

## PERSPECTIVE

# Androgen receptor polyglutamine repeat number: models of selection and disease susceptibility

Calen P. Ryan\* and Bernard J. Crespi

Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada

**Keywords**

androgens, balancing selection, cancer, disease susceptibility, infertility, sexual conflict, tandem-repeat accumulation

**\*Correspondence**Calen P. Ryan, Department of Biological Sciences, Simon Fraser University, 8888 University Dr, Burnaby, BC V5A 1S6, Canada.  
Tel.: +778 782 3986;  
fax: +778 782 3496;  
e-mail: calen\_ryan@sfu.ca

Received: 18 January 2012

Accepted: 4 May 2012

doi:10.1111/j.1752-4571.2012.00275.x

**Abstract**

Variation in polyglutamine repeat number in the androgen receptor (AR CAGn) is negatively correlated with the transcription of androgen-responsive genes and is associated with susceptibility to an extensive list of human disease. Only a small portion of the heritability for many of these diseases is explained by conventional SNP-based genome-wide association studies, and the forces shaping AR CAGn among humans remains largely unexplored. Here, we propose evolutionary models for understanding selection at the AR CAG locus, namely balancing selection, sexual conflict, accumulation-selection, and antagonistic pleiotropy. We evaluate these models by examining AR CAGn-linked susceptibility to eight extensively studied diseases representing the diverse physiological roles of androgens, and consider the costs of these diseases by their frequency and fitness effects. Five diseases could contribute to the distribution of AR CAGn observed among contemporary human populations. With support for disease susceptibilities associated with long and short AR CAGn, balancing selection provides a useful model for studying selection at this locus. Gender-specific differences AR CAGn health effects also support this locus as a candidate for sexual conflict over repeat number. Accompanied by the accumulation of AR CAGn in humans, these models help explain the distribution of repeat number in contemporary human populations.

**Introduction**

An applied understanding of the evolutionary forces shaping human health and disease susceptibility has profound medical implications, providing clinical insights and suggesting novel, testable hypotheses (Di Rienzo 2006; Nesse 2011). Robust integration of evolutionary theory with human medicine must continue to address topics whose resolution eludes research in each field independently. In particular, identifying the extent and regions of the human genome most directly subjects to balancing selection (Andrés et al. 2009), and mutation-selection balance (Keller and Miller 2006; Haerty and Golding 2010b), and explaining the 'missing heritability' of complex diseases (Manolio et al. 2009), have remained problematic, despite their central importance for the interfaces of disease biology with human evolution.

Tandem-repeat polymorphisms can be key components of human phenotypic and disease-related variation (Ashley

and Warren 1995; Sutherland and Richards 1995; Koshy and Zoghbi 1997; Reddy and Housman 1997; Gatchel and Zoghbi 2005). These genomic elements are largely undetectable using SNP-based genome-wide association (GWA) studies (Hannan 2010), yet they constitute about 3% of the human genome, inhabit as many as 10–20% of human genes and promoters, exhibit mutation rates that are orders of magnitude higher than those of SNPs, show high levels of within-population variation, and exert graded as well as discrete functional effects on gene expression and activity via a suite of documented molecular mechanisms (Kashi and King 2006; Gemayel et al. 2010; Kelkar et al. 2011). Over 40 human diseases, as well as continuous variation in phenotypes related to morphology, physiology and behavior, have been associated with tandem-repeat variation (Fondon and Garner 2004; Pearson et al. 2005; Fondon et al. 2008; Gemayel et al. 2010). As such, tandem-repeat loci represent excellent candidates for the effects of balancing selection (Mularoni et al. 2010) and mutation-selection

balance (Haerty and Golding 2010b), as well as mediating notable potential for rapid, adaptive evolution (Birge et al. 2010).

Two well-studied human tandem-repeat polymorphisms are situated within the X-linked androgen receptor (AR) gene, which codes for a transcription factor that mediates binding of the androgens testosterone (T) and dihydrotestosterone. Androgens play an integral role in the organizational and ontogenic processes involved in sexual differentiation and male-sexual development during embryogenesis (Swerdlow et al. 1992; Chang et al. 2002), although the AR remains widely expressed in a range of tissues in both male and female adults (Ruizeveld de Winter et al. 1991). Among other functions, AR-mediated gene transcription is integral to skeletal (Kenny and Raisz 2003), muscular (Dillon et al. 2010), and nerve cell (Arnold and Breedlove 1985; Hammond et al. 2001) development and maintenance, and in the regulation of cognition and behavior (Cherrier and Craft 2003; Eisenegger et al. 2011).

Within exon 1 of the AR gene are two trinucleotide repeats (CAG and GGN) whose numbers vary both within and between human populations (Figure S1; Edwards et al. 1992; Ackerman et al. 2012). Independent studies support an inverse relationship between the transcriptional activity of the AR and polyglutamine repeat length (AR CAGn; Kazemi-Esfarjani et al. 1995; Choong et al. 1996; Buchanan et al. 2004), and at least 64 different diseases and phenotypes have been investigated in relation to AR CAGn (Table S1). Given the functionally diverse roles of the AR in human development, physiology, and behavior (Chang et al. 1995; Chang 2002), coupled with more general effects of polyglutamine repeat number on mutability and function (Kashi and King 2006), the spectrum of putative effects of polymorphism in AR CAGn is well founded. However, the factors responsible for shaping the frequency and distribution of AR CAGn, which varies in number between eight and thirty-seven repeats in healthy humans (Zitzmann et al. 2003; Lindstrom et al. 2010), and its potential role in unexplained heritability in those diseases have received little attention (but see Hannan 2010). Here, we suggest AR CAGn as a promising, paradigmatic candidate helping explain missing heritability in human disease mediated by androgen levels, and we propose and evaluate four potential models of selection acting on the AR CAG locus to maintain intermediate repeat frequency in human populations: balancing selection, sexual conflict, accumulation-selection, and antagonistic pleiotropy.

First, balancing selection contributes to human illness when deleterious alleles are maintained in the population through heterozygote advantage or when disease risk results from extremes of bidirectional variation in gene expression (review in Crespi 2010). To the extent that the diseases described here are related to variations in CAGn

and the transcriptional activity in the AR, balancing selection should modulate the distribution of AR CAGn within and among populations. Under these conditions, selection will be strongest at the genetic and phenotypic extremes of AR CAGn, and the accompanying fitness costs and benefits should vary with the environment in which the AR is expressed (i.e. genetically, physiologically, ecologically).

Second, sexual conflict may be a selective force acting on AR CAGn, specifically if fitness optima for AR responsiveness and polyglutamine repeat number differ between men and women owing to differences in the benefits or disease-related costs of repeat number. A model for sexual selection, antagonistic pleiotropy, and sexual conflict at the AR CAG locus has been proposed (Summers and Crespi 2008), whereby an early advantage to male fertility and reproduction arising from shorter repeat number entails increased risk of cancer for both sexes. While in men the fertility benefits of short repeat number may compensate for the costs of greater prostate cancer risk, the absence of such benefits in women may lead to genomic conflict over repeat number between the sexes (Summers and Crespi 2008). Still, to our knowledge, sex differential fitness costs and the potential for sexual conflict accompanying the remaining 60+ diseases and phenotypes putatively linked to AR CAGn (Table S1) have not yet been proposed. The extent of sexual conflict and the magnitude of the costs arising from conflict over AR CAGn will, like balancing selection, be dictated by the contribution of repeat number to disease susceptibility, etiology, and their impacts on survival and reproduction for each sex.

Third, the expansion of tandem repeats is a common feature in the human genomic landscape, and an important force shaping the frequency of AR CAGn alleles in human populations (Rubinsztein et al. 1995; Gatchel and Zoghbi 2005). Although the mechanisms of accumulation of tandem repeats are becoming clear, the reasons behind the tendency toward expansion and repeat fidelity in humans (compared to other mammalian and primate lineages) are not well understood (Vowles and Amos 2006; Kelkar et al. 2008). Tandem-repeat expansions in protein coding regions are remarkably widespread, even though expansion is frequently associated with disease-related health effects (Buschiazzi and Gemmell 2006). If disease-associated selection is biased against phenotypes arising from longer than average AR CAG repeat number, then in a modified mutation-selection model, selection would primarily oppose the pressures of expansion of repeat number. Repeat number variation may then represent a fluctuating equilibrium between the strength of selection against longer AR CAGn and the propensity toward repeat accumulation, the rate of which is influenced by a collection of factors and which includes, intriguingly, the number and sequence

purity of repeats already present at that locus (e.g. Fig. 2 in Buschiazio and Gemmell 2006).

Finally, the nature and magnitude of phenotypic effects and disease susceptibilities putatively associated with AR CAGn should vary within an individual's lifetime, suggesting the potential for antagonistic pleiotropic effects in any or all of the above models of selection. Traits linked to AR CAGn which favor mating success and fertility early in life may be in conflict with disease-associated costs later on (e.g. Summers and Crespi 2008; Carter and Nguyen 2011); these costs may appear exaggerated in contemporary societies owing to increases in modal life span (Gurven and Kaplan 2007).

To evaluate these four non-exclusive hypotheses for explaining the distribution and variability in AR CAGn within and among human populations, we review eight of the best-studied diseases from a range of phenotypic classes putatively associated with AR repeat length, evaluate the evidence for and against associations between repeat length and disease, and appraise the relative strength and direction of selection for each disease. We also consider the health costs accompanying longer or shorter extremes in repeat number, the potential for sexual conflict arising from sex differences in the costs of, and the possible effects of antagonistic pleiotropy on, AR CAGn-associated disease. We conclude by discussing the accumulation of AR CAG repeat number in the human lineage in the context of human evolution and human disease susceptibility, and the potential role of AR CAGn as a component of the missing heritability of diseases linked to circulating androgen levels.

## Methods

### *Literature search*

We obtained data on the role of AR CAGn in human health and disease, using three online databases and one comprehensive review (Rajender et al. 2007) to compile a list of phenotypes and diseases with published, putative associations with AR CAG repeat length (Table S1). Database sources were the following: the AR Mutations Database (Gottlieb et al. 2004; ARDB; <http://androgendb.mcgill.ca/>, accessed 28 May, 2011), the online mendelian inheritance in man (OMIM) database (MIM ID \*313700, accessed 30 May, 2011), and the genetic association database (GAD; <http://geneticassociationdb.nih.gov>, accessed 21 May, 2011). All phenotypes associated with polymorphisms in AR CAGn were initially included, but mutations or polymorphisms at other loci in the AR (e.g. AR GGNn) were discarded. Using phenotypes and references obtained from these databases, subsequent searches focused on disease phenotypes and risk associations, and studies investigating the molecular structure and function of the AR were included only for interpreting mechanisms of pathology or

their role in selection on AR CAGn frequency in human populations. Additional literature was collected using Web of Science, Google Scholar, and PubMed databases using combinations of the search terms: 'AR' or 'AR', 'CAG', 'CAG repeat' or 'Polyglutamine repeat', and individual disease names (e.g. 'Prostate Cancer').

### *Disease and study selection*

To explore how susceptibility to disease risks associated with AR CAGn may shape the frequency of this polymorphism in humans, eight diseases were evaluated as follows: four putatively linked to longer repeat length, four to shorter repeat length. Necessary disease inclusion criteria were that the disease be referred to in at least three of the four main sources (the three databases and review) mentioned above and that it be the subject of at least one large-scale (500+ participants) study or meta-analysis examining its relationship to AR CAGn. Diseases that matched these criteria and were the focus of  $\geq 10$  independent studies were considered, as was colorectal cancer, for which we deemed the relationship between this disease and AR CAGn to be well supported by two particularly large (>3500 subjects for colon cancer; >1800 subjects for rectal cancer) case-control studies (Slattery et al. 2005). We were left with 12 remaining disease susceptibilities that we considered of equal interest but beyond the scope of the current study; therefore, we included eight diseases, which well represent the diversity of androgen function and disease class phenotypes described in Table S1. Of the remaining diseases not included but of interest are endometrial cancer, ovarian cancer, polycystic ovarian syndrome, and muscle mass/obesity and type II diabetes/metabolic syndrome. We consider the eight diseases ultimately chosen to evaluate hypotheses about the models of selection on AR CAGn illustrative and characteristic of disease relationships with the AR, and potential focal points for further research on the role of AR CAGn in missing heritability and on the models of selection proposed.

### *Strength of selection*

Although fitness costs of disease under ancestral conditions may be only loosely associated with clinical severity in contemporary human populations (Di Rienzo 2006), we evaluated the capacity for a disease to act as a selective agent based on three criteria, all of which would be reasonably expected to affect lifetime reproductive success. These were as follows: the frequency of disease occurrence; the average age of onset of the disease; and the effect of the disease on survival, fertility, and reproduction. Common diseases with profound reproductive or survivability effects, and with earlier onset in contemporary human populations, were predicted to have the greatest capacity to exert direct selective effects on AR CAGn. However, negative selection

against susceptibility alleles associated with typically 'late-onset' diseases is likely greater than previously believed, as a result of marked variability in age at onset and a number of factors other than direct selection (e.g. effects on survival and reproduction of kin; Pavard and Metcalf 2007). Additionally, extensive allelic variability and the potential for rapid and reversible changes in tandem-repeat purity and number (Kashi et al. 1997) make exaggerated selection in short timescales a distinct possibility at the AR CAGn locus. While the prevalence of a number of the diseases discussed (e.g. osteoporosis, cardiac diseases, some types of cancer) in past populations has been difficult to estimate, data emerging from paleoepidemiological studies support the presence of these diseases in human history, although contributing environmental factors in those populations may have differed (Zimmerman 1993; Mays 1998; Faltas 2010). Factors with the capacity to modulate the severity or progression of a disease, such as diet or lifestyle, may play an important role in disease-related costs and are described further in Table 1 and Table S1.

## Results

### *Disease risk and longer AR CAGn*

#### *Spinal bulbar muscular atrophy*

Spinal bulbar muscular atrophy (SBMA) shows an unequivocal relationship with AR CAGn. Patients with SBMA invariably have longer repeat number than is observed in the general population, typically between 38 and 62 repeats (La Spada et al. 1991; Amato et al. 1993; Brooks and Fischbeck 1995). Symptoms include late-onset muscular weakness and atrophy, frequently accompanied by androgen insensitivity and hypogonadism (Dejager et al. 2002; Palazzolo et al. 2008), believed to be a result of AR protein aggregation resulting in apoptosis of affected cells (Grierson et al. 1999; Ellerby et al. 2002; Vismara et al. 2009). Women may act as carriers of higher repeat number, experiencing mild if any symptoms of the disease, and toxicity appears to remain low even among women homozygous for high numbers of repeats (Mariotti et al. 2000; Greenland et al. 2004; Katsuno et al. 2010). Disease onset is typically later in life (30–60 years of age), although longer AR CAGn is predictive of earlier disease onset, which is often preceded by less severe symptoms including muscle fatigue and cramping (Atsuta et al. 2006). Despite the relatively rare nature of this disease (roughly 1/40 000 men), repeat numbers in the SBMA range bear formidable negative health effects. Risk of aspiration pneumonia (the most common cause of death in SBMA patients; Katsuno et al. 2010), muscle degeneration, and loss of mobility would likely have been strongly selected against under most ancestral conditions. The rescuing effect of a second less

toxic allele with shorter repeat number in women, accompanied by the reduction in symptoms in homozygotes, means that the costs of high repeat number associated with this disease differ for men and women, conditions which would contribute to sexual conflict over AR CAG repeat length. There are no known fitness benefits associated with repeat numbers in the SBMA range.

#### *Male infertility*

The essential role of androgens in male virility and spermatogenesis (Collins and Chang 2002), and the association between SBMA and infertility (e.g. Arbizu et al. 1983) have led to a number of investigations into the differences in AR sensitivity arising from variations in CAGn and idiopathic male infertility (e.g. Table 1 and references therein). Repeat number at the AR CAGn among infertile patients has been variously found to be longer (e.g. Tut et al. 1997; Dowsing et al. 1999; Lim et al. 2000; von Eckardstein et al. 2001; Davis-Dao et al. 2007; Nenonen et al. 2010), shorter (Komori et al. 1999; Nenonen et al. 2010), or not significantly different (e.g. Dadze et al. 2000; Meyts et al. 2002; Thangaraj et al. 2002; Yong et al. 2003) from those of controls, with ethnic or population level differences potentially confounding the results. A large-scale meta-analysis provides good support for a link between longer AR CAGn and infertility (Davis-Dao et al. 2007), but the average contribution of each additional repeat to infertility has not been empirically demonstrated. Still, the actual difference between patients and controls is likely to underestimate the effect of repeat number, given the fact that an unknown proportion of patients with repeat numbers in the shorter, 'normal' range will be infertile because of other unknown causes (Davis-Dao et al. 2007). A non-linear relationship between infertility and AR CAGn has also been proposed, such that men with longer *or* shorter AR CAGn than the median (22–23 repeats) are at a 20% increased risk of infertility (Nenonen et al. 2010). If this pattern is true, then stabilizing selection around intermediate repeat frequency could arise from male infertility alone. Given the relative commonness of male infertility (estimated to be approximately 7%; Meacham et al. 2007), and the age-independent and thus potentially profound effects of AR CAGn on male lifetime reproductive success and fitness, longer (and possibly shorter) AR CAGn should be under strong selection from infertility in men.

#### *Breast cancer*

Endogenous androgen steroid levels have been recognized as modulating factors associated with breast cancer (BC) (Adams 1998; Ferro et al. 2002; Kaaks et al. 2005), and between 60% and 70% of BCs express the AR as well as androgen-dependent proteins (e.g. PSA and GCDFFP-15; Díaz-Chico et al. 2007). *In vitro* studies support a

**Table 1.** Susceptibility for eight diseases putatively associated with AR exon 1 polyglutamine repeat number (AR CAGn).

Disease putatively associated with AR CAG	Disease prevalence	Age of onset	Susceptible sex	Health effects	Risk factors	References
Longer						
Spinal bulbar muscular atrophy	Rare	Mid-late reproductive	Males	Survival and reproduction	Androgens levels, pneumonia	1–8
Infertility	Common	Early reproductive	Males	Reproduction	Ethnicity, SHBG, epigenetics	9–20
Breast cancer	Very common	Late reproductive	Predominantly females	Survival	Other genes, hormone therapy, family history, parity	21–39
Osteoporosis, decrease BMD	Very common	Late reproductive	Both sexes	Survival	Age, gender, SHBG	40–45
Shorter						
Prostate cancer	Very common	Mid-Late reproductive	Males	Survival and reproduction	Androgen levels, other genes	46–73
Cardiac diseases	Very common	Mid reproductive	Both sexes	Survival	Diet, lifestyle	74–80
Colorectal cancer	Common	Mid-late reproductive	Both sexes*	Survival	Diet, gender, other genes/hormones	81–84
Cognition and behaviour disorders	Common	Pre-reproductive	Males	Survival?	Age, gender, environment	85–93

AR, androgen receptor; BMD, bone mass density; SHBG, sex hormone-binding globulin.

Diseases grouped by proposed direction of association, and prevalence is based on data from contemporary American society ( $\leq 0.0001/100$  people = rare,  $0.01–0.1/100$  = common, and  $\geq 0.1/100$  = very common. Age of onset of disease pertains to the age at which health effects most likely become evident relative to reproductive age. Effects of disease on health and the sex most susceptible are described, and possible risk factors are provided based on the references provided.

\*Direction of association with colorectal cancer may differ for each sex, see text.

References: 1. Amato et al. (1993). 2. Kazemi-Esfarjani et al. (1995). 3. Mariotti et al. (2000). 4. DeJager et al. (2002). 5. Greenland et al. (2004). 6. Atsuta et al. (2006). 7. Katsuno et al. (2010). 8. La Spada et al. (1991). 9. Asatiani et al. (2003). 10. Tut et al. (1997). 11. Thangaraj et al. (2002). 12. Dowsing et al. (1999). 13. Davis-Dao et al. (2007). 14. Komori et al. (1999). 15. Lim et al. (2000). 16. von Eckardstein et al. (2001). 17. Meyts et al. (2002). 18. Yong et al. (2003). 19. Lazaros et al. (2008). 20. Verhoeven et al. (2010). 21. Comings et al. (2003). 22. Díaz-Chico et al. (2007). 23. Elhaji et al. (2001). 24. Ferro et al. (2002). 25. Giguère et al. (2001). 26. Goode et al. (2002). 27. Haiman et al. (2002). 28. Haiman et al. (2003). 29. Hao et al. (2010). 30. Iobagiu et al. (2005). 31. Kaaks et al. (2005). 32. Lillie et al. (2004). 33. Maclean et al. (2004). 34. López-Otín and Diamandis (1998). 35. Rebbeck et al. (1999). 36. Spurdle et al. (2005). 37. Suter et al. (2003). 38. Wang et al. (2005). 39. Yu et al. (2000). 40. Gennari et al. (2007). 41. Guadalupe-Grau et al. (2010). 42. Langdahl et al. (2003). 43. Limer et al. (2009). 44. Tofteng et al. (2003). 45. Zitzmann et al. (2001a, b). 46. Coetzee and Ross (1994). 47. Monroe et al. (1995). 48. Tilley et al. (1996). 49. Giovannucci et al. (1997). 50. Stanford et al. (1997). 51. Pettaway (1999). 52. Beilin et al. (2001). 53. Kittles et al. (2001). 54. Latil et al. (2001). 55. Panz et al. (2001). 56. Shibata et al. (2001). 57. Azzouzi et al. (2002). 58. Chang et al. (2002). 59. Mononen et al. (2002). 60. Schatzl et al. (2002). 61. dos Santos et al. (2003). 62. Buchanan et al. (2004). 63. Gilligan et al. (2004). 64. Giwercman et al. (2004). 65. Zeegers et al. (2004). 66. Alvarado et al. (2005). 67. Forrest et al. (2005). 68. Freedman et al. (2005). 69. Platz et al. (2005). 70. Andersson et al. (2006). 71. Summers and Crespi (2008). 72. Lindstrom et al. (2010). 73. Kumar et al. (2011). 74. Hersberger et al. (2005). 75. Pausova et al. (2010). 76. Zitzmann et al. (2001a). 77. Alevizaki et al. (2003). 78. Lind et al. (2008). 79. Page et al. (2006). 80. Rexrode et al. (2008). 81. Slattey et al. (2005). 82. Gillessen et al. (2010). 83. Ferro et al. (2000). 84. Di Fabio et al. (2009). 85. Cheng et al. (2006). 86. Rajender et al. (2008). 87. Jönsson et al. (2001). 88. Westberg et al. (2009). 89. Manuck et al. (2010). 90. Fondon et al. (2008). 91. Yaffe et al. (2003). 92. Seidman et al. (2001). 93. Colangelo et al. (2007).

protective effect of androgens on hormone-independent BC-cell lineage proliferation (Di Monaco et al. 1995; Gatto et al. 1996; Szelei et al. 1997), and low premenopausal androgen levels have been associated with susceptibility to this disease (Adams 1998; Wang et al. 2000). Longer AR CAGn has been correlated with BC risk (Giguère et al. 2001; Haiman et al. 2002; Liede et al. 2003; Suter et al. 2003; but see Hao et al. 2010), younger age at onset (Rebbeck et al. 1999) and tumor aggressiveness (Yu et al. 2000) and grade (Elhaji et al. 2001; Maclean et al. 2004). Because the effects of AR CAGn on BC risk interact with a number

of other factors, including nutrition (Kaaks et al. 2005), hormone treatment (Suter et al. 2003; Lillie et al. 2004), polymorphisms at other loci (Rebbeck et al. 1999; Suter et al. 2003), family history (Rebbeck et al. 1999; Haiman et al. 2002), and ethnicity (with Caucasians showing the highest risk and the longest average repeat number; Altek-ruse et al. 2007), it is difficult to infer the magnitude of selection against longer AR CAGn resulting from BC. Still, the high prevalence (lifetime risk approximately 12%; Altek-ruse et al. 2007) and variance in age at onset of the disease (Pavard and Metcalf 2007), and the importance of

aloparental care from post-reproductive women suggest that the fitness costs of susceptibility to BC in ancestral environments could have been significant and could have contributed to the distribution in AR CAGn repeat number we see in contemporary populations. Although a positive relationship between repeat number and the occurrence and grade of BC has also been observed in men (Maclean et al. 2004), selection owing to disease susceptibility in this sex is unlikely to contribute to AR CAGn owing to the very low occurrence of this disease in men (approximately 0.13%; Altekruuse et al. 2007).

#### *Osteoporosis and bone mass density*

The general role of androgens in bone metabolism, loss of bone mass in cases of hypogonadism, and reduction in bone turnover with testosterone treatment all lead to predictions for a decrease in bone mass density (BMD) and increase in osteoporosis (femoral neck BMD  $<0.56 \text{ g/cm}^2$ ) with longer AR CAGn (Zitzmann et al. 2001b; Zitzmann 2009). A relationship between polyglutamine repeat lengths among premenopausal (but not postmenopausal) women with lower BMD has been shown (Yamada et al. 2004), as has a relationship between BMD and AR CAGn on the longer of the two alleles in women, with significantly longer AR CAGn among female patients compared to controls (Langdahl et al. 2003). In healthy men, AR CAGn is a negative predictor of BMD, and the effect of age on bone loss is greater in subjects with longer repeat length (22–31 repeats) compared to those with shorter repeat lengths (14–21 repeats; Zitzmann et al. 2001b). Several studies have reported the opposite, however, finding both a negative (Limer et al. 2009) or both positive and negative relationship between BMD, bone mineral content (BMC), and AR CAGn, but only in conjunction with AR GGNn (another amino acid repeat polymorphisms in the AR; Guadalupe-Grau et al. 2010), or under the modulation of steroid hormone binding globulin (SHBG; Tofteng et al. 2003).

Independent of other factors, longer AR CAGn may contribute to BMD and risk of osteoporosis, and osteoporosis prevalence differs considerably between ethnic groups as well as between the sexes (Melton 2001). The Third National Health and Nutrition Examination Study recorded 20% of postmenopausal Caucasian women as osteoporotic, compared to only 5% in African American women. By contrast, only 4% of Caucasian men over 50 years of age were defined as osteoporotic, compared to 2% of African American men (Looker et al. 1997). These differences in prevalence mirror ethnic AR CAGn frequencies, with African-Americans possessing shorter repeat lengths (Figure S1), which supports a role for longer AR CAGn in osteoporosis risk.

Although often considered to be a 'disease of civilization' (Karasik 2008), a body of archeological evidence suggests

that osteoporosis may be more common among human history than once believed (Mays 1998 and references therein; Poulsen et al. 2001; Cho and Stout 2011; but see Agarwal and Grynepas 1996). In fact, poor nutrition and extended lactation may have contributed to even higher rates of osteoporosis or earlier onset of age-related losses in BMD in some regions, despite the positive effect of greater physical activity in those populations (Turner-Walker et al. 2001). While hip fractures in past populations appear to be rare (possibly owing to the shorter overall lifespan or selection against low BMD earlier on in life; Mays 1998), other bone-density related fractures might have been more common and accompanied by poorer functional outcomes (Mays 2006). Also, the relatively porous nature of human vertebral bone (compared to our closest living relatives) makes it particularly susceptible to fracture with even modest losses in BMD (Cotter et al. 2011), and the health consequences of these types of acute trauma and debilitation no doubt exceeded those of current mechanized societies. The poor functional outcomes accompanying even relatively 'minor' fractures associated with losses in BMD, potentially earlier age at onset from deficient childhood nutrition or extended lactation, and the importance of alloparental care from post-reproductive women (Hawkes 2003) suggest that osteoporosis associated with longer AR CAGn has the capacity to act as a selective force AR CAGn repeat number frequency. As for a modulating effect of AR CAGn in the missing heritability of osteoporosis (Karasik 2011), at least one intriguing study has identified heritable components of canine skeletal morphology with glutamine repeat number at another locus: heritability not detected using traditional SNP-based approaches (Fondon and Garner 2004), pointing to the importance of tandem repeats in bone structure and function.

#### **Disease risk and shorter AR CAGn**

##### *Prostate cancer*

Evidence strongly supports a role for androgens and AR CAGn in prostate cancer (PC) risk and progression. Castrated or hypogonadic men (including men suffering SBMA) rarely develop prostate cancer, and chronic exogenous androgen administration in rats can induce the disease (Henderson and Feigelson 2000; Hsing et al. 2008). Prostate cancer progression is sensitive to androgen deprivation (a common therapy), and a crucial stage in disease progression is the evolution of androgen-independent cancer cell lineages (Tilley et al. 1996; Henderson and Feigelson 2000; Grönberg 2003; Ross et al. 2005). Inheritance patterns also support a contribution of X-linked genes (which include the AR), with brothers of individuals succumbing to the disease showing greater risk of developing prostate cancer themselves than sons of individuals with the disease

(Monroe et al. 1995). Shorter AR CAGn repeat number has been associated with disease risk (Irvine et al. 1995; Panz et al. 2001; Andersson et al. 2006; but see Forrest et al. 2005), age at onset/diagnosis (Beilin et al. 2001; Latil et al. 2001; dos Santos et al. 2003), and prostate cancer grade, stage, metastasis and fatality resulting from the disease (Giovannucci et al. 1997; Hakimi et al. 1997; Shibata et al. 2001). Shortening of AR CAGn is also commonly associated with PC progression (Alvarado et al. 2005), and the AR itself has become a key target for therapeutic research (Berger et al. 2011). Additionally, ethnic differences in AR CAGn (like BC and osteoporosis) mirror racial susceptibility to prostate cancer, with men of African origin displaying the shortest CAGn and the highest incidence of prostate cancer, with the opposite being true of Asians (Figure S1; Edwards et al. 1992; Coetzee and Ross 1994; Pettaway 1999; Kittles et al. 2001; Panz et al. 2001). A 2004 meta-analysis confirmed a significant difference between cases and controls, although the differences do appear to be modest (<1 repeat difference between patients and controls; Zeegers et al. 2004).

In contrast to the patterns described above, the largest study to examine prostate cancer and AR CAGn (Lindstrom et al. 2010) did not detect any relationship between these two traits, nor did several other large-scale studies multi-ethnic cohort study (Mononen et al. 2002; Freedman et al. 2005). One explanation for the difference between earlier and more recent studies has been diagnostic technologies for identifying prostate cancer in its early stages. The widespread use of prostate-specific antigen (PSA) beginning in the early nineties has shifted detection to less aggressive manifestations and earlier stages of the disease, which do not always progress to advanced stages of the disease, or may do so much more slowly (Platz et al. 2005).

There is no doubt that factors other than AR CAGn are important in prostate cancer risk and disease etiology (e.g. AR GGNn; Hakimi et al. 1997; Stanford et al. 1997). Still, the relatively robust connection between AR CAGn and disease susceptibility and the high occurrence of the disease in men (approximately 16% lifetime risk; Altekruse et al. 2007), coupled with successful reproduction to relatively old ages in human men, mean that even modest increases in the susceptibility to prostate cancer could engender considerable fitness costs and be an important contributor to AR CAGn polymorphism frequencies in contemporary human populations. Antagonistic pleiotropy arising from selection for prostate-expressed genes and their implication on prostate cancer have been formally proposed (Summers and Crespi 2008), but the promising role of AR CAGn in explaining missing heritability by modulating prostate cancer risk and progression via interactions with circulating androgen levels may deserve increased attention.

#### *Cardiac diseases and atherosclerosis*

Testosterone has been suggested as a contributing factor to the higher rates of atherosclerosis and cardiac disease among men than among women and may affect a range of risk factors contributing to susceptibility to these diseases (Hanke et al. 2001; Weidemann and Hanke 2002; Wu and Eckardstein 2003). Higher sympathetic vasomotor tone, blood pressure, and intra-abdominal fat, all factors known to contribute to cardiac disease rate, were found among French Canadian boys (aged 12–18 years) with shorter AR CAGn when compared to boys with longer CAGn (Pausova et al. 2010). Similarly, European men with shorter repeat length associated positively with obesity and stenosis of the arteries (Alevizaki et al. 2003) and negatively with high-density lipoprotein cholesterol levels (Hersberger et al. 2005) and flow-mediated dilatation (Zitzmann et al. 2001a). Shorter AR CAGn also correlated with higher LDL-cholesterol in Spanish women (Rodríguez-González et al. 2009), and ventricular hypertrophy in men (Lind et al. 2008), suggesting an adverse affect of short AR CAGn for cardiac and atherosclerotic diseases. However, Page et al. (2006) failed to detect any relationship between body mass, heart disease, and HDL, even over a 15-year follow-up period, nor did another study on American women (Rexrode et al. 2008). Several protective parameters for cardiac disease, including lower body fat mass and insulin levels, have also been associated with shorter AR CAGn (Zitzmann et al. 2003; but see Gustafson et al. 2003), making interpreting the role of AR CAGn in cardiac diseases difficult. With all European studies supporting an effect of AR CAGn on at least some cardiac disease risk factors (Zitzmann et al. 2001a, 2003; Alevizaki et al. 2003; Hersberger et al. 2005), and one small and one large American study finding no effect (Page et al. 2006; Rexrode et al. 2008), potential population-level differences merit further consideration. Although extremely common in contemporary populations (nearly one-half and one-third lifetime disease risk by the age of 40 for American men and women, respectively), heart disease and atherosclerosis are largely modulated by diet and lifestyle, which also explain some of the missing heritability and population level differences. The mismatch between current and ancestral diet and lifestyle is a major contributor to cardiac and vascular diseases, making selection from these disease susceptibilities unlikely to contribute significantly to AR CAGn frequencies in ancestral human populations. Still, cardiac diseases and atherosclerosis appear common among certain ancient Chinese and Egyptian social classes as well as among some ethnic groups (e.g. the Inuit) and have been identified in 5000 +-year-old mummified remains (Murphy et al. 2003), suggesting that these diseases should not be dismissed outright as 'diseases of civilization' (David et al. 2010; Allam et al. 2011).

### *Colon and rectal cancer*

Androgens regulate growth and differentiation in colon and rectal tissue, and there is support for an association between low testosterone levels and colon cancer in laboratory animals (Xiao et al. 2007; Gu et al. 2009). Studies in animals suggest a protective role of androgens in colon tumorigenesis (Ferro et al. 2002), and prostate cancer patients undergoing long-term androgen deprivation therapy were at a greater risk of developing colorectal cancer (Gillessen et al. 2010). While longer AR CAGn corresponds to the risk of colon cancer in men, longer repeat length appears to be protective in women (Slattery et al. 2005). Women with long repeat number in another polymorphic gene, the  $\beta$ -estrogen receptor, in addition to long AR CAGn, also had a higher risk of disease than women with shorter repeat numbers for both alleles (Slattery et al. 2005). It is worth noting that African American men show lower lifetime risk for the disease compared to Caucasian women and that this relationship is reversed for African American and Caucasian women (Altekruse et al. 2007), corresponding to shorter and longer AR CAGn, respectively. These findings suggest that susceptibility to colon cancer is associated with both longer *and* shorter AR CAGn, *depending on gender*, which correspond to ethnic level differences in mean AR CAG repeat number (Figure S1). The protective effects of shorter AR CAGn in men versus the increased risk of disease in women means that optimal repeat number for AR CAGn with respect to colon cancer alone may differ for each sex. The relatively high frequency of colon cancer occurrence in contemporary populations (approximately 5% lifetime risk in Americans; Altekruse et al. 2007) implies that, in addition to being a strong candidate disease for sexual conflict over optimal AR CAGn, susceptibility to colorectal cancer could contribute to variation in AR CAGn among human populations. Colorectal cancer is another case where the sources of heritable susceptibility remains unclear (Lascorz et al. 2010), and where modulation of the effects of androgens on AR CAGn may be an important factor in susceptibility to this disease (Slattery 2006).

### *Cognitive and behavioral disorders*

Although not mutually exclusive, at least three psychological traits show some support for a role of the AR CAGn: (i) aggression, violence, and criminal activity; (ii) cognitive functioning and general intelligence; and (iii) depression. Other socio-behavioral traits, including social and sexual behavior, have been linked to AR CAGn as well as to other repeat polymorphisms in non-human mammals (Hammock and Young 2005; Fondon et al. 2008).

Consistent, though, non-significant trends toward aggressive and dominant behavior were first associated with shorter AR CAGn (Jönsson et al. 2001) and subsequently

spurred interest in this area. AR CAGn has since been significantly correlated with both aggression and risky behaviors among boys (Vermeersch et al. 2010) and inmates (Aluja et al. 2011) with shorter repeat length, and a study on Taiwanese criminals found that a significantly larger proportion of violent criminals carried short alleles (<17) than did controls (Cheng et al. 2006). Among Indian men, Rajender et al. (2008) observed significantly shorter repeat length among murderers and rapists than controls, and convicts of both murder and rape had significantly shorter repeat length than criminals who committed murder or rape, but not both. More recently, a study looking at amygdala reactivity among Caucasian American men using fMRI found higher reactivity to facial displays of negative affect among men with shorter AR CAGn (Manuck et al. 2010), although a questionnaire-based study of college students did not detect personality differences with respect to polymorphisms at the AR (Hurd et al. 2011). AR CAGn has also been linked to dominance and status, variables associated with intrasexual competition for mates (Simmons and Roney 2011), and the response of men to potential mates (Roney et al. 2010).

Cognitive functioning in elderly, community-dwelling men was inversely correlated with AR CAGn for all three of the cognitive tests originally examined by Yaffe et al. (2003), but no such relationship has been observed in middle-aged and aging European men (Lee et al. 2010) nor in a sample of healthy Chinese volunteers of varying ages (Kovacs et al. 2009). Manning (2007) has suggested a hypothesis for an effect of AR CAGn on neuronal transmission rate and general intelligence (*g*), which proposes that *g* increases with repeat length observed along the mammalian lineage is constrained in humans by the negative effects of SBMA and impaired sperm production. This 'gain in function' for general intelligence with longer AR CAGn has little empirical support, but given the highly expressed nature of the AR in areas of the brain associated with visual and verbal memory (Cherrier and Craft 2003), cognitive functioning (Kovacs et al. 2009), and neurological development and neuroprotection (Hammond et al. 2001; Perrin et al. 2008), this line of thinking should not be dismissed and could provide insights into cognitive and disease susceptibility differences observed between the sexes.

In a recent study of adolescent boys, free testosterone levels have been associated with aggressive and non-aggressive risk-taking behaviors, self-esteem, and inversely correlated to depressive symptoms, but these relationships were highly dependent on AR CAGn (Vermeersch et al. 2010). A significant interaction between total testosterone and depressive symptoms has also been observed in men with the short, but not men with long, AR CAGn (Seidman et al. 2001), with similar findings in black (who tend to have shorter AR CAGn lengths), but not white, American



men (Colangelo et al. 2007). Collectively, these data suggest that short AR CAGn may contribute to the risk of depression, particularly when testosterone levels are low.

The data described above suggest that men with shorter AR CAGn are more generally intelligent, violent, and aggressive, and less inclined toward depression, but that this relationship may be largely dependant on circulating testosterone levels. The modulating effect of AR CAGn is particularly intriguing, given the reciprocal relationship between dominance and testosterone; testosterone levels not only affect, but are also affected by, dominant social behavior (Mazur and Booth 1998). As a result, the psychological responses to competitive or goal-directed behavior may be mediated by testosterone, but the psychological costs and benefits of high or low testosterone levels may be greater for men with short AR CAGn. Based on these data, it is also interesting to consider a role of sexual selection for cognitive and behavioral traits of testosterone, which may encompass both mood-oriented and cognitive effects modulated by AR CAGn and AR transcriptional sensitivity, although the typically small, multigenic, and environmentally influenced the effect of repeat variation like AR CAGn make it an ongoing challenge to detect their effects (Fondon et al. 2008).

#### Disease risk and the accumulation of tandem repeats in the human genome

There is a well-documented increase in coding single tandem repeats, like the AR CAGn, which accompanies the evolution of mammals, primates, and humans (Rubinsztein et al. 1995; Andrés et al. 2004; Vowles and Amos 2006; Kehrer-Sawatzki and Cooper 2007; Mularoni et al. 2010). For the AR CAGn, repeat accumulation along the evolutionary trajectory of mammals is close to exponential (Choong et al. 1998); accumulations that persist in humans, even though long repeat number is frequently associated with pathology, including neurodegenerative disease such as SBMA (Gatchel and Zoghbi 2005). The expansion rate for trinucleotide repeats like the AR CAGn is influenced by sequence repeat purity (Buschiazzo and Gemmell 2006), and while there is evidence for selection acting on trinucleotide repeats (Hancock et al. 2001; Haerty and Golding 2010a), polyglutamine repeats in the human genome are more common and retain a higher degree of sequence fidelity than predicted by neutral expectations alone (Gemayel et al. 2010; Haerty and Golding 2010b; Mularoni et al. 2010). Thus, even though selection should favor point mutations modifying the repeat sequence purity, which would decrease the propensity toward further disease-causing repeat accumulation, repeat sequence in disease-associated coding regions is more highly conserved than in nearby non-coding regions. More

intriguing is the fact that tandem repeats with conserved sequences tend to be concentrated on certain amino acids and in certain categories of genes (e.g. glutamine repeats and transcription factors, as for the AR CAGn; Hancock et al. 2001; Gemayel et al. 2010) and that repeat sequence purity, and hence the propensity for repeat accumulation, may differ among human populations (Sobczak and Krzyzosiak 2004), consistent with some adaptive, functional role for repeat accumulation in the human genome.

#### Discussion

Balancing selection, sexual conflict, and accumulation-selection are empirically supported as forces with the potential to contribute to AR CAGn frequencies among human populations. These conclusions are based on the causal links between AR CAGn and disease susceptibility, the prevalence the diseases in question, their age at onset, their effects on survival, fertility and reproduction, as well as sex-dependent differences in the putative fitness costs associated with each disease. Of the eight disease susceptibilities evaluated, there is empirical support for a role of AR CAGn in disease susceptibility to five of these diseases: SBMA, infertility, and BC for longer AR CAGn, prostate cancer for shorter AR CAGn, and cancer of the colon for both long and short repeat length, depending on the sex of the carrier. The strong association between osteoporosis and AR CAGn is tempered by equivocal paleopathological data regarding prevalence in past populations (Ekenman et al. 1995; Agarwal and Grynepas 1996; Mays 1998; Karasik 2008), making it more difficult to infer the role of this disease on AR CAGn distribution. The diseases associated with later age at onset (i.e. prostate cancer, BC, osteoporosis; Table 1) could still be significant contributors to AR CAGn distributions owing to the generally underappreciated capacity for negative selection at late-onset disease susceptibility alleles. Variability in age at onset, the contribution of children born to women of 39 years and older in pre-industrial societies (e.g. 6–11% of lifetime reproductive success; Pavard and Metcalf 2007), and indirect contributions of late survival to fitness (e.g. grandmothering; Hawkes 2003; but see Kachel et al. 2011) imply that while antagonistic pleiotropy may be a less important factor to explain AR CAGn distributions than originally hypothesized (Summers and Crespi 2008; Carter and Nguyen 2011), the late-onset diseases discussed still have the capacity to exert effects on lifetime reproductive success. To the extent that repeat number in the AR contributes, either directly or indirectly, to susceptibility to these complex diseases, AR CAGn and comparable microsatellite loci are promising candidates for helping us to explain some of the missing heritability of disease risk not currently accounted for in traditional SNP-based GWA studies.

Compelling evidence for disease risk accompanying AR CAGn at both long and short AR CAGn implies that balancing selection is involved in CAGn number distributions in human populations. To address this hypothesis, the sum benefits, costs, and susceptibilities of long AR CAGn diseases must balance the sum benefits, costs and susceptibilities of short AR CAGn diseases, which should be reflected in the mean number of repeats in the population of interest. Balancing selection at the population level would maintain high levels of intermediate repeat number, with fewer individuals possessing high or low repeat number, one possible explanation for the 'bell-shaped' distribution for AR CAGn observed (Figure S1). The variance and mean repeat number in a population will then reflect susceptibilities for each disease and their fitness costs, which must therefore be interpreted in the physiological, geographical, and ecological context in which they have, and presumably are, evolving.

There is also support for sexual conflict over AR CAGn locus, given sex differences in susceptibility and disease type (Summers and Crespi 2008). For SBMA, infertility, and susceptibility to colon cancer, men should favor shorter AR CAGn, whereas women may favor shorter (BC) or longer (colon cancer) repeat lengths. In fact, if the AR CAGn affects susceptibility to colon cancer in men and women differently, then this disease alone could contribute to conflict over repeat number between the sexes. In men, the benefits of short repeat number may exceed the costs (i.e. from prostate cancer), and as there appear to be fewer benefits of short repeat number for women, our findings appear to support a hypothesis of sexual conflict over AR CAGn. The costs and benefits of short repeat length may also differ owing to other sex-dependant modulating factors, such as circulating androgen levels, which could reasonably explain part of the ethnic variation in AR CAGn (Figure S1).

The tendency toward polyglutamine repeat accumulation within mammalian, primate, and human lineages, and the well-documented disease risks accompanying longer AR CAGn (Choong et al. 1998; Gatchel and Zoghbi 2005; Kellkar et al. 2008; Mularoni et al. 2010) provide support for a model of accumulation-selection at this locus. The accumulation of repeats, driven by events such as slippage during replication (Buschiazzo and Gemmell 2006), could be offset in the AR by selection against longer repeat number from SBMA, infertility, and colon cancer and BC. This hypothesis provides an alternative, but not necessarily exclusive, explanation to balancing selection and sexual conflict when considering the distribution of repeat numbers among human populations, particularly if the propensity for repeat accumulation does indeed differ between populations (Sobczak and Krzyzosiak 2004). Our understanding of, and interest in, the functional role and evolutionary context of trinucleotide repeats continues to grow (Vismara et al. 2009; Castel et al. 2010; Haerty and Golding

2010a; Luo et al. 2012), and if the suggestion that accumulation of trinucleotide repeats like the AR CAGn is a non-neutral process, contributing to genetic variability for rapidly evolving traits (Birge et al. 2010), then trade-offs between adaptive trait variation and the costs of disease susceptibility may be pivotal in the proposed accumulation-selection model. If an 'equilibrium' of costs and benefits in repeat number exists, then AR CAGn distribution among populations becomes particularly interesting in the light of changing selective pressures and medical intervention in locus-associated diseases. With human intervention in some diseases potentially outpacing others, the outcome over many generations may be changes in median population repeat number and shifts in phenotypes and the susceptibility to other diseases associated with AR CAGn. Given the push toward detailed phenotypic data collection for large clinical cohorts and the implementation of novel evolutionary models to track phenotypic and disease-related changes (Stearns et al. 2010), genes like the AR CAGn may be useful targets in tracing effects of human-induced changes and gene-culture co-evolution.

## Conclusions

Models taking into account balancing selection, sexual conflict, antagonistic pleiotropy, and accumulation-selection will be instrumental to our understanding of disease susceptibility associated with repeat number at the AR CAGn and other loci. Testing these hypotheses requires accurate estimates of disease susceptibility and fitness costs (and benefits) of repeat length, or at least quantitative proxies for these metrics (Polanski et al. 1998), as well as studies of molecular-evolutionary forces affecting such loci within and among human populations. One of the great challenges is a quantitative measure of disease susceptibility and costs from AR CAGn, especially considering genetic correlations and interactions, gene-by-environment interactions, and thus myriad contributions to disease risk and clinical outcomes. Mismatch between the ancestral environments in which repeat number evolved and the current environment, and the differences in disease rates and costs that accompany this mismatch add yet another dimension to an already formidable task.

Still, the challenge of testing evolutionary hypotheses is not restricted to hypotheses of human health and disease (Gluckman et al. 2011), and the AR and CAGn provide an excellent system in which to explore these complex and elusive forms of selection in the human genome. The growing number of sequenced human genomes (The 1000 Genomes Project Consortium 2010; <http://www.1000genomes.org>), the recent construction of a large-scale, annotated database of expressed trinucleotide repeats like the AR CAGn (Luo et al. 2012), and large, long-term, multigenerational studies

(Stearns et al. 2010) may open up new possibilities for studying the evolutionary and functional context of accumulations at the AR CAGn, and for resolving some of the disease-associated consequences of their expansion (Haerty and Golding 2010a).

These kinds of advances may also provide powerful insights into the role of the AR CAGn in the missing heritability of complex diseases and phenotypes modulated by androgens. Tandem repeats such as the AR CAGn may be more informative than SNPs at the individual level owing to their functional role and greater standing genetic variation in human populations, but have been largely neglected in GWA studies owing to the statistical power and high-throughput assays required to incorporate them (Ku et al. 2010). Yet finer-scale analyses of the genetic architecture of the human genome, including tandem repeats like the AR CAGn, are becoming an increasingly important goal in the pursuit of missing heritability for complex phenotypes and disease (Eichler et al. 2010).

A more comprehensive picture of the heritability of human disease susceptibility must also account for interactions between genes and between genes and the environment (Eichler et al. 2010; Stearns et al. 2010). As an evolvable, dynamic, yet robust, interface between cellular responses and the physiological and ecological environment, the endocrine system and its receptors are ideally situated to mediate a wide range of disease susceptibilities and health-related effects. Incorporating tandem repeats, particularly those with known functional roles like those found in the AR CAGn, into the current GWA study framework may unveil genetic and environmental interactions confounding current efforts to explain disease risk and etiology (Hannan 2010; Ku et al. 2010). While understanding the mechanistic and functional consequences of polymorphisms in tandem-repeat number are vital, the evolutionary forces upon which that genetic and functional variation is superimposed are inextricable from phenotypic and disease-associated manifestations. Applied as a component of more comprehensive GWA study design or therapeutically in relation to conventional (e.g. androgen supplementation or ablation) or novel (e.g. targeting instable repeats) personalized disease treatments, tandem repeats like the AR CAGn hold great promise for the effective identification and treatment of disease. In each case, the costs and benefits of polymorphisms in tandem-repeat number variation are fundamentally embedded in their evolutionary legacies.

## Acknowledgements

This manuscript was greatly improved by the insightful comments of Tony D. Williams and two anonymous reviewers. C.P.R. and B.J.C thank the Natural Sciences and Engineering Council of Canada (NSERC) for the financial

support. C.P.R. was also supported by a fellowship from Simon Fraser University, Dean of Graduate Studies.

## Literature cited

- Ackerman, C. M., L. P. Lowe, H. Lee, M. G. Hayes, A. R. Dyer, B. E. Metzger, W. L. Lowe et al. 2012. Ethnic variation in allele distribution of the androgen receptor (AR) (CAG)<sub>n</sub> repeat. *Journal of Andrology* **33**:210–215.
- Adams, J. B. 1998. Adrenal androgens and human breast cancer: A new appraisal. *Breast Cancer Research and Treatment* **51**:183–188.
- Agarwal, S. C., and M. D. Grynbas. 1996. Bone quantity and quality in past populations. *The Anatomical Record* **246**:423–432.
- Allam A.H., R.C. Thompson, L.S. Wann, M.I. Miyamoto, A. el-Halim Nur el-Din, G.A. el-Maksoud, M. Al-Tohamy Soliman et al. 2011. Atherosclerosis in Ancient Egyptian Mummies: The Horus Study. *Journal of the American College of Cardiology: Cardiovascular Imaging* **4**:315:327.
- Alevizaki, M., A. T. Cimponeriu, M. Garofallaki, H.-L. Sarika, C. C. Alevizaki, C. Papamichael, G. Philippou et al. 2003. The androgen receptor gene CAG polymorphism is associated with the severity of coronary artery disease in men. *Clinical Endocrinology* **59**:749–755.
- Altekruse, S. F., C. L. Kosary, M. Krapcho, N. Neyman, R. Aminou, W. Waldron, J. Ruhl et al. (eds). SEER Cancer Statistics Review, 1975–2007. National Cancer Institute. Bethesda, MD, [http://seer.cancer.gov/csr/1975\\_2007/](http://seer.cancer.gov/csr/1975_2007/), based on November 2009 SEER data submission, posted to the SEER web site, 2010.
- Aluja, A., L. F. Garcia, A. Blanch, and J. Fibla. 2011. Association of androgen receptor gene, CAG and GGN repeat length polymorphism and impulsive-disinhibited personality traits in inmates: the role of short-long haplotype. *Psychiatric Genetics* **21**:229–239.
- Alvarado, C., L. K. Beitel, K. Sircar, A. Aprikian, M. Trifiro, and B. Gottlieb. 2005. Somatic mosaicism and cancer: a micro-genetic examination into the role of the androgen receptor gene in prostate cancer. *Cancer Research* **65**:8514–8518.
- Amato, A. A., T. W. Prior, R. J. Barohn, P. Snyder, A. Papp, and J. R. Mendell. 1993. Kennedy's disease. *Neurology* **43**:791.
- Andersson, P., E. Varenhorst, and P. Söderkvist. 2006. Androgen receptor and vitamin D receptor gene polymorphisms and prostate cancer risk. *European Journal of Cancer* **42**:2833–2837.
- Andrés, A. M., M. Soldevila, O. Lao, V. Volpini, N. Saitou, H. T. Jacobs, I. Hayasaka, F. Calafell, and J. Bertranpetit. 2004. Comparative genetics of functional trinucleotide tandem repeats in humans and apes. *Journal of Molecular Evolution* **59**:329–339.
- Andrés, A. M., M. J. Hubisz, A. Indap, D. G. Torgerson, J. D. Degenhardt, A. R. Boyko, R. N. Gutenkunst et al. 2009. Targets of balancing selection in the human genome. *Molecular Biology and Evolution* **26**:2755–2764.
- Arbizu, T., J. Santamaría, J. Gomez, A. Quílez, and J. P. Serra. 1983. A family with adult spinal and bulbar muscular atrophy, X-linked inheritance and associated testicular failure. *Journal of the Neurological Sciences* **59**:371–382.
- Arnold, A. P., and S. M. Breedlove. 1985. Organizational and activational effects of sex steroids on brain and behavior: A reanalysis. *Hormones and Behavior* **19**:469–498.
- Asatiani, K., S. Eckardstein, M. Simoni, J. Gromoll, and E. Nieschlag. 2003. CAG repeat length in the androgen receptor gene affects the risk of male infertility. *International Journal of Andrology* **26**:255–261.
- Ashley, C.T., and S.T. Warren. 1995. Trinucleotide Repeat Expansion and Human Disease. *Annual Review of Genetics* **29**:703–728.

- Atsuta, N., H. Watanabe, M. Ito, H. Banno, K. Suzuki, M. Katsuno, F. Tanaka, A. Tamakoshi, and G. Sobue. 2006. Natural history of spinal and bulbar muscular atrophy (SBMA): a study of 223 Japanese patients. *Brain* **129**:1446–1455.
- Azzouzi, A., B. Cochand-Priollet, P. Mangin, G. Fournier, P. Berthon, A. Latil, and O. Cussenot. 2002. Impact of constitutional genetic variation in androgen/oestrogen-regulating genes on age-related changes in human prostate. *European Journal of Endocrinology* **147**:479–484.
- Beilin, J., L. Harewood, M. Frydenberg, H. Mameghan, R. F. Martyres, S. J. Farish, C. Yue, D. R. Deam, K. A. Byron, and J. D. Zajac. 2001. A case-control study of the androgen receptor gene CAG repeat polymorphism in Australian prostate carcinoma subjects. *Cancer* **92**: 941–949.
- Berger, M. F., M. S. Lawrence, F. Demichelis, Y. Drier, K. Cibulskis, A. Y. Sivachenko, A. Sboner et al. 2011. The genomic complexity of primary human prostate cancer. *Nature* **470**:214–220.
- Birge, L., M. Pitts, R. Baker, and G. Wilkinson. 2010. Length polymorphism and head shape association among genes with polyglutamine repeats in the stalk-eyed fly, *Teleopsis dalmanni*. *BMC Evolutionary Biology* **10**:227.
- Brooks, B. P., and K. H. Fischbeck. 1995. Spinal and bulbar muscular atrophy: a trinucleotide-repeat expansion neurodegenerative disease. *Trends in Neurosciences* **18**:459–461.
- Buchanan, G., M. Yang, A. Cheong, J. M. Harris, R. A. Irvine, P. F. Lambert, N. L. Moore et al. 2004. Structural and functional consequences of glutamine tract variation in the androgen receptor. *Human Molecular Genetics* **13**:1677–1692.
- Buschiazzo, E., and N. J. Gemmell. 2006. The rise, fall and renaissance of microsatellites in eukaryotic genomes. *BioEssays* **28**:1040–1050.
- Carter, A. J. R., and A. Q. Nguyen. 2011. Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. *BMC Medical Genetics* **12**:160.
- Castel, A. L., J. D. Cleary, and C. E. Pearson. 2010. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nature Reviews Molecular Cell Biology* **11**:165–170.
- Chang, C. 2002. *Androgens and androgen receptor: mechanisms, functions, and clinical applications*. 1st edn. Kluwer Academic Publishers, Boston.
- Chang, J. A., H. T. Nguyen, and T. F. Lue. 2002. Androgens in penile development, penile erection, and erectile dysfunction. In C. Chang, ed. *Androgens and Androgen Receptor: Mechanisms, Functions, and Clinical Applications*, 1st edn. pp. 289–298. Kluwer Academic Publishers, Boston, MA.
- Cheng, D., C.-J. Hong, D.-L. Liao, and S.-J. Tsai. 2006. Association study of androgen receptor CAG repeat polymorphism and male violent criminal activity. *Psychoneuroendocrinology* **31**:548–552.
- Cherrier, M. M., and S. C. Craft. 2003. Androgens and cognition. In C. J. Bagatell, and W. J. Bremner, eds. *Androgens in Health and Disease*, pp. 291–309. Humana Press, Totawa, NJ
- Cho, H., and S. D. Stout. 2011. Age-associated bone loss and intraskeletal variability in the Imperial Romans. *Journal of Anthropological Sciences* **89**:109–125.
- Choong C., J. Kempainen, Z. Zhou, and E. Wilson. 1966. Reduced androgen receptor gene expression with first exon CAG repeat expansion. *Molecular Endocrinology* **10**(12):1527–1535.
- Choong, C. S., J. A. Kempainen, and E. M. Wilson. 1998. Evolution of the primate androgen receptor: a structural basis for disease. *Journal of Molecular Evolution* **47**:334–342.
- Coetzee, G. A., and R. K. Ross. 1994. Re: Prostate cancer and the androgen receptor. *Journal of the National Cancer Institute* **86**:872–873.
- Colangelo, L. A., L. Sharp, P. Kopp, D. Scholtens, B. C.-H. Chiu, K. Liu, and S. M. Gapstur. 2007. Total testosterone, androgen receptor polymorphism, and depressive symptoms in young black and white men: The CARDIA Male Hormone Study. *Psychoneuroendocrinology* **32**:951–958.
- Collins, L. L., and C. Chang. 2002. Androgens and the androgen receptor in male sex development and fertility. In C. Chang, ed. *Androgens and Androgen Receptor: Mechanisms, Functions, and Clinical Applications*, 1st edn. pp. 299–323. Kluwer Academic Publishers, Boston, MA.
- Comings, D. E., R. Gade-Andavolu, L. A. Cone, D. Muhleman, and J. P. MacMurray. 2003. A multigene test for the risk of sporadic breast carcinoma. *Cancer* **97**:2160–2170.
- Cotter, M. M., D. A. Loomis, S. W. Simpson, B. Latimer, and C. J. Hernandez. 2011. Human evolution and osteoporosis-related spinal fractures. *PLoS ONE* **6**(10): e26658. doi:10.1371/journal.pone.0026658.
- Crespi, B. J. 2010. The origins and evolution of genetic disease risk in modern humans. *Annals of the New York Academy of Sciences* **1206**:80–109.
- Dadze, S., C. Wieland, S. Jakubiczka, K. Funke, E. Schröder, B. Royer-Pokora, R. Willers, and P. Wieacker. 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 Caucosoid sample of German origin. *Molecular Human Reproduction* **6**:207–214.
- David A.R., A. Kershaw, and A. Heagerty. 2010. Atherosclerosis and diet in ancient Egypt. *The Lancet* **375**(9716):718–719.
- Davis-Dao, C. A., E. D. Tuazon, R. Z. Sokol, and V. K. Cortessis. 2007. Male infertility and variation in CAG repeat length in the androgen receptor gene: a meta-analysis. *Journal of Clinical Endocrinology and Metabolism* **92**:4319–4326.
- Dejager, S., H. Bry-Gaillard, E. Bruckert, B. Eymard, F. Salachas, E. LeGuern, S. Tardieu, R. Chadarevian, P. Giral, and G. Turpin. 2002. A comprehensive endocrine description of Kennedy's disease revealing androgen insensitivity linked to CAG repeat length. *Journal of Clinical Endocrinology and Metabolism* **87**:3893–3901.
- Di Fabio, F., C. Alvarado, A. Gologan, E. Youssef, L. Voda, E. Mitmaker, L. K. Beitel, P. H. Gordon, and M. Trifiro. 2009. Somatic mosaicism of androgen receptor CAG repeats in colorectal carcinoma epithelial cells from men. *Journal of Surgical Research* **154**:38–44.
- Di Monaco, M., L. Leonardi, V. Gatto, M. Gallo, E. Brignardello, and G. Boccuzzi. 1995. Dihydrotestosterone affects the growth of hormone-unresponsive breast cancer cells: An indirect action. *Anticancer Research* **15**:2581–2584.
- Di Rienzo, A. 2006. Population genetics models of common diseases. *Current Opinion in Genetics & Development* **16**:630–636.
- Díaz-Chico, N., F. Germán Rodríguez, A. González, R. Ramírez, C. Bilbao, A. Cabrera de León, A. Aguirre Jaime, R. Chirino, D. Navarro, and J. C. Díaz-Chico. 2007. Androgens and androgen receptors in breast cancer. *The Journal of Steroid Biochemistry and Molecular Biology* **105**:1–15.
- Dillon, E. L., W. J. Durham, R. J. Urban, and M. Sheffield-Moore. 2010. Hormone treatment and muscle anabolism during aging: androgens. *Clinical Nutrition* **29**:697–700.
- Dowsing, A. T., E. Yong, M. Clark, R. I. McLachlan, D. M. de Kretser, and A. O. Trounson. 1999. Linkage between male infertility and trinucleotide repeat expansion in the androgen-receptor gene. *The Lancet* **354**:640–643.
- von Eckardstein, S., A. Syska, J. Gromoll, A. Kamischke, M. Simoni, and E. Nieschlag. 2001. Inverse correlation between sperm concentration

- and number of androgen receptor CAG Repeats in normal men. *Journal of Clinical Endocrinology and Metabolism* **86**:2585–2590.
- Edwards, A., H. A. Hammond, L. Jin, C. T. Caskey, and R. Chakraborty. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**:241–253.
- Eichler, E. E., J. Flint, G. Gibson, A. Kong, S. M. Leal, J. H. Moore, and J. H. Nadeau. 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics* **11**:446–450.
- Eisenegger, C., J. Haushofer, and E. Fehr. 2011. The role of testosterone in social interaction. *Trends in Cognitive Sciences* **15**:263–271.
- Ekenman, I., S. A. V. Eriksson, and J. U. Lindgren. 1995. Bone density in medieval skeletons. *Calcified Tissue International* **56**:355–358.
- Elhaji, Y. A., B. Gottlieb, R. Lumbroso, L. K. Beitel, W. D. Foulkes, L. Pinsky, and M. A. Trifiro. 2001. The polymorphic CAG repeat of the androgen receptor gene: a potential role in breast cancer in women over 40. *Breast Cancer Research and Treatment* **70**:109–116.
- Ellerby, L. M., A. S. Hackam, S. S. Propp, H. M. Ellerby, S. Rabizadeh, N. R. Cashman, M. A. Trifiro et al. 2002. Kennedy's disease. *Journal of Neurochemistry* **72**:185–195.
- Faltas, B. 2010. Cancer is an ancient disease: the case for better palaeoepidemiological and molecular studies. *Nature Reviews Cancer* **11**:76.
- Ferro P., M. G. Catalano, M. Raineri, G. Reato, R. dell' Eva, M. Risio, R. Foà, N. Fortunati, and U. Pfeffer. 2000. Somatic alterations of the androgen receptor CAG repeat in human colon cancer delineate a novel mutation pathway independent of microsatellite instability. *Cancer Genetics and Cytogenetics* **123**(1):35–40.
- Ferro, P., M. G. Catalano, R. Dell'Eva, N. Fortunati, and U. Pfeffer. 2002. The androgen receptor CAG repeat: a modifier of carcinogenesis? *Molecular and Cellular Endocrinology* **193**:109–120.
- Fondon, J. W., and H. R. Garner. 2004. Molecular origins of rapid and continuous morphological evolution. *Proceedings of the National Academy of Sciences of the United States of America* **101**:18058–18063.
- Fondon, J. W., E. A. D. Hammock, A. J. Hannan, and D. G. King. 2008. Simple sequence repeats: genetic modulators of brain function and behavior. *Trends in Neurosciences* **31**:328–334.
- Forrest, M. S., S. M. Edwards, R. Houlston, Z. Kote-Jarai, T. Key, N. Allen, M. A. Knowles et al. 2005. Association between hormonal genetic polymorphisms and early-onset prostate cancer. *Prostate Cancer and Prostatic Disease* **8**:95–102.
- Freedman, M. L., C. L. Pearce, K. L. Penney, J. N. Hirschhorn, L. N. Kolonel, B. E. Henderson, and D. Altshuler. 2005. Systematic evaluation of genetic variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study. *The American Journal of Human Genetics* **76**:82–90.
- Gatchel, J. R., and H. Y. Zoghbi. 2005. Diseases of unstable repeat expansion: mechanisms and common principles. *Nature Reviews Genetics* **6**:743–755.
- Gatto, V., M. D. Monaco, E. Brignardello, L. Leonardi, M. Gallo, and G. Boccuzzi. 1996. Indirect growth inhibition of the MDA-MB-231 hormone-independent breast cancer cell line by dihydrotestosterone. *Annals of the New York Academy of Sciences* **784**:439–442.
- Gennari L., V. De Paola, D. Merlotti, G. Martini, and R. Nuti. 2007. Steroid hormone receptor gene polymorphisms and osteoporosis: a pharmacogenomic review. *Expert Opinion on Pharmacotherapy* **8**(5):537–553.
- Gemayel, R., M. D. Vines, M. Legendre, and K. J. Verstrepen. 2010. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics* **44**:445–477.
- Giguère, Y., E. Dewailly, J. Brisson, P. Aytote, N. Laflamme, A. Demers, V.-I. Forest, S. Dodin, J. Robert, and F. Rousseau. 2001. Short polyglutamine tracts in the androgen receptor are protective against breast cancer in the general population. *Cancer Research* **61**:5869–5874.
- Gillesen, S., A. Templeton, G. Marra, Y.-F. Kuo, E. Valtorta, and V. B. Shahinian. 2010. Risk of colorectal cancer in men on long-term androgen deprivation therapy for prostate cancer. *Journal of the National Cancer Institute* **102**:1760–1770.
- Gilligan, T., J. Manola, O. Sartor, S. Weinrich, J. Moul, and P. Kantoff. 2004. Absence of a correlation of androgen receptor gene CAG repeat length and prostate cancer risk in an African-American population. *Clinical Genitourinary Cancer* **3**:98–103.
- Giovannucci, E., M. J. Stampfer, K. Krithivas, M. Brown, A. Brufsky, J. Talcott, C. H. Hennekens, and P. W. Kantoff. 1997. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America* **94**:3320–3323.
- Giwerzman, Y. L., J. Richthoff, H. Lilja, C. Anderberg, P. Abrahamsson, and A. Giwerzman. 2004. Androgen receptor CAG repeat length correlates with semen PSA levels in adolescence. *The Prostate* **59**:227–233.
- Gluckman, P. D., F. M. Low, T. Buklijas, M. A. Hanson, and A. S. Beeble. 2011. How evolutionary principles improve the understanding of human health and disease. *Evolutionary Applications* **4**:249–263.
- Gottlieb, B., L. K. Beitel, J. H. Wu, and M. Trifiro. 2004. The androgen receptor gene mutations database (ARDB): 2004 update. *Human Mutation* **23**:527–533.
- Greenland, K. J., J. H. Beilin, J. Castro, P. N. Varghese, and J. D. Zajac. 2004. Polymorphic CAG repeat length in the androgen receptor gene and association with neurodegeneration in a heterozygous female carrier of Kennedy's disease. *Journal of Neurology* **251**:35–41.
- Grierson, A. J., R. C. Mootoosamy, and C. C. J. Miller. 1999. Polyglutamine repeat length influences human androgen receptor/c-Jun mediated transcription. *Neuroscience Letters* **277**:9–12.
- Grönberg, H. 2003. Prostate cancer epidemiology. *The Lancet* **361**:859–864.
- Gu, S., N. Papadopoulou, E.-M. Gehring, O. Nasir, K. Dimas, S. K. Bhavsar, M. Föller, K. Alevizopoulos, F. Lang, and C. Stourmaras. 2009. Functional membrane androgen receptors in colon tumors trigger pro-apoptotic responses in vitro and reduce drastically tumor incidence in vivo. *Molecular Cancer* **8**:114.
- Guadalupe-Grau, A., F. Rodriguez-Gonzalez, J. Ponce-Gonzalez, C. Dorado, H. Olmedillas, T. Fuentes, J. Perez-Gomez, J. Sanchis-Moysi, B. Diaz-Chico, and J. Calbet. 2010. Bone mass and the CAG and ggn androgen receptor polymorphisms in young men. *PLoS ONE* **5**(7):e11529. doi:10.1371/journal.pone.0011529.
- Curven, M., and H. Kaplan. 2007. Longevity among hunter-gatherers: a cross-cultural examination. *Population and Development Review* **33**:321–365.
- Gustafson, D. R., M. J. Wen, and B. M. Koppanati. 2003. Androgen receptor gene repeats and indices of obesity in older adults. *International Journal of Obesity* **27**:75–81.
- Goode E.L., A.M. Dunning, B. Kuschel, C.S. Healey, N.E. Day, B.A.J. Ponder, D.F. Easton, and P.P.D. Pharoah. 2002. Effect of Germ-Line Genetic Variation on Breast Cancer Survival in a Population-based Study. *Cancer Research* **62**(11):3052–3057.
- Haerty, W., and G. B. Golding. 2010a. Low-complexity sequences and single amino acid repeats: not just "junk" peptide sequences. *Genome* **53**:753–762.
- Haerty, W., and G. B. Golding. 2010b. Genome-wide evidence for selection acting on single amino acid repeats. *Genome Research* **20**:755–760.
- Haiman, C. A., M. Brown, S. E. Hankinson, D. Spiegelman, G. A. Colclitz, W. C. Willett, P. W. Kantoff, and D. J. Hunter. 2002. The

- androgen receptor CAG repeat polymorphism and risk of breast cancer in the Nurses' Health Study. *Cancer Research* **62**:1045–1049.
- Haiman, C. A., S. E. Hankinson, I. D. Vivo, C. Guillemette, N. Ishibe, D. J. Hunter, and C. Byrne. 2003. Polymorphisms in steroid hormone pathway genes and mammographic density. *Breast Cancer Research and Treatment* **77**:27–36.
- Hakimi, J. M., M. P. Schoenberg, R. H. Rondinelli, S. Piantadosi, and E. R. Barrack. 1997. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clinical Cancer Research* **3**:1599–1608.
- Hammock, E. A. D., and L. J. Young. 2005. Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* **308**:1630–1634.
- Hammond, J., Q. Le, C. Goodyer, M. Gelfand, M. Trifiro, and A. LeBlanc. 2001. Testosterone-mediated neuroprotection through the androgen receptor in human primary neurons. *Journal of Neurochemistry* **77**:1319–1326.
- Hancock, J. M., E. A. Worthey, and M. F. Santibáñez-Koref. 2001. A role for selection in regulating the evolutionary emergence of disease-causing and other coding CAG repeats in humans and mice. *Molecular Biology and Evolution* **18**:1014–1023.
- Hanke, H., C. Lenz, B. Hess, K.-D. Spindler, and W. Weidemann. 2001. Effect of testosterone on plaque development and androgen receptor expression in the arterial vessel wall. *Circulation* **103**:1382–1385.
- Hannan, A. J. 2010. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. *Trends in Genetics* **26**:59–65.
- Hao, Y., R. Montiel, B. Li, E. Huang, L. Zeng, and Y. Huang. 2010. Association between androgen receptor gene CAG repeat polymorphism and breast cancer risk: a meta-analysis. *Breast Cancer Research and Treatment* **124**:815–820.
- Hawkes, K. 2003. Grandmothers and the evolution of human longevity. *American Journal of Human Biology* **15**:380–400.
- Henderson, B. E., and H. S. Feigelson. 2000. Hormonal carcinogenesis. *Carcinogenesis* **21**:427–433.
- Hersberger, M., J. Muntwyler, H. Funke, J. Marti-Jaun, H. Schulte, G. Assmann, T. F. Luscher, and A. von Eckardstein. 2005. The CAG repeat polymorphism in the androgen receptor gene is associated with HDL-cholesterol but not with coronary atherosclerosis or myocardial infarction. *Clinical Chemistry* **51**:1110–1115.
- Hsing, A. W., L. W. Chu, and F. Z. Stanczyk. 2008. Androgen and prostate cancer: is the hypothesis dead? *Cancer Epidemiology Biomarkers & Prevention* **17**:2525–2530.
- Hurd, P. L., K. L. Vaillancourt, and N. L. Dinsdale. 2011. Aggression, digit ratio and variation in androgen receptor and monoamine oxidase genes in men. *Behavior Genetics* **41**(4):543–556.
- Iobagiu, C., C. Lambert, M. Normand, and C. Genin. 2005. Microsatellite profile in hormonal receptor genes associated with breast cancer. *Breast Cancer Research and Treatment* **95**:153–159.
- Irvine, R. A., M. C. Yu, R. K. Ross, and G. A. Coetzee. 1995. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Research* **55**:1937–1940.
- Jönsson, E. G., C. von Gerten, J. P. Gustavsson, Q.-P. Yuan, K. Lindblad-Toh, K. Forslund, G. Rylander, M. Mattila-Evenden, M. Åsberg, and M. Schalling. 2001. Androgen receptor trinucleotide repeat polymorphism and personality traits. *Psychiatric Genetics* **11**(1):19–23.
- Kaaks, R., S. Rinaldi, T. J. Key, F. Berrino, P. H. M. Peeters, C. Biessy, L. Dossus et al. 2005. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. *Endocrine Related Cancer* **12**:1071–1082.
- Kachel, A. F., L. S. Premo, and J.-J. Hublin. 2011. Grandmothering and natural selection. *Proceedings of the Royal Society of London. Series B, Biological sciences* **278**:384–391.
- Karasik, D. 2008. Osteoporosis: an evolutionary perspective. *Human Genetics* **124**:349–356.
- Karasik, D. 2011. How pleiotropic genetics of the musculoskeletal system can inform genomics and phenomics of aging. *Age* **33**(1):49–62.
- Kashi, Y., and D. G. King. 2006. Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* **22**:253–259.
- Kashi, Y., D. G. King, and M. Soller. 1997. Simple sequence repeats as a source of quantitative genetic variation. *Trends in Genetics* **13**:74–78.
- Katsuno, M., H. Banno, K. Suzuki, H. Adachi, F. Tanaka, and G. Sobue. 2010. Clinical features and molecular mechanisms of spinal and bulbar muscular atrophy (SBMA). In *Diseases of DNA Repair*, 685:64–74–74. *Advances in Experimental Medicine and Biology* **685**:64–74.
- Kazemi-Esfarjani, P., M. A. Trifiro, and L. Pinsky. 1995. Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenetic relevance for the (CAG)<sub>n</sub>-expanded neuropathies. *Human Molecular Genetics* **4**:523–527.
- Kehrer-Sawatzki, H., and D. N. Cooper. 2007. Understanding the recent evolution of the human genome: insights from human–chimpanzee genome comparisons. *Human Mutation* **28**:99–130.
- Kelkar, Y. D., S. Tyekucheva, F. Chiaromonte, and K. D. Makova. 2008. The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Research* **18**:30–38.
- Kelkar, Y. D., K. A. Eckert, F. Chiaromonte, and K. D. Makova. 2011. A matter of life or death: how microsatellites emerge in and vanish from the human genome. *Genome Research* **21**(12):2038–2048.
- Keller, M. C., and G. Miller. 2006. Resolving the paradox of common, harmful, heritable mental disorders: Which evolutionary genetic models work best? *Behavioral and Brain Sciences* **29**:385–404; discussion 405–52.
- Kenny, A. M., and L. G. Raisz. 2003. Androgens and bone. In C. J. Bagatell, and W. J. Bremner, eds. *Androgens in Health and Disease*, pp. 221–232. Humana Press, Totawa, NJ.
- Kittles, R., D. Young, S. Weinrich, J. Hudson, G. Argyropoulos, F. Ukoli, L. Adams-Campbell, and G. Dunston. 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.
- Komori, S., H. Kasumi, R. Kanazawa, K. Sakata, Y. Nakata, H. Kato, and K. Koyama. 1999. CAG repeat length in the androgen receptor gene of infertile Japanese males with oligozoospermia. *Molecular Human Reproduction* **5**:14–16.
- Koshy B.T., and H.Y. Zoghbi. 1997. The CAG/Polyglutamine Tract Diseases: Gene Products and Molecular Pathogenesis. *Brain Pathology* **7**(3):927–942.
- Kovacs, D., E. Vassos, X. Liu, X. Sun, J. Hu, G. Breen, P. Tompa, D. A. Collier, and T. Li. 2009. The androgen receptor gene polyglycine repeat polymorphism is associated with memory performance in healthy Chinese individuals. *Psychoneuroendocrinology* **34**:947–952.
- Ku, C. S., E. Y. Loy, A. Salim, Y. Pawitan, and K. S. Chia. 2010. The discovery of human genetic variations and their use as disease markers: past, present and future. *Journal of Human Genetics* **55**:403–415.
- Kumar, R., H. Atamna, M. N. Zakharov, S. Bhasin, S. H. Khan, and R. Jasuja. 2011. Role of the androgen receptor CAG repeat polymorphism in prostate cancer, and spinal and bulbar muscular atrophy. *Life Sciences* **88**:565–571.
- La Spada, A. R., E. M. Wilson, D. B. Lubahn, A. E. Harding, and K. H. Fischbeck. 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**:77–79.

- Langdahl, B. L., L. Stenkjaer, M. Carstens, C. L. Tofteng, and E. F. Erikssen. 2003. A CAG repeat polymorphism in the androgen receptor gene is associated with reduced bone mass and increased risk of osteoporotic fractures. *Calcified Tissue International* **73**:237–243.
- Lascorz, J., A. Försti, B. Chen, S. Buch, V. Steinke, N. Rahner, E. Holinski-Feder et al. 2010. Genome-wide association study for colorectal cancer identifies risk polymorphisms in German familial cases and implicates MAPK signalling pathways in disease susceptibility. *Carcinogenesis* **31**:1612–1619.
- Latil, A. G., R. Azzouzi, G. S. Cancel, E. C. Guillaume, B. Cochran-Priollet, P. L. Berthon, and O. Cussenot. 2001. Prostate carcinoma risk and allelic variants of genes involved in androgen biosynthesis and metabolism pathways. *Cancer* **92**:1130–1137.
- Lazaros, L., N. Xita, A. Kaponis, K. Zikopoulos, N. Sofikitis, and I. Georgiou. 2008. Evidence for association of sex hormone-binding globulin and androgen receptor genes with semen quality. *Andrologia* **40**:186–191.
- Lee, D. M., A. Ulubaev, A. Tajar, S. R. Pye, N. Pendleton, N. Purandare, T. W. O'Neill et al. 2010. Endogenous hormones, androgen receptor CAG repeat length and fluid cognition in middle-aged and older men: results from the European Male Ageing Study. *European Journal of Endocrinology* **162**:1155–1164.
- Liede, A., W. Zhang, M. L. D. L. Matsuda, A. Tan, and S. A. Narod. 2003. Androgen receptor gene polymorphism and breast cancer susceptibility in the Philippines. *Cancer Epidemiology Biomarkers & Prevention* **12**:848–852.
- Lillie, E. O., L. Bernstein, S. A. Ingles, W. J. Gauderman, G. E. Rivas, V. Gagalang, T. Krontiris, and G. Ursin. 2004. Polymorphism in the androgen receptor and mammographic density in women taking and not taking estrogen and progestin therapy. *Cancer Research* **64**:1237–1241.
- Lim, H. N., H. Chen, S. McBride, A. M. Dunning, R. M. Nixon, I. A. Hughes, and J. R. Hawkins. 2000. Longer polyglutamine tracts in the androgen receptor are associated with moderate to severe undermasculinized genitalia in XY males. *Human Molecular Genetics* **9**:829–834.
- Limer, K. L., S. R. Pye, W. Thomson, S. Boonen, H. Borghs, D. Vanderschueren, I. T. Huhtaniemi et al. 2009. Genetic variation in sex hormone genes influences heel ultrasound parameters in middle-aged and elderly men: results from the European Male Aging Study (EMAS). *Journal of Bone and Mineral Research* **24**:314–323.
- Lind, J. M., C. Chiu, J. Ingles, L. Yeates, S. E. Humphries, A. K. Heather, and C. Semsarian. 2008. Sex hormone receptor gene variation associated with phenotype in male hypertrophic cardiomyopathy patients. *Journal of Molecular and Cellular Cardiology* **45**:217–222.
- Lindstrom, S., J. Ma, D. Altshuler, E. Giovannucci, E. Riboli, D. Albanes, N. E. Allen et al. 2010. A large study of androgen receptor germline variants and their relation to sex hormone levels and prostate cancer risk. Results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium. *Journal of Clinical Endocrinology and Metabolism* **95**:E121–E127.
- Looker, A. C., E. S. Orwoll, C. C. Johnston Jr, R. L. Lindsay, H. W. Wahner, W. L. Dunn, M. S. Calvo, T. B. Harris, and S. P. Heyse. 1997. Prevalence of low femoral bone density in older U.S. Adults from NHANES III. *Journal of Bone and Mineral Research* **12**:1761–1768.
- López-Otín, C. L., and E. P. Diamandis. 1998. Breast and prostate cancer: an analysis of common epidemiological, genetic, and biochemical features. *Endocrine Reviews* **19**:365–396.
- Luo, H., K. Lin, A. David, H. Nijveen, and J. A. M. Leunissen. 2012. ProRepeat: an integrated repository for studying amino acid tandem repeats in proteins. *Nucleic Acids Research* **40**(Database issue):D394–399.
- Macleod, H. E., R. W. Brown, J. Beilin, G. L. Warne, and J. D. Zajac. 2004. Increased frequency of long androgen receptor CAG repeats in male breast cancers. *Breast Cancer Research and Treatment* **88**:239–246.
- Manning, J. T. 2007. The androgen receptor gene: A major modifier of speed of neuronal transmission and intelligence? *Medical Hypotheses* **68**:802–804.
- Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorf, D. J. Hunter, M. I. McCarthy et al. 2009. Finding the missing heritability of complex diseases. *Nature* **461**:747–753.
- Manuck, S. B., A. L. Marsland, J. D. Flory, A. Gorka, R. E. Ferrell, and A. R. Hariri. 2010. Salivary testosterone and a trinucleotide (CAG) length polymorphism in the androgen receptor gene predict amygdala reactivity in men. *Psychoneuroendocrinology* **35**:94–104.
- Mariotti, C., B. Castellotti, D. Pareyson, D. Testa, M. Eoli, C. Antozzi, V. Silani et al. 2000. Phenotypic manifestations associated with CAG-repeat expansion in the androgen receptor gene in male patients and heterozygous females: a clinical and molecular study of 30 families. *Neuromuscular Disorders* **10**:391–397.
- Mays, S. A. 1998. Osteoporosis in earlier human populations. *Journal of Clinical Densitometry* **2**:71–78.
- Mays, S. A. 2006. A palaeopathological study of Colles' fracture. *International Journal of Osteoarchaeology* **16**:415–428.
- Mazur, A., and A. Booth. 1998. Testosterone and dominance in men. *Behavioral and Brain Sciences* **21**:353–363.
- Meacham, R. B., G. F. Joyce, M. Wise, A. Kparker, and C. Niederberger. 2007. Male infertility. *The Journal of Urology* **177**:2058–2066.
- Melton, L. J. 2001. The prevalence of osteoporosis: gender and racial comparison. *Calcified Tissue International* **69**:179–181.
- Meys, E. R.-D., H. Leffers, J. H. Petersen, A.-G. Andersen, E. Carlsen, N. Jørgensen, and N. E. Skakkebaek. 2002. CAG repeat length in androgen-receptor gene and reproductive variables in fertile and infertile men. *The Lancet* **359**:44–46.
- Mononen, N., T. Ikonen, V. Autio, A. Rökman, M. Matikainen, T. Tammele, O.-P. Kallioniemi, P. Koivisto, and J. Schleutker. 2002. Androgen receptor CAG polymorphism and prostate cancer risk. *Human Genetics* **111**:166–171.
- Monroe, K. R., M. C. Yu, L. N. Kolonel, G. A. Coetzee, L. R. Wilkens, R. K. Ross, and B. E. Henderson. 1995. Evidence of an X-linked or recessive genetic component to prostate cancer risk. *Nature Medicine* **1**:827–829.
- Mularoni, L., A. Ledda, M. Toll-Riera, and M. M. Albà. 2010. Natural selection drives the accumulation of amino acid tandem repeats in human proteins. *Genome Research* **20**:745–754.
- Murphy W.A. Jr, D. zur Nedden, P. Gostner, R. Knapp, W. Recheis, and H. Seidler. 2003. The iceman: discovery and imaging. *Radiology* **226**(3):614–629.
- Neonen, H. A., A. Giwercman, E. Hallengren, and Y. L. Giwercman. 2010. Non-linear association between androgen receptor CAG repeat length and risk of male subfertility – a meta-analysis. *International Journal of Andrology* **34**:327–332.
- Nesse, R. M. 2011. Ten questions for evolutionary studies of disease vulnerability. *Evolutionary Applications* **4**:264–277.
- Page, S. T., V. Kupelian, W. J. Bremner, and J. B. McKinlay. 2006. The androgen receptor gene CAG repeat polymorphism does not predict increased risk of heart disease: longitudinal results from the Massachusetts Male Ageing Study. *Clinical Endocrinology* **65**:333–339.
- Palazzolo, I., A. Gliozzi, P. Rusmini, D. Sau, V. Crippa, F. Simonini, E. Onesto, E. Bolzoni, and A. Poletti. 2008. The role of the polyglutamine tract in androgen receptor. *The Journal of Steroid Biochemistry and Molecular Biology* **108**:245–253.

- Panz, V. R., B. I. Joffe, I. Spitz, T. Lindenberg, A. Farkas, and M. Haffeejee. 2001. Tandem CAG repeats of the androgen receptor gene and prostate cancer risk in black and white men. *Endocrine* **15**:213–216.
- Pausova, Z., M. Abrahamowicz, A. Mahboubi, C. Syme, G. T. Leonard, M. Perron, L. Richer, S. Veillette, D. Gaudet, and T. Paus. 2010. Functional variation in the androgen-receptor gene is associated with visceral adiposity and blood pressure in male adolescents. *Hypertension* **55**:706–714.
- Pavard, S., and C. J. E. Metcalf. 2007. Negative selection on BRCA1 susceptibility alleles sheds light on the population genetics of late-onset diseases and aging theory. *PLoS ONE* **2**:e1206.
- Pearson, C. E., K. N. Edamura, and J. D. Cleary. 2005. Repeat instability: mechanisms of dynamic mutations. *Nature Reviews Genetics* **6**:729–742.
- Perrin, J. S., P.-Y. Hervé, G. Leonard, M. Perron, G. B. Pike, A. Pitiot, L. Richer, S. Veillette, Z. Pausova, and T. Paus. 2008. Growth of white matter in the adolescent brain: role of testosterone and androgen receptor. *The Journal of Neuroscience* **28**:9519–9524.
- Pettaway, C. A. 1999. Racial differences in the androgen/androgen receptor pathway in prostate cancer. *Journal of the National Medical Association* **91**:653–660.
- Platz, E. A., M. F. Leitzmann, N. Rifai, P. W. Kantoff, Y.-C. Chen, M. J. Stampfer, W. C. Willett, and E. Giovannucci. 2005. Sex steroid hormones and the androgen receptor gene CAG Repeat and subsequent risk of prostate cancer in the prostate-specific antigen era. *Cancer Epidemiology Biomarkers & Prevention* **14**:1262–1269.
- Polanski, A., R. Chakraborty, M. Kimmel, and R. Deka. 1998. Dynamic balance of segregation distortion and selection maintains normal allele sizes at the myotonic dystrophy locus. *Mathematical Biosciences* **147**:93–112.
- Poulsen, L., D. Qvesel, K. Brixen, A. Vesterby, and J. Boldsen. 2001. Low bone mineral density in the femoral neck of medieval women: a result of multiparity? *Bone* **28**:454–458.
- Rajender, S., L. Singh, and K. Thangaraj. 2007. Phenotypic heterogeneity of mutations in androgen receptor gene. *Asian Journal of Andrology* **9**:147–179.
- Rajender, S., G. Pandu, J. D. Sharma, K. P. C. Gandhi, L. Singh, and K. Thangaraj. 2008. Reduced CAG repeats length in androgen receptor gene is associated with violent criminal behavior. *International Journal of Legal Medicine* **122**:367–372.
- Rebbeck, T. R., P. W. Kantoff, K. Krithivas, S. Neuhausen, M. A. Blackwood, A. K. Godwin, M. B. Daly et al. 1999. Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. *The American Journal of Human Genetics* **64**:1371–1377.
- Reddy P.S., and D.E Housman. 1997. The complex pathology of trinucleotide repeats. *Current Opinion in Cell Biology* **9**(3):364–372.
- Rexrode, K. M., P. M. Ridker, H. H. Hegener, J. E. Buring, J. E. Manson, and R. Y. L. Zee. 2008. Genetic variation of the androgen receptor and risk of myocardial infarction and ischemic stroke in women. *Stroke* **39**:1590–1592.
- Rodríguez-González, G., R. Ramírez-Moreno, P. Pérez, C. Bilbao, L. López-Ríos, J. C. Díaz-Chico, P. C. Lara, L. Serra-Majem, R. Chirino, and B. N. Díaz-Chico. 2009. The GGN and CAG repeat polymorphisms in the exon-1 of the androgen receptor gene are, respectively, associated with insulin resistance in men and with dyslipidemia in women. *The Journal of Steroid Biochemistry and Molecular Biology* **113**:202–208.
- Roney, J. R., Z. L. Simmons, and A. W. Lukaszewski. 2010. Androgen receptor gene sequence and basal cortisol concentrations predict men's hormonal responses to potential mates. *Proceedings of the Royal Society of London. Series B, Biological sciences* **277**:57–63.
- Ross, R. K., L. Pearce, J. K. V. Reichardt, and G. A. Coetzee. 2005. Androgens and prostate cancer etiology: sorting through the evidence. In J. J. Li, S. A. Li, and A. Llombart-Bosch, eds. *Hormonal Carcinogenesis*, Vol. 4: pp. 183–196. Springer-Verlag, New York, NY.
- Rubinsztein, D. C., W. Amos, J. Leggo, S. Goodburn, S. Jain, S.-H. Li, R. L. Margolis, C. A. Ross, and M. A. Ferguson-Smith. 1995. Microsatellite evolution—evidence for directionality and variation in rate between species. *Nature Genetics* **10**:337–343.
- Ruizeveld de Winter, J. A., J. Trapman, M. Vermey, E. Mulder, N. D. Zegers, and T. H. van der Kwast. 1991. Androgen receptor expression in human tissues: an immunohistochemical study. *Journal of Histochemistry & Cytochemistry* **39**:927–936.
- dos Santos, M. L., Á. S. Sarkis, I. N. Nishimoto, and M. A. Nagai. 2003. Androgen receptor CAG repeat polymorphism in prostate cancer from a Brazilian population. *Cancer Detection and Prevention* **27**:321–326.
- Schatzl, G., S. Madersbacher, A. Gsur, M. Preyer, G. Haidinger, A. Haitel, C. Vutuc, M. Micksche, and M. Marberger. 2002. Association of polymorphisms within androgen receptor, 5 $\alpha$ -reductase, and PSA genes with prostate volume, clinical parameters, and endocrine status in elderly men. *The Prostate* **52**:130–138.
- Seidman, S. N., A. B. Araujo, S. P. Roose, and J. B. McKinlay. 2001. Testosterone level, androgen receptor polymorphism, and depressive symptoms in middle-aged men. *Biological Psychiatry* **50**:371–376.
- Shibata, A., T. A. Stamey, J. E. McNeal, I. Cheng, and D. M. Peehl. 2001. Genetic polymorphisms in the androgen receptor and type-II 5 $\alpha$ -reductase genes in prostate enlargement. *The Journal of Urology* **166**:1560–1564.
- Simmons, Z. L., and J. R. Roney. 2011. Variation in CAG repeat length of the androgen receptor gene predicts variables associated with intrasexual competitiveness in human males. *Hormones and Behavior* **60**:306–312.
- Slattery, M. L. 2006. Single nucleotide polymorphisms in sporadic colorectal cancer. *Current Colorectal Cancer Reports* **2**:200–205.
- Slattery, M. L., C. Sweeney, M. Murtaugh, K. N. Ma, R. K. Wolff, J. D. Potter, B. J. Caan, and W. Samowitz. 2005. Associations between ER $\alpha$ , ER $\beta$ , and AR genotypes and colon and rectal cancer. *Cancer Epidemiology Biomarkers & Prevention* **14**:2936–2942.
- Sobczak, K., and W. J. Krzyzosiak. 2004. Patterns of CAG repeat interruptions in SCA1 and SCA2 genes in relation to repeat instability. *Human Mutation* **24**:236–247.
- Spurdle, A. B., A. C. Antoniou, D. L. Duffy, N. Pandeya, L. Kelemen, X. Chen, S. Peock et al. 2005. The androgen receptor CAG repeat polymorphism and modification of breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Research* **7**:R176–R183.
- Stanford, J. L., J. J. Just, M. Gibbs, K. G. Wicklund, C. L. Neal, B. A. Blumstein, and E. A. Ostrander. 1997. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Research* **57**:1194–1198.
- Stearns, S. C., S. G. Byars, D. R. Govindaraju, and D. Ewbank. 2010. Measuring selection in contemporary human populations. *Nature Reviews Genetics* **11**:611–622.
- Summers, K., and B. J. Crespi. 2008. The androgen receptor and prostate cancer: A role for sexual selection and sexual conflict. *Medical Hypotheses* **70**:435–443.
- Sutherland G.R., and R.I. Richards. 1995. Simple tandem DNA repeats and human genetic disease. *Proceedings of the National Academy of Sciences of the United States of America* **92**(9):3636–3641.



- Suter, N. M., K. E. Malone, J. R. Daling, D. R. Doody, and E. A. Ostrander. 2003. Androgen receptor (CAG)<sub>n</sub> and (GGC)<sub>n</sub> polymorphisms and breast cancer risk in a population-based case-control study of young women. *Cancer Epidemiology Biomarkers & Prevention* **12**:127–135.
- Swerdlow, R. S., C. Wang, M. Hines, and R. Gorski. 1992. Effect of androgens on the brain and other organs during development and aging. *Psychoneuroendocrinology* **17**:375–383.
- Szelei, J., J. Jimenez, A. M. Soto, M. F. Luizzi, and C. Sonnenschein. 1997. Androgen-induced inhibition of proliferation in human breast cancer MCF7 cells transfected with androgen receptor. *Endocrinology* **138**:1406–1412.
- Thangaraj, K., M. B. Joshi, A. G. Reddy, N. J. Gupta, B. Chakravarty, and L. Singh. 2002. CAG repeat expansion in the androgen receptor gene is not associated with male infertility in Indian populations. *Journal of Andrology* **23**:815–818.
- Tilley, W. D., G. Buchanan, T. E. Hickey, and J. M. Bentel. 1996. Mutations in the androgen receptor gene are associated with progression of human prostate cancer to androgen independence. *Clinical Cancer Research* **2**:277–285.
- Tofteng, C. L., A. Kindmark, H. Brandstrom, B. Abrahamsen, S. Petersen, F. Stiger, L. S. Stilgren et al. 2003. Polymorphisms in the CYP19 and AR genes? relation to bone mass and longitudinal bone changes in postmenopausal women with or without hormone replacement therapy: the danish osteoporosis prevention study. *Calcified Tissue International* **74**:25–34.
- Turner-Walker, G., U. Syversen, and S. Mays. 2001. The archaeology of osteoporosis. *European Journal of Archaeology* **4**:263–269.
- Tut, T. G., F. J. Ghadessy, M. A. Trifiro, L. Pinsky, and E. L. Yong. 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.
- Verhoeven, G., A. Willems, E. Denolet, J. V. Swinnen, and K. De Gendt. 2010. Androgens and spermatogenesis: lessons from transgenic mouse models. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**:1537–1556.
- Vermeersch, H., G. T'Sjoen, J. M. Kaufman, J. Vincke, and M. Van Houtte. 2010. Testosterone, androgen receptor gene CAG repeat length, mood and behaviour in adolescent males. *European Journal of Endocrinology* **163**:319–328.
- Vismara, G., F. Simonini, E. Onesto, M. Bignamini, V. Miceli, L. Martini, and A. Poletti. 2009. Androgens inhibit androgen receptor promoter activation in motor neurons. *Neurobiology of Disease* **33**:395–404.
- Vowles, E. J., and W. Amos. 2006. Quantifying ascertainment bias and species-specific length differences in human and chimpanzee microsatellites using genome sequences. *Molecular Biology and Evolution* **23**:598–607.
- Wang, D. Y., D. S. Allen, B. L. D. Stavola, I. S. Fentiman, J. Brussen, R. D. Bulbrook, B. S. Thomas, J. L. Hayward, and M. J. Reed. 2000. Urinary androgens and breast cancer risk: results from a long-term prospective study based in Guernsey. *British Journal of Cancer* **82**:1577–1584.
- Wang, W., E.M. John, and S.A. Ingles. 2005. Androgen Receptor and Prostate-Specific Antigen Gene Polymorphisms and Breast Cancer in African-American Women. *Cancer Epidemiology Biomarkers & Prevention* **14**(12):2990–2994.
- Weidemann, W., and H. Hanke. 2002. Cardiovascular effects of androgens. *Cardiovascular Drug Reviews* **20**:175–198.
- Westberg, L., S. Henningsson, M. Landén, K. Annerbrink, J. Melke, S. Nilsson, R. Rosmond, G. Holm, H. Anckarsäter, and E. Eriksson. 2009. Influence of androgen receptor repeat polymorphisms on personality traits in men. *Journal of Psychiatry & Neuroscience: JPN* **34**:205–213.
- Wu, F. C. W., and A. von Eckardstein. 2003. Androgens and coronary artery disease. In C. J. Bagatell, and W. J. Bremner, eds. *Androgens in Health and Disease*, pp. 191–220. Humana Press, Totawa, NJ.
- Xiao, R., L. J. Hennings, T. M. Badger, and F. A. Simmen. 2007. Fetal programming of colon cancer in adult rats: correlations with altered neonatal growth trajectory, circulating IGF-I and IGF binding proteins, and testosterone. *Journal of Endocrinology* **195**:79–87.
- Yaffe, K., E. R. Edwards, L.-Y. Lui, J. M. Zmuda, R. E. Ferrell, and J. A. Cauley. 2003. Androgen receptor CAG repeat polymorphism is associated with cognitive function in older men. *Biological Psychiatry* **54**:943–946.
- Yamada, Y., F. Ando, N. Niino, and H. Shimokata. 2004. Association of polymorphisms of the androgen receptor and klotho genes with bone mineral density in Japanese women. *Journal of Molecular Medicine* **83**:50–57.
- Yong, E. L., C. J. Loy, and K. S. Sim. 2003. Androgen receptor gene and male infertility. *Human Reproduction Update* **9**:1–7.
- Yu, H., B. Bharaj, E. J. K. Vassilikos, M. Giai, and E. P. Diamandis. 2000. Shorter CAG repeat length in the androgen receptor gene is associated with more aggressive forms of breast cancer. *Breast Cancer Research and Treatment* **59**:153–161.
- Zeegeers, M. P., L. A. L. M. Kiemeny, A. M. Nieder, and H. Ostrer. 2004. How strong is the association between CAG and GGN repeat length polymorphisms in the androgen receptor gene and prostate cancer risk? *Cancer Epidemiology Biomarkers & Prevention* **13**:1765–1771.
- Zimmerman, M. R. 1993. The paleopathology of the cardiovascular system. *Texas Heart Institute Journal* **20**:252–257.
- Zitzmann, M. 2009. The role of the CAG repeat androgen receptor polymorphism in andrology. *Frontiers of Hormone Research* **37**:52–61.
- Zitzmann, M., M. Brune, B. Kornmann, J. Gromoll, J. S. von Eckardstein, A. von Eckardstein, and E. Nieschlag. 2001a. The CAG repeat polymorphism in the AR gene affects high density lipoprotein cholesterol and arterial vasoreactivity. *Journal of Clinical Endocrinology and Metabolism* **86**:4867–4873.
- Zitzmann, M., M. Brune, B. Kornmann, J. Gromoll, R. Junker, and E. Nieschlag. 2001b. The CAG repeat polymorphism in the androgen receptor gene affects bone density and bone metabolism in healthy males. *Clinical Endocrinology* **55**:649–657.
- Zitzmann, M., J. Gromoll, E. von Eckardstein, and E. Nieschlag. 2003. The CAG repeat polymorphism in the androgen receptor gene modulates body fat mass and serum concentrations of leptin and insulin in men. *Diabetologia* **46**:31–39.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** AR CAGn distribution among human ethnic populations.

**Table S1.** Phenotypes and diseases of androgens and the AR CAGn.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.