues it causes spinobulbar muscular atrophy, a motor neuron disease, due to a gain of function. Variations in polyGln tract length have been associated as risk factors with prostate, breast, uterine, endometrial and colorectal cancer, as well as male infertility.

Introduction

The androgen receptor [AR; Online Mendelian Inheritance in Man (OMIM) code 31370] is a member of the superfamily of nuclear receptors that function as ligand-dependent transcription factors. The gene for the AR is approx. 90 kb, with eight exons, and lies on the X chromosome at Xq11–12. Like other nuclear receptors, the AR protein contains four major domains, the N-terminal domain, the DNA-binding domain (DBD), the hinge region and the androgen- or ligand-binding domain (LBD). The DBD and LBD show considerable homology with other nuclear receptors. The DBD contains two zinc fingers and is required for androgen-response-element recognition. The 253 residue C-terminal LBD contains 12 α-helices and a highly hydrophobic ligand-binding site. Intracellular AR is essential for androgen action, whether the ligand is testosterone or its 5α-reduced derivative, 5α-dihydrotestosterone. Hence, the AR is essential for normal primary male sexual development before birth (masculinization), and for normal secondary male sexual development around puberty (virilization). AR dysfunctions in XY individuals result in androgen-insensitivity syndromes (AISs).

Structure–function relationships of the AR protein and gene

In the cell cytoplasm the unliganded AR is associated with a number of heat-shock proteins (hsp). The AR is activated by the binding of an androgen molecule, usually 5α-dihydrotestosterone, and the release of hsp. Ligand binding also promotes AR hyperphosphorylation. The transformed AR assumes an altered conformation in which two zinc fingers in the DBD are exposed (Figure 1). It is then translocated into the nucleus, where homodimerization occurs and the AR homodimer binds to specific androgen-response elements in certain genes acquiring the ability to regulate the rate of transcription of these genes (Figure 1). To exert such regulation, a complex of an androgen and an AR must also interact with transcriptionally active proteins that bind upstream and downstream of the androgen-response
element. These interactions, with a number of co-regulators, including basal transcription factors and core promoter element-binding proteins, determine the vectorial control over the transcriptional expression of a given androgen target gene. The mechanism of action of the AR are reviewed in a recent paper by Gobinet et al. [1]. The nature and possible function of these co-regulators have been reviewed by Heinlein and Chang [2], and a comprehensive database of these co-regulatory proteins is available in the AR Gene Mutations Database (http://www.mcgill.ca/androgendb).

While the AR is about 90 kb long, only approx. 2.75 kb, divided into eight exons, codes for amino acids. The arrangement of the exons and introns is shown in Figure 2. The variable length of the AR protein reflects the fact that its N-terminal transregulation modulatory portion (approx. 537 amino acids) contains two homopolymeric amino acid ‘repeats’ that are polymorphic in size (Figure 2): the polyglutamine (polyGln) repeat size varies from 11 to 37 [3];

![Figure 2. Major structure–function domains and putative subdomains of the AR gene and protein](image)

Note that the thicker portions of the lines below the protein structure indicate the most likely location of a given functional domain, with the exception of the phosphorylation sites, which are at specific numbered residues. Zn$^{2+}$ refers to the zinc fingers; AF refers to the transcriptional activation function regions. The nature of these domains is reviewed in a number of papers [1,7,10].
the other, polyglycine, varies from 12 to 29 [4]. The central DBD (approx. amino acids 557–616) is encoded by exons 2 and 3. Adjacent to the DBD, C-terminally, there is a hinge region, (amino acids 617–663) encoded by exons 3 and 4. Finally, the C-terminal LBD (approx. 253 amino acids, starting at amino acid 664) is encoded by exons 4–8. In addition to their principal functions, the LBD, DBD, hinge and N-terminal domains embody subsidiary functions that affect dimerization (involving N- and C-terminal and LBD–LBD interactions), nuclear localization, transcriptional regulation, hsp binding and phosphorylation [1] (Figure 2). In addition, a number of activation function domains (AF) has been identified (Figure 2). Thus the tetramodular (domain) concept of AR function is a simplification; instead, domain interaction, and interaction with co-regulatory proteins, hold the secrets to a full understanding of an AR’s structure–function properties.

Diseases as a result of mutations in the AR gene

Loss-of-function diseases: AISs

AISs (OMIM # 300068) can be subdivided into three phenotypes: complete AIS (CAIS), partial AIS (PAIS) and mild AIS (MAIS). AR mutations that severely impair the amount, structure or function of the AR cause the complete androgen insensitivity phenotype. Standard references quote rates of 2–5 in 100 000 for complete androgen insensitivity. Subjects are born looking unambiguously female because 5α-dihydrotestosterone-dependent masculinization of the external genital primordia is totally absent. These individuals are typically not suspected of being abnormal until the onset of puberty, when breast development is normal, but pubic and axillary hair are not. Menarche, initially considered ‘late’, never occurs. Müllerian duct regression, being androgen-independent, is normal. Their testes may or may not be inguinal.

Because all XY subjects with complete androgen insensitivity are sterile (genetic lethals), one-third of their mutant alleles should represent new mutations. A recent report on single-case families with complete or partial androgen insensitivity gene mutations [5] revealed a de novo AR mutation rate of close to 30% (8 out of 30), therebyaffirming the theoretical expectation of 33% for an X-linked recessive genetic lethal [6].

Partial androgen insensitivity has a highly variable phenotypic expression. At one extreme, the external genitalia are near-normal female, except for clitoromegaly and/or posterior labial fusion; at the other extreme, the genitalia may be morphologically normal male, but small, or there may be simple coronal hypospadias or a prominent midline raphe of the scrotum. In between these extremes are all grades of frank external genital ambiguity that are, nonetheless, predominantly masculine or predominantly feminine. Indeed, for some mutations in the LBD, such variable expressivity may be the rule, not the exception
[7]. Furthermore, in rare families with partial androgen insensitivity, the expressivity may vary markedly, from near-normal male to near-normal female.

Mild androgen insensitivity takes two phenotypic forms at puberty: in one, spermatogenesis and fertility are impaired [8]; in the other, spermatogenesis is normal, or sufficient to preserve fertility [9]. In both, gynaecomastia, high-pitched voice, sparse sex hair and impotence may be noted. In the form where fertility is preserved, one presumes that the dysfunction of the mutant AR is sufficiently mild so that it can be overcome by collaboration with the set of co-regulatory proteins that is active in Sertoli cells [10].

Phenotype–genotype correlation of AR mutations
The present version of the Androgen Receptor Gene Mutation Database [11] (see http://www.mcgill.ca/androgendb), contains 450 entries of mutations causing AIS representing over 300 different AR mutations from more than 600 patients with AIS and demonstrates an unequal distribution of these mutations along the length of the AR, as shown in Table 1. It has previously been suggested that mutation-dense regions are hot spots that reflect the high density of mutable CpG sites in the region [12]. It is also apparent that the types of mutations differ along the length of the AR. In particular, nearly all mutations in exon 1 (Table 1) cause complete androgen insensitivity, and nearly all are of the premature translation termination variety, whether by direct mutation to a stop codon or indirectly by frameshifts after small deletions or insertions. To date, only 54 mutations have been reported in exon 1 of the AR in patients suffering from some form of AIS, despite the fact that it encodes more than half of the AR protein [11], and even fewer in splicing and untranslated regions of the AR gene (Table 1).

In the LBD there is a striking preponderance of missense mutations with a significantly greater number of CAIS than PAIS cases (Table 1). In an effort to better understand the structure–function relationship of how specific mutations in the LBD cause AIS, the X-ray crystal structure of the AR LBD was solved [13]. This revealed a structure that consisted of 12 α-helices (Figure 3). However, the putative crystal structure only identified 18 specific residues in the LBD that were in close contact with the ligand-binding pocket when modelled with the synthetic ligand R1881. This meant that the vast majority of AIS mutants whose ARs exhibited reduced or even no ligand binding could not be explained by the crystal structure. A selection of CAIS, PAIS and MAIS mutations are shown in Figure 3, which shows that the nature of their AIS phenotypes has little to do with their distance from the ligand-binding pocket. Therefore, to try to elucidate how such mutations could effect the structure of the ligand-binding pocket, and so explain the lack of ligand binding in these mutants, we have recently used molecular dynamic modelling techniques over extended periods of time (up to 4 ns) to, in effect, create four-dimensional structures of AR mutants [14].
Table 1. Nature and distribution of unique AR mutations that cause disease


<table>
<thead>
<tr>
<th>Loss-of-function disease</th>
<th>Type of mutation</th>
<th>Number of mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N-terminal domain</td>
</tr>
<tr>
<td>CAIS</td>
<td>Single-base substitution</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Premature termination</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Complete gene deletion</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Partial gene deletion</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Deletion (1–4 bases)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Insertion</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Duplication</td>
<td>2</td>
</tr>
<tr>
<td>PAIS</td>
<td>Single-base substitution</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Multiple-base substitution</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Premature termination</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Deletion (1–4 bases)</td>
<td>1</td>
</tr>
<tr>
<td>MAIS</td>
<td>Single-base substitution</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Partial gene deletion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deletion (1–4 bases)</td>
<td></td>
</tr>
</tbody>
</table>

continued
<table>
<thead>
<tr>
<th>Loss-of-function disease</th>
<th>Type of mutation</th>
<th>Number of mutations</th>
<th>N-terminal domain</th>
<th>DBD</th>
<th>Hinge region</th>
<th>LBD</th>
<th>Splice site</th>
<th>UTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate cancer</td>
<td>Single-base substitution</td>
<td>25</td>
<td>6</td>
<td>2</td>
<td>38</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Premature termination</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deletion (1–4 bases)</td>
<td>5</td>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Insertion</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>Single-base substitution</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larynx cancer</td>
<td>Deletion (30 bases)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2004 The Biochemical Society
In these studies a CAIS mutant not close to the ligand-binding pocket causes a local structural distortion that also affects the ligand-binding pocket. Thus, modelling has resulted in a better understanding of how alterations in residues that are not close to the putative binding pocket affect ligand binding. In future, such studies could ultimately be used to create androgen analogues with which to treat AIS patients.

The issue of variable expressivity is one of the most puzzling in trying to understand the genotype–phenotype relationship of many LBD mutations. The traditional explanation for such variable expressivity of an AR mutation is that the level or competence of co-regulatory proteins act as genetic ‘background’ factors in determining the overall clinical outcome [7]. Recently, however, it has been appreciated that somatic mosaicism (more or less covert), may account for some variable expressivity [15]. The simplest origin of such mosaicism would be forward mutation of an inherited normal allele to a mutant allele in a subject with a negative family history. However, in a family with multiple affected individuals, a relatively mild clinical outcome could reflect a back mutation of an inherited mutation to a normal allele.
Gain-of-function diseases

Prostate cancer (CaP)
To date 85 AR mutations have been found in CaP tissue, almost all being single-base substitutions due to somatic rather than germline mutations, and the role of the AR in CaP has recently been reviewed by Debes and Tindall [16]. As can be seen in Table 1, while the majority occur in the LBD (approx. 45%), a substantial number occur in exon 1 (approx. 30%). Originally it was thought that AR was not expressed in CaP, but this does not appear to be the case. In fact, the level of AR expression is significantly higher in hormone-resistant tumours compared with hormone-sensitive tumours [17]. Considerable controversy has revolved around conflicting studies that only sometimes find a significant number of AR mutations in CaP [18]. It has been argued that AR mutations only appear during the latter stages of CaP and, in addition, some studies have indicated that anti-androgen treatments have resulted in AR mutations. There is evidence from study of somatic AR mutations in prostatic carcinoma that certain ones of the missense variety not only permit the AR to bind unusual androgens or other steroids [19], promiscuously, but also allow it to be activated by them [20] in a manner that allows these unorthodox steroid–receptor complexes to be effective in transcriptional regulation of certain androgen target genes. Further, recent studies have found that a significant percentage of recurrent CaP following treatment have somatic mutations in their ARs and that the localization of mutations are influenced by the type of treatment [21]. However, by far the most significant event from a clinical prospective is that all CaP tumours eventually become androgen-resistant or -independent [22], and that in a number of cases this coincides with the appearance of mutations in the AR.

Male breast cancer
To date only two cases of AR mutations associated with male breast cancer have been reported and in both cases they have been in individuals suffering from PAIS [23]; the mutations being in adjacent codons in exon 3. It therefore seems likely that the AR mutations are not a primary cause of male breast cancer, which rather may possibly be due to the increased incidence of gynaecomastia associated with AIS.

Diseases directly associated with AR CAG repeat-length variation: spinobulbar muscular atrophy (SBMA; Kennedy disease)
Kennedy disease (OMIM no. 31320) is a spinobulbar motor neuronopathy associated with mild androgen insensitivity, and is one of the classic trinucleotide-repeat-expansion diseases that cause inherited neurogenerative disorders [24]. It is caused by expansion of the glutamine-coding (CAG)$_{8-35}$ CAA tract in exon 1 of the AR to a total of 38 or more [25]. It is an adult-onset...
motor neuronopathy that typically causes slowly progressive, symmetric wasting and weakness, initially of the proximal muscles of the hip and shoulder. Muscle cramps, hand tremors and fasciculations are often associated. Eventually, motor neurons of the brainstem become involved, leading to speech and swallowing difficulties. Male hypogonadism, usually represented by gynaecomastia and testicular atrophy, and attributable to mild androgen insensitivity, is not infrequent. The mild androgen-insensitivity component of SBMA may reflect a loss of AR transcriptional regulatory activity by virtue of a pathologically expanded polyGln tract. It should be noted that, in SBMA, the androgen-insensitivity phenotype is quite variable.

Since subjects with complete androgen insensitivity, including those with complete AR deletions, do not develop SBMA, this knowledge mandated the logic that the polyCAG-expanded AR or the polyGln-expanded AR protein is somehow motor neuronotoxic by a gain, not a loss, of function.

It is now clear that the polyGln-expanded AR protein, and probably the expanded polyGln tracts themselves, with certain flanking segments of the various parental proteins, are the essential pathogenetic agents. Expansion to beyond approx. 38 repeats endows a polyGln tract with a threshold property (or more than one) that must be selectively lethal to certain neurons. The biochemical, histopathological and neurophysiological features of SBMA are, unremarkably, those secondary to motor denervation. A number of possible causes for this gain of function are listed in Table 2, and these were analysed and discussed in a recent review by La Spada and Taylor [24].

The mild androgen-insensitivity component of SBMA is clearly due to a loss of function by the polyGln-expanded AR protein. Since androgens are both motor neuronotropic and motor neuronotrophic, it is possible that the polyGln-expanded AR protein loses a function that is necessary, but not sufficient, for the motor neuronopathy of SBMA generally or for death of certain motor neurons specifically.

The polyGln-expanded proteins or fragments thereof may oligomerize, directly or through intermediates, either non-covalently (by hydrogen-bonded polar zippering) or covalently (by transglutaminase-catalysed isodipeptide formation), to yield inclusions (aggregates) that accumulate in or around the nucleus of certain neurons [25]. Why neurons are selectively vulnerable to the toxicity of polyGln-expanded proteins when these proteins (e.g. the AR) are widely distributed in many non-neuronal cells and why only certain motor neurons are affected remains entirely speculative.

**AR CAG tract-length variation as a risk factor for disease**

**Female breast cancer**

Previous investigations into the relationship of CAG repeat lengths in the AR with female breast cancer have yielded somewhat confusing results [26]. There
Table 2. Diseases associated with AR CAG-tract-length variation

<table>
<thead>
<tr>
<th>Direct association</th>
<th>CAG tract</th>
<th>Androgen sensitivity</th>
<th>Gain of function: possible causes</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBMA</td>
<td>&gt;38</td>
<td>Reduced</td>
<td>1. Misfolding</td>
<td>1. Adult-onset motor neuropathy of proximal hip and shoulder muscles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Truncation</td>
<td>2. Hypogonadism results in gynaecomastia and testicular atrophy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Aggregation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Sequestration of AR protein/ transcription factors</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Proteosome inhibition</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Mitochondrial dysfunction</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indirect association</th>
<th>CAG tract</th>
<th>Androgen sensitivity</th>
<th>Associated risk factors</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaP</td>
<td>Shorter length</td>
<td>Increased</td>
<td>Ethnicity and family history</td>
<td>Inconclusive studies: possible somatic alterations</td>
</tr>
<tr>
<td>Male infertility</td>
<td>Longer length</td>
<td>Reduced</td>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Female breast cancer</td>
<td>Longer length</td>
<td>Reduced</td>
<td>BRCA1 mutation carriers</td>
<td>Inconclusive studies: possible somatic alterations</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>Longer length</td>
<td>Reduced</td>
<td>Somatic alterations</td>
<td></td>
</tr>
<tr>
<td>Colon cancer</td>
<td>Shorter length</td>
<td>Increased</td>
<td>Selective growth advantage: somatic alterations</td>
<td></td>
</tr>
</tbody>
</table>

is also a suggestion that CAG repeats might be a significant modifier to the
BRCA1 mutation associated with breast cancer risk [27], although recently
even this observation has been questioned [28]. Interestingly, decreased AR
transactivational activity lowers androgen/oestrogen balance, and may thereby
effect functional hypooestrogenicity. This may promote the pathogenesis of
breast cancer. To elucidate whether longer CAG repeats of the AR, which
correlate with lower transactivational activity of the AR, are associated with
breast cancer in women over 40, we have examined the distribution of CAG
repeat lengths in breast cancer tissue from this population [29]. The breast
cancer tissue was histologically graded as grade 1, or well differentiated, grade
2, or moderately differentiated, and grade 3, or poorly differentiated. Analysis
showed significant differences as compared with controls when CAG lengths
greater than 21 were examined, and that alleles with 26 repeats or more were
2.4-fold more frequent in breast cancer samples than in constitutional samples
from a normal population. A significant shift to greater CAG repeat lengths
appeared in grade 1 and 2 tumours. Our results give some indication as to the
progression of breast cancer by suggesting that hypotransactive ARs with long
polyGln tracts may have a role in the initiation and/or progression of breast
cancer, which has also been noted in a recent review [30]. In addition grade 3
tumours, perhaps due to increased genomic instability, tended to have shorter
than normal CAG repeat lengths. In this case it is hypothesized that the ARs
have now become hypertransactive, possibly coinciding with the oestrogen
resistance that is associated with grade 3 tumours. Whether this shift is of
germline or somatic origin was not clear, though the appearance in 14% of the
breast cancer samples of a third CAG repeat length indicates that it may be
somatic [26].

**Male infertility**
In Singapore, otherwise normal males with 28 or more CAG repeats in their
AR have been reported to have more than a 4-fold increased risk of impaired
spermatogenesis, and the more severe the spermatogenic defect, the greater the
chance of finding a long repeat [31]. While some additional studies have
supported this observation that longer CAG repeats are associated with
infertility [32], not all have done so [33].

**CaP**
Due to androgen activity being inversely proportional to polyGln tract length
[34], there has been considerable speculation as to a possible relationship
between tract length and CaP. The only significant association seems to be
related to the fact that the length of the tract shows ethnic variation, and that
this may be one cause for the higher risk of CaP development in, for example,
African-Americans [35]. However, to date, the results have been largely
inconclusive, as reported in an extensive review by Ferro et al. [26]. Their
overall conclusion is that only in combination with other polymorphisms such
as that of the prostate-specific antigen (PSA) protein could CAG-tract-length variation in the AR be considered a significant risk factor. To date, however, the most severe limitations of the studies is that they have not examined somatic alterations in AR CAG repeats in CaP tissues, as previous studies have shown that AR CAG repeats have a high degree of somatic instability [36]. It is hoped that with the ability to examine repeat length in very specific CaP tissues from specific prostate tumours, a better understanding of any possible relationship will emerge.

Uterine endometrial cancer
Only one study to date has examined the possible effect of a CAG-repeat expansion as a risk factor, which may reduce the antagonist effect of the androgens in counteracting the proliferative effect of oestrogens [37].

Colorectal cancer
Recently it has been shown that in 10% of colon cancer samples there was a somatic AR CAG-repeat reduction [26]. The exact mechanism of how such instability could lead to tumour growth was strictly the subject of speculation, although the suggestion was made that it might be due to such cells having a selective growth advantage [26].

Conclusions
The ability to correlate specific gene alterations with both phenotype and structure–function of the AR is likely to lead to considerable insight into how these alterations directly cause the expression of disease phenotypes. This is likely to be accomplished using a number of new techniques such as laser capture microdissection and four-dimensional protein modelling. Of particular interest will be our ability to begin to explain such concepts as phenotypic variable expressivity, gain-of-function disease phenotypes and the influence on phenotype of somatic versus germline gene alterations. Finally, the importance of somatic alterations in polymorphic trinucleotide repeats, as risk factors for certain diseases, is likely to be further elucidated.

Summary
• Gene alterations in the AR cause AIS with a range of phenotypic expression from complete to mild AIS.
• Some individual AR alterations have a significant degree of variable expressivity which in some cases is caused by somatic mosaicism.
• Somatic AR alterations have been found in CaP tissues in which the AR protein appears to have an acquired gain of function.
• Expansion of a CAG repeat in exon 1 of the AR results in the motor neuron disease SBMA due to a gain of function of the AR.
• Variations in the length of the AR CAG tract have been identified as a possible risk factor for prostate, female breast, uterine endometrial and colon cancer, as well as male infertility.

We thank the Canadian Institutes of Health Research for supporting our own work on AR mutations and the various diseases associated with them.

References


