LUND UNIVERSITY

Androgenic impact on prostate cancer risk

Bentmar Holgersson, Magdalena

2017

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Bentmar Holgersson, M. (2017). *Androgenic impact on prostate cancer risk*. [Doctoral Thesis (compilation), Department of Translational Medicine]. Lund University: Faculty of Medicine.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

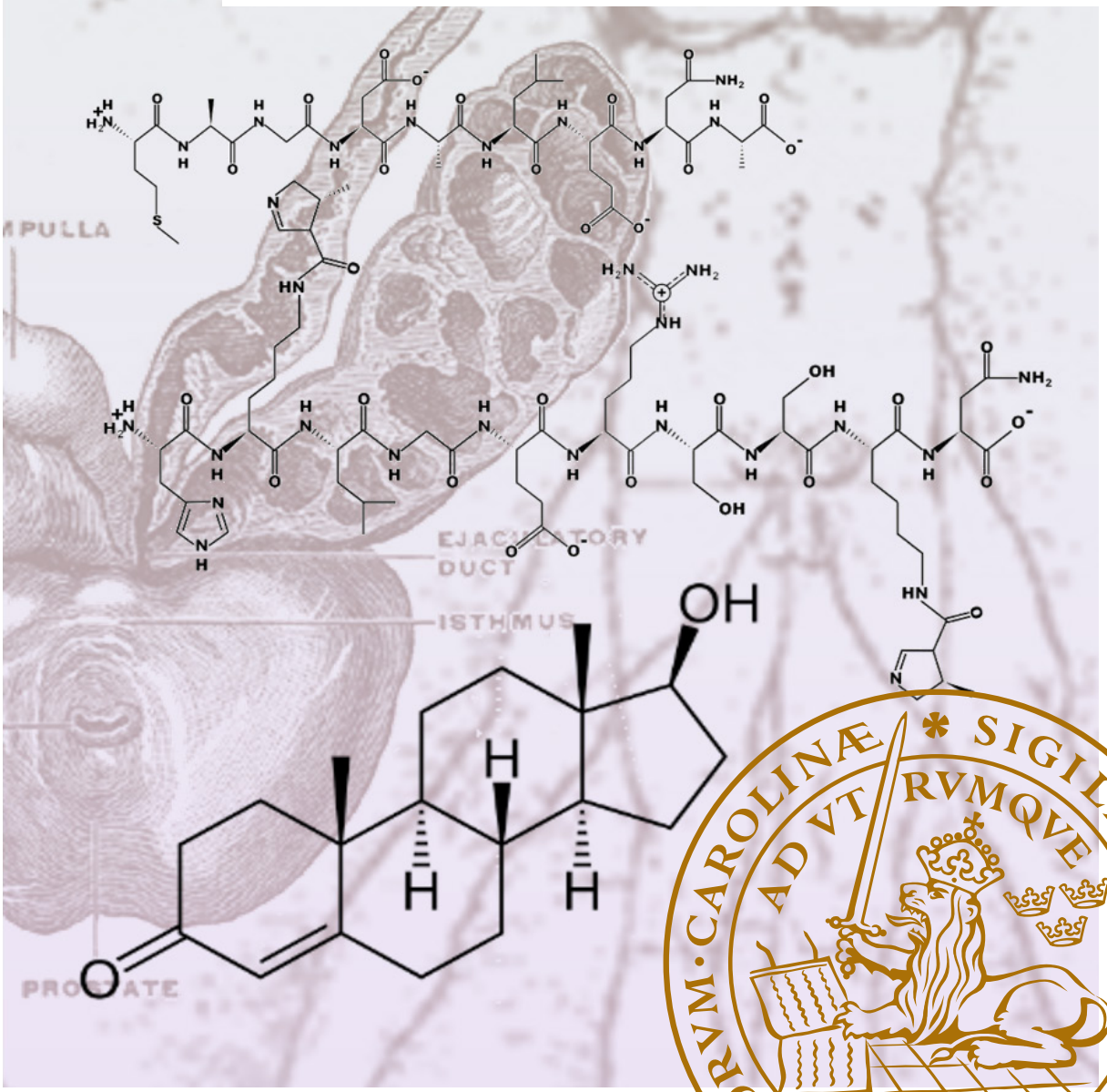
Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Androgenic impact on prostate cancer risk

MAGDALENA BENTMAR HOLGERSSON | FACULTY OF MEDICINE | LUND UNIVERSITY



Androgenic impact on prostate cancer risk

Magdalena Bentmar Holgersson

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Kvinnoklinikens aula, SUS, Malmö.

Friday 10th of November 2017 at 9.00.

Faculty opponent
Pernilla Wikström
Department of Medical Biosciences,
Faculty of Medicine, Umeå University, Umeå, Sweden

Organization
LUND UNIVERSITY
Faculty of Medicine
Department of Translational Medicine

Division of Molecular Genetic Reproductive Medicine

Author: Magdalena Bentmar Holgersson

Document name
DOCTORAL DISSERTATION

Date of issue
November 10th 2017

Sponsoring organization

Title: Androgenic impact on prostate cancer risk

Abstract
Prostate cancer (PCa) is the most common cancer in Swedish men and approximately 16% of all men will receive a PCa diagnosis before his 75th birthday. Testosterone (T) is crucial for the growth of both the healthy prostate and PCa, and is also important in the activation of the expression prostate specific antigen (PSA). The central role of T in PCa, and in the making of PSA, a clinical marker for prostatic disease, has led to the common belief that high T is a risk factor for PCa. T and its metabolite 5-alpha dihydrotestosterone (DHT) acts through the androgen receptor (AR), and genetic variants of the AR have been associated with PCa in several studies. The aims of this thesis were to investigate if T concentrations or the combination of genetic variants in the AR could influence the risk of being diagnosed with PCa and if AR-variants could affect the risk of having PSA above clinical thresholds on the suspicion of PCa in men without PCa. Additionally, the long term effect of extreme levels of T on mortality risk was also investigated, as previous studies have indicated an association between low T and mortality, but without determining the direction of the association. For the association studies regarding AR-variants and PSA concentrations, the European Male Ageing study (EMAS) was used (n after exclusion=1,804). For the association studies regarding AR-variants and PCa risk the EMAS cohort was combined with men from the Swedish Osteoporotic fractures in men (MrOS; n after exclusion=1,120) and a nested case-control sample of the Malmö diet and cancer study (MDCS; n after exclusion=883) to a final dataset of n=689 PCa cases and n=3214 men without PCa. Odds ratios (OR) and 95% confidence intervals (95%CI) were calculated for the risk of PSA above 3 or 4 ng/mL in the n=1,744 men in EMAS without PCa, and for the risk of PCa diagnosis in the combined dataset. For the associations between T concentration and risk of PCa and mortality, all men from whom T measurements were analyzed (n=4278) at the Department of Clinical Chemistry, Skåne University Hospital, Malmö, Sweden 1987-1992 were linked to national Cancer Registry and Death Registry. Hazard ratios (HR) and 95%CIs were calculated for the risk of mortality or being diagnosed with prostate cancer after more than 20 years of follow-up, for the men with the 5% highest and 5% lowest T concentrations compared to the men in the >10% to <90% range. All analyses were adjusted for age.The results revealed two dominant AR haplotypes, H1 and H2 with a frequency of ~85 and ~15%, respectively, in European men. The H2 haplotype had statistically significantly shorter CAG-repeat lengths (*p<0.001*), statistically significantly lower risk of PCa: OR (95%CI) = 0.70 (0.54-0.91; p=*0.007*), and statistically significantly increased risk of having PSA above 3 and 4 ng/mL; OR (95%CI) = 1.69 (1.13-2.52, p=*0.011*) and 1.99 (1.21-3.29; *p=0.007*), respectively. After 20 years of follow-up, no differences in PCa risk were detected for the men with the 5% lowest T (*p=0.122*) or 5% highest T (*p=0.282*). An increased risk of all-cause mortality was detected in the lowest 5% group, HR (95%CI) = 1.39 (1.10-1.75; *p=0.006*) and when divided into younger (<50 years) and older (>50 years) at T measurement, the increased risk was found to be statistically significant only in the younger men; HR (95%CI) = 2.31 (1.48-3.60; *p<0.001*). In conclusion, this thesis show a haplotype difference in PCa risk and risk of having PSA above clinical cutoffs, which might indicate a subgroup of men for which the PSA-test should be haplotype-adjusted. In addition, high T was not a risk factor for PCa, but low T might be an indicator of future risk of all-cause mortality, calling for closer monitoring of T by clinicians.

Key words: Testosterone, androgen receptor, genetic variants, prostate cancer, mortality

Classification system and/or index terms (if any)

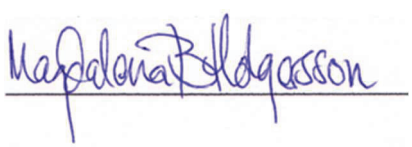
Supplementary bibliographical information | Language English

ISSN and key title 1652-8220 | ISBN 978-91-7619-532-1

Recipient's notes | Number of pages 94 | Price

Security classification

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature  Date 5/10-2017

Androgenic impact on prostate cancer risk

Magdalena Bentmar Holgersson

LUND
UNIVERSITY

Coverphoto collage:

"Vitruvian Man", Leonardo da Vinci, ~1490

"Anatomy of the Human Body: figure 1160", Henry Gray, 1918

"Structure of testosterone", Wikimedia commons

Copyright © Magdalena Bentmar Holgersson

Lund University, Faculty of Medicine Doctoral Dissertation Series 2017:149
Department of Translational Medicine

ISBN 978-91-7619-532-1
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2017

F--k cancer

Preface

When I started my work on this thesis in 2013, the double edged sword of the prostate specific antigen (PSA) screening test was one of the main issues being discussed. Early detection of prostate cancer (PCa) can be crucial for the survival of the patient, but can also lead to over-diagnosis and overtreatment of men, who should perhaps never had become PCa patients as they would never have noticed their tumor during their life-time, had it not been for the increased PSA concentration in their blood.

In addition, the role of testosterone in the etiology and progression of PCa was debated, as was the benefit-risk ratio of testosterone supplementation to men with low testosterone.

Being able to find commonalities between unique patients can hopefully reduce their suffering by reducing their risk of aggressive disease and reducing their treatment related side-effects by creating a tailor-made risk profile for each individual patient based on the knowledge gathered on a population level.

The aim of my work has been to investigate whether genetic variants of the androgen receptor or levels of testosterone in younger men could affect the risk of PCa with the hope that the identification of preexisting variables affecting PCa risk in the future could be used to adjust the diagnostic tests and make the diagnosis more specific.

Content

Preface 7

Content 8

Abbreviations 10

List of papers 12

Review of the literature 13

- Part I: Androgens 13
  - Short history of testosterone 13
  - Testosterone synthesis 14
  - The hypothalamic-Pituitary-Gonadal (HPG) axis 16
  - The importance of testosterone 17
  - Low testosterone 19
- Part II: The androgen receptor 20
  - Short history of the discovery of the androgen receptor and androgen receptor related diseases and disorders 20
  - Nuclear receptors 22
  - The androgen receptor gene 25
  - The androgen receptor protein 31
- Part III: Prostate cancer 34
  - Short history of prostate cancer 34
  - The healthy prostate 35
  - Prostate cancer 37
  - Testosterone concentration and prostate cancer 37
  - Risk factors, incidence and mortality 38
  - Prostate specific antigen 40

Aims 41

Materials and Methods 43

- Subjects for genetic associations 43
  - Genotypes 46
  - Haplotype construction 48
- Subjects without genotype information 50
  - Testosterone studies 50

Statistics 54
Results and discussion 56
Testosterone concentrations and risk of PCa 56
Testosterone concentrations and risk of mortality 57
Androgen receptor haplotypes 60
Potential causes for haplotype risk differences 68
Populärvetenskaplig sammanfattning 71
Acknowledgements 73
References 75
Original publications 94

Abbreviations

ABP	Androgen binding protein
AC	Adenylyl cyclase
AF	Activation function
AIS	Androgen insensitivity syndrome
AFR	African superpopulation of 1000 genomes
AMR	Ad Mixed American superpopulation of 1000 genomes
AR	Androgen receptor
ARE	Androgen response elements
BPH	Benign prostate hypertrophy
cAMP	Cyclic adenosine monophosphate
CI	Confidence interval
CRPC	Castration resistant prostate cancer
CYP11A1	Cytochrome P450 family 11 subfamily A member 1
D-box	Distal box
DBD	DNA-binding domain
DHT	Dihydrotestosterone
DNA	Deoxyribonucleic acid
E_2	Estradiol
EAS	East Asian superpopulation of 1000 genomes
EMAS	European Male Ageing Study
EPIC	European Prospective Investigation into Diet and Cancer
EUR	European superpopulation of 1000 genomes
FSH	Follicle stimulating hormone
GnRH	Gonadotropin-releasing hormone
HPG	Hypothalamic-Pituitary-Gonadal
HR	Hazard ratio
HSP	Heat-shock proteins
ICD	International classification of diseases
KLK3	Kallikrein-3
LBD	Ligand binding domain
LD	Linkage disequilibrium
LH	Luteinizing hormone
LHR	Luteinizing hormone receptor
M	Swedish military conscripts
MDCS	Malmö Diet and Cancer Study

MrOS	Osteoporotic fractures in men study
NLS	Nuclear localization signal
NR	Nuclear receptor
NTD	Amino-terminal domain
OR	Odds ratio
P-box	Proximal box
PAP	Prostatic acid phosphatase
PCa	Prostate cancer
PKA	Protein kinase A
PSA	Prostate-Specific Antigen
REDUCE	Reduction by Dutasteride of Prostate Cancer Events
SAS	South Asian superpopulation of 1000 genomes
SBMA	Spinal and Bulbar Muscular Atrophy
SHBG	Sex hormone binding globulin
SNP	Single nucleotide polymorphism
SP1	Specificity factor 1
StAR	Steroidogenic Acute Regulatory
TC	Testicular Cancer
TGCC	Testicular germ cell cancer
TURP	Transurethral resection of the prostate
UTR	Untranslated region

List of papers

This thesis is based on the following original publications, referred to in the following text by their Roman numerals.

I. **Bentmar Holgersson, M.,** Giwercman, A., Bjartell, A., Wu, F. C., Huhtaniemi, I. T., O'Neill, T. W., Pendleton, N., Vanderschueren, D., Lean, M. E., Han, T. S., Finn, J. D., Kula, K., Forti, G., Casanueva, F. F., Bartfai, G., Punab, M. and Lundberg Giwercman, Y. Androgen receptor polymorphism-dependent variation in prostate-specific antigen concentrations of European men. *Cancer Epidemiol Biomarkers Prev, 2014.* ***23****(10): p. 2048-56.*

II. **Bentmar Holgersson, M.,** Ruhayel, Y., Karlsson, M., Giwercman, A., Bjartell, A., Ohlsson, C., Mellstrom, D., Ljunggren, O., Haghsheno, M. A., Damber, J. E. and Lundberg Giwercman, Y.
Lower prostate cancer risk in Swedish men with the androgen receptor E213 A-allele. *Cancer Causes Control, 2017.* ***28****(3): p. 227-233.*

III. **Bentmar Holgersson, M.,** Malm, J. and Lundberg Giwercman, Y. Serum Testosterone Levels in Early Prediction of Prostate Cancer Risk. *Eur Urol, 2017.* ***71****(6): p. 992-994.*

IV. **Bentmar Holgersson, M.,** Landgren, F., Rylander, L. and Lundberg Giwercman, Y.
Mortality Is Linked to Low Serum Testosterone Levels in Younger and Middle-aged Men. *Eur Urol, 2017.* ***71****(6): p. 991-992.*

V. **Bentmar Holgersson, M.** and Lundberg Giwercman, Y.
Genetic associations regarding prostate cancer, within and beyond the androgen receptor.
Manuscript

Reprints were made with the permission from the publishers.

Review of the literature

Part I: Androgens

Short history of testosterone

The effect of testosterone deficiency has in an indirect way been known for a long time. Castration, which in older times meant removal of the testes, was historically used in wartime to prevent the defeated group of men to carry on their lineage and is often seen as a way to humiliate the castrate. In ancient China for instance, several criminal offenses were punished by castration and thereafter enslavement, to create obedient servants, unable to have offspring and thereby unable to compete with the lineage of the emperor[1]. Castration could also be performed to create servants able to perform a specific social function, e.g. as guardians of harems or as castrate singers.

Even though the effect of the removal of the testes was well-known, the cause of the changes in the men who had been castrated wouldn't be identified until much later.

The first clue came in 1771, when the English surgeon John Hunter transplanted the testicles from a rooster into a hen, with masculinization of the hen as a result. In 1849, the experiment was repeated, but with castrated roosters (capons) acting as the transplant receivers, and results were published by Arnold A. Berthold [2].

In the late 1800s, Dr. Brown-Séquard, a physician known to perform experiments on himself in order to understand human biology became interested in the effect of castration on Eunuchs. In particular, he was interested in the behavior and health of these men, as he saw the same symptoms in older non-castrated men[3]. To investigate whether the testicles contained a substrate, potent enough to rejuvenate older men, he created an elixir containing extracts of newly removed testicles. At first he tried

the extracts on elderly dogs, and thereafter he tried it on himself with an increase in strength and energy as the result [3].

Although the results of Brown-Séquard soon were found to be most likely the effect of placebo (a recent study showed that the method of extraction he used resulted in a testosterone concentration four orders of magnitude lower than that required to give a biological effect[4]), the hypothesis behind the experiment still was valid. However, it wasn't until 1927, when McGee and Koch extracted the lipid fraction of bull testicles and injected capons with the extract, resulting in secondary sexual characteristics developing in the capons, similar to those from a non-castrated rooster. The extraction process, however, required 29 kg of bull testes to produce 20 mg of the substance[5] which deemed it impractical for isolation aimed for medical use, but pharmaceutical companies had become interested and initiated steroid research laboratories in the early 1930s.

In 1931 Butenandt was able to isolate 15 mg of an androgen from 15 000 l of urine from German policemen which was named androsteron[6]. This androgen was however not as powerful as the previously extracted substance and was later found to be one of the most common metabolites of testosterone.

Pure testosterone was first isolated and named by the Laquer laboratory in May of 1935[7], and a few months later, in September 1935, two groups independently published methods for synthesis of testosterone[8,9]. The two researchers responsible for the two latter studies, Butenandt and Ruzicka, were rewarded the Nobel Prize in medicine in 1939 for their work.

Soon after, in 1937, testosterone was introduced to the pharmaceutical market[10] and nowadays, perhaps due to increased incidence of subnormal testosterone, higher awareness or the large influence of pharmaceutical companies, men are prescribed testosterone for symptoms such as fatigue and loss of libido at an increasing rate.

Testosterone synthesis

Testosterone is synthesized in the Leydig cells of the testes and, to a smaller extent, in the adrenal cortex located above the kidneys. Synthesis starts when luteinizing hormone (LH) binds to its receptor (LHR) in the cell membrane of the Leydig cells.

The LHR is a transmembrane, G protein-coupled receptor, and upon binding of LH in the Leydig cell membrane a chain of reactions is started with the activation of adenylyl cyclase (AC; **Figure 1**). The activated AC will catalyze the conversion of ATP to cyclic adenosine monophosphate (cAMP), an important second messenger,

which triggers the activation of protein kinase A (PKA)[11]. This in turn will phosphorylate the Steroidogenic Acute Regulatory Protein (StAR) in the mitochondrial membrane, which activates it and allows it to transport the severely hydrophobic cholesterol molecule across the outer and inner mitochondrial membrane.

The transportation of cholesterol into the mitochondrion is considered to be one of the rate-limiting step in the steroidogenesis process. Inside the mitochondrion, the CYP11A1 enzyme transforms cholesterol into pregnenolone, a progestogen which after formation exits the mitochondrion and enters the smooth endoplasmic reticulum where it, through the steroidogenic cascade, is further altered until it becomes testosterone.

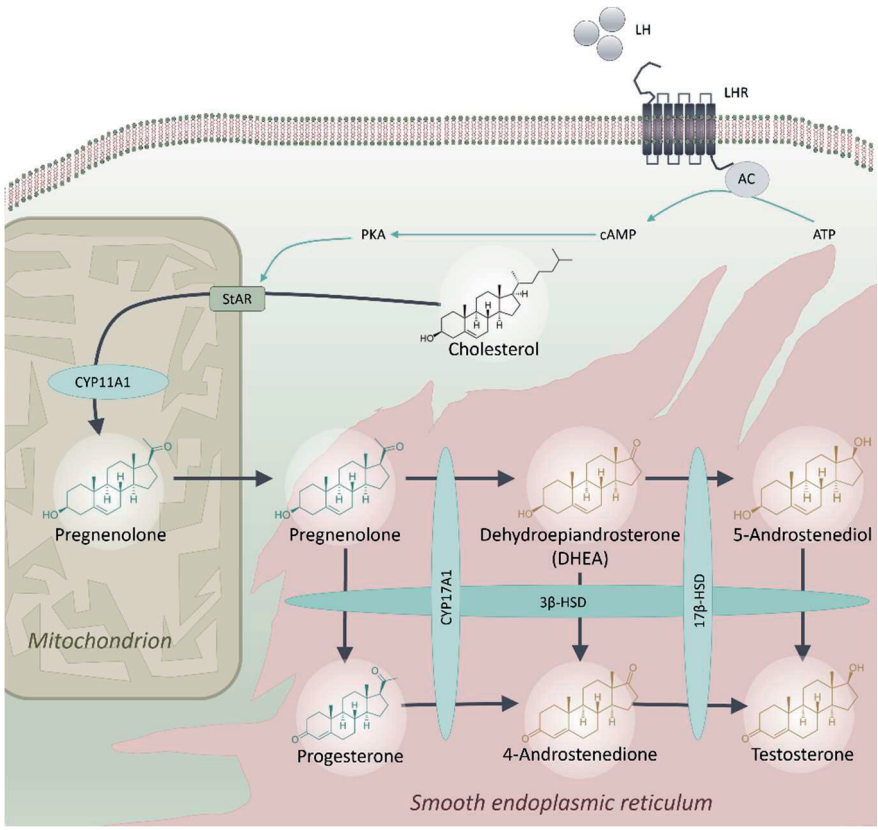


Figure 1. The mechanism behind testosterone formation after LH stimulation in the Leydig cells of the testes.

The hypothalamic-Pituitary-Gonadal (HPG) axis

The production of testosterone is strongly regulated by the Hypothalamus-Pituitary-Gonadal axis (**Figure 2**). The Hypothalamus synthesizes and releases Gonadotropin-releasing hormone (GnRH) in a pulsatile fashion, which will stimulate the anterior pituitary to release LH and follicle stimulating hormone (FSH) into the blood stream[12]. When LH and FSH reaches the testes, LH will bind to the LH receptors on the Leydig cells, stimulating testosterone synthesis, while FSH will bind to the FSH receptors on the Sertoli cells, stimulating sperm production.

Following synthesis, testosterone will diffuse out if the Leydig cells and fill up the interstitial compartment, from which it can diffuse into the blood stream and act on distant androgen responsive targets or into the closely located Sertoli cells.

The Sertoli cells are responsive for sperm production, a process that is dependent on testosterone and FSH but also release Inhibin B and Follistatin upon testosterone stimulation which, together with testosterone, will downregulate the release of GnRH, and thereby also the release of FSH and LH.

The Sertoli cells require extremely high concentrations of intratesticular testosterone[13] to stimulate full spermatogenesis and therefore also secrete androgen binding protein (ABP), a glycoprotein consisting of a slightly altered sex hormone binding globulin (SHBG), which binds androgens to stabilize a high androgen concentration in the seminiferous tubules and epididymis and thereby ensure spermatogenesis.

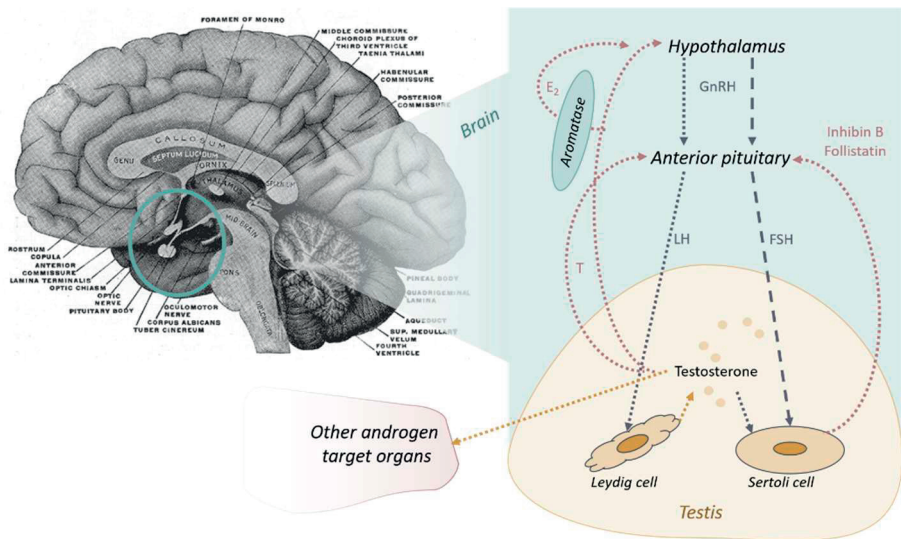


Figure 2 The regulatory pathways of the Hypothalamic-Pituitary-Testis axis. Red dotted lines indicate negative feedback.

The importance of testosterone

In males, androgens are crucial for the development and maintenance of both the primary and secondary sex characteristics and will affect the general well-being and sexual health from cradle to grave.

Upon formation, in general, only about 2% of testosterone will circulate freely, the rest will to a large extent be bound with high affinity to SHBG, deeming it basically bio-unavailable, and with a lower affinity to albumin, from which it can easily disengage and become free, bio-available testosterone.

Testosterone is capable of inducing androgenic response in peripheral target cells on its own, by binding to the androgen receptor (AR), a ligand dependent transcription factor capable of inducing expression of androgen response genes. However, much of the seen testosterone effects can also be attributed to its metabolites, 5α-dihydrotestosterone (DHT) and estradiol (E_2) (**Figure 3**).

During fetal development, testosterone is in general considered to be responsible for the development of in internal male genitalia, with the exception of the prostate gland. During puberty and adult life, testosterone stimulates skeletal muscle growth and the elongation of the larynx causing the deeper voice in males. The anabolic features of testosterone has been abused in the form of anabolic steroids, due to its ability to stimulate muscle growth and performance.

In some specific androgen target cells, e.g. in the prostate, skin and hair follicles, the enzyme 5-alpha reductase is expressed, and in these cells a large part of the testosterone will be converted to the more potent DHT, with a higher affinity for the AR and a higher capacity for induction of androgen response[13]. During fetal development, DHT is important for the development of the external male genitalia but it is also important for much of the development and function of the prostate. In older men, suffering from benign prostate hypertrophy (BPH), treatment often consists of 5-alpha reductase inhibitors such as finasteride and dutasteride, which block the conversion of testosterone to DHT and relieve the symptoms by reducing prostatic growth. The 5-alpha reductase inhibitors are also used by dermatologists to prevent and treat male pattern baldness, androgenetic alopecia, as the main factor involved in the progression of male pattern baldness is DHT.

Testosterone can also be converted into the estrogen E_2 in a reaction catalyzed by the enzyme CYP19A1. Although estrogens are generally described as the female sex hormone, E_2 plays an important role also in men[14], as it stimulates epiphyseal closure during puberty and maintains bone mass in adult life but also maintains spermatogenesis and libido. One of the E_2 target organs in males, expressing the estrogen receptors to which E_2 binds, is the prostate.

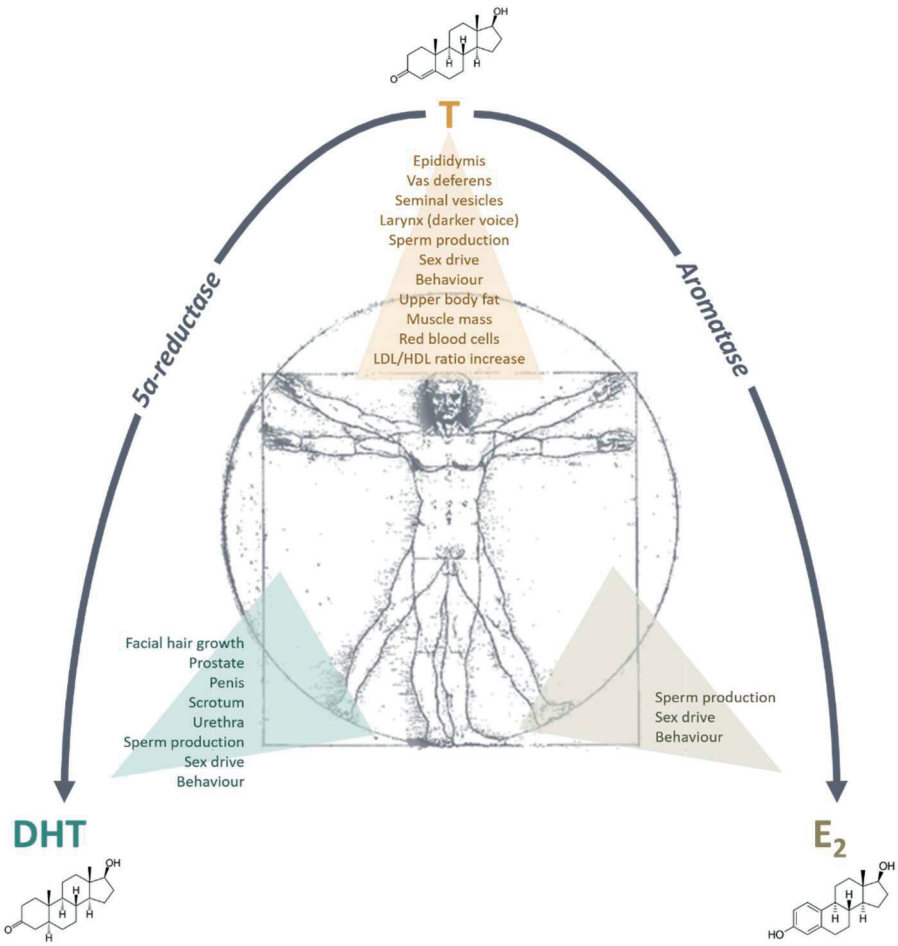


Figure 3 Effect of testosterone, dihydrotestosterone and estradiol on the development and maintenance of the male phenotype

Low testosterone

Factors affecting testosterone secretion and blood concentration is glucose load[15], diurnal variation, with highest concentrations in the morning, and an almost 50% lower concentration in the evening for younger men[16-18] and age, as testosterone levels are lower in older men[19-21]. The age-dependent decrease in testosterone in combination with the rapid growth of the older population size has made testosterone deficiency and subsequent co-morbidities a large field of interest in the medical community and testosterone measurements and prescribed testosterone replacement therapies have reached new heights in the last decade[22,23].

As an inter-individual difference in androgen sensitivity exists, dependent on for instance the concentrations of aromatase, SHBG and 5-alpha reductase, low testosterone in itself is not always problem but when the testosterone concentrations are not sufficient to carry on androgen dependent processes in the body, testosterone deficiency can become symptomatic.

Low testosterone, also called hypogonadism can be divided into several categories[24]. Primary hypogonadism, where testosterone concentrations are low and LH concentrations are high, is usually a sign that the testes are not responding efficiently to the LH-signaling[25]. A similar more recently described condition is the compensated hypogonadism, where testosterone concentrations are kept at normal or low normal levels by high concentrations of LH, These men often present as a mild hypogonadism and a less distinct clinical profile. If the hypogonadism is characterized by low testosterone as well as low LH and FSH, as in the case of secondary hypogonadism, the problem usually resides in the signaling between the hypothalamus and the pituitary.

Symptoms of hypogonadism are different depending on the severity and the age of onset but usually include sexual problems such as loss of libido, impotence, and infertility but also low muscle mass, anxiety, depression and osteoporosis[26,27].

Hypogonadism appear to not only affect quality of life but has also been associated with several conditions such as metabolic syndrome, diabetes mellitus type 2 and cardiovascular disease[28-33]. Additionally, low testosterone has been associated with an increased risk of all-cause mortality[34,35], but whether low testosterone is causing these conditions or the other way around is still not known.

Part II: The androgen receptor

Short history of the discovery of the androgen receptor and androgen receptor related diseases and disorders

As long as there have been humans, a small proportion of children have most likely been born with ambiguous genitalia. These people have historically been called hermaphrodites after the Greek god Hermaphroditus, the son of Hermes and Aphrodite in Greek mythology. According to the myth, he was a remarkably beautiful boy who attracted the love of the water nymph Salmacis, who in turn prayed to be united with him forever. When her prayers were answered by the gods, the two were transformed into one androgynous form.

Historical records on intersex people are surprisingly common, and mentions can be found both in laws and myths. Depending on cultural perception and religious beliefs, they are sometimes described as godlike with fortunetelling abilities and sometimes as monsters[36]. The law usually cover whether they should inherit as men or women and which sex they should be penalized as, should they commit crimes[37]. With 20th century medicine came surgical measures to "cure" the state of intersexuality in an attempt to avoid later gender identity confusion, an action that in recent years have been questioned[38].

At the same time, researchers had begun to investigate the scientific explanations for these traits. In 1942, Fuller Albright described the concept of peripheral hormone resistance, opening up a new field of research on tissue response to hormones[39]. In 1947, Reifenstein, a student of Albright, reported on a family with hereditary pseudohermaphrodism, which subsequentially led to the condition being named after him[40], but also made researchers interested in the genetics behind the syndrome.

Since the severity of the symptoms of Reifenstein syndrome were ill-defined, several reports on different degrees of hermaphrodism were published, giving the syndrome several names, e.g. Gilbert-Dreyfus syndrome (1957)[41], Lub's syndrome (1959)[42] and Rosewater syndrome (1965)[43].

In 1970, Lyon and Hawkes published a study were they described X-linked testicular feminization in mice[44] and 4 years later the research community debated the pattern of inheritance after Bremner published a study describing a family displaying autosomal inherited pseudohermaphrodism[45]. The same year, Wilson suggested that pseudohermaphrodism should be divided into the autosomally inherited Type 2 and the X-linked inherited Type 1[46]. Wilson also concluded that the syndromes described by Rosewater, Lubs, Gilbert-Dreyfus and Reinfenstein all were caused by the same genetic defect and belonged to the X-linked androgen

resistance type 1 and that the families did not appear to be affected by defects of androgen synthesis, but instead defect androgen action.

In 1977 the X-linked Riefenstein syndrome was renamed androgen insensitivity syndrome (AIS) after Amrhein had reported differences in cytoplasmic DHT-binding, further strengthening the hypothesis that deficiency in androgen action is the underlying cause[47]. The patients in the study could be divided into three categories; weaker DHT-binding (partial AIS), no DHT binding (complete AIS) and normal binding, where the cause of the androgen insensitivity was unknown.

In 1979, the partial AIS and complete AIS categories were complemented with a third category, when Aiman et al. described mild AIS as a cause of infertility in otherwise healthy men[48].

While AIS by now was believed to be caused by a deficiency in androgen binding, the genetic locus of the receptor for androgens was not yet known. In 1981, Migeon et al were able to narrow down the location to Xq11-Xq13[49]. Seven years later, two groups independently reported successful cloning of the AR[50,51], and the first mutations of the AR, causing AIS were described[52].

In 1989. Brown et al. reported the exact locus of the AR [53] while Lubahn et al added a description of the sequence of the intron/exon junctions within the AR [54].

The detection of the AR gene also led to the finding by La Spada et al in 1991; that Kennedy's disease, a slowly progressing muscular atrophic disease with patients often displaying symptoms of mild AIS, is caused by an increased size of the polymorphic tandem CAG repeat, located in exon 1 of the AR gene[55].

The number of published mutations in the AR or AIS patients quickly grew, and in 1994 a web based database collecting all published mutations of the AR was launched[56]. In 2012 the AR gene mutations database reported 1,209 registered mutations found in both AIS and PCa patients [57].

Nuclear receptors

Transcription factors

Transcription factors are regulators of gene expression, and are thereby crucial in all cellular processes from conception and birth, until the death of an organism. Cellular differentiation, development, DNA repair and morphogenesis are all processes that are dependent on a complex chain of responses to internal or external stimuli causing gene expression, and thereby protein synthesis, to be turned on or off (**Figure 4**).

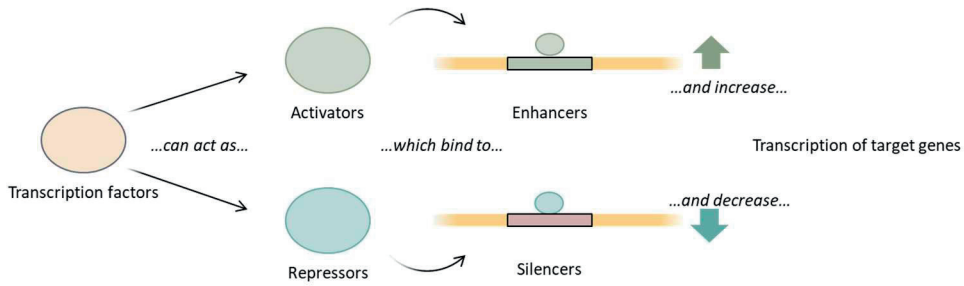


Figure 4 Schematic view of the mechanism of transcription factors.

Nuclear receptor superfamily

The AR is part of the nuclear receptor (NR) superfamily[58], one of the largest transcription factor groups. The nuclear receptor superfamily in humans contains both receptors that are activated upon binding of a specific ligand and orphan receptors, for which specific ligand have not been identified so far. In humans, 49 genes for nuclear receptors have been identified and 48 of these are expressed[59]. Of these, 20 are considered orphans with no known ligand[60].

The other 28 nuclear receptors are known to recognize certain small hydrophobic ligands in the form of endogenous hormones and vitamins or xenobiotic endocrine disruptors. In common they all (with a few exceptions) have a certain structure in the form of regulatory domains. The N-terminal domain (NTD), the DNA-binding domain (DBD), the hinge region, the Ligand binding domain (LBD). Upon ligand binding to the LBD, the receptors form homodimers or heterodimers and activate transcription by the DBD binding to hormone responsive elements of the target gene.

The structural similarities of the NRs indicate a common ancestral NR, from which all NRs have evolved[61]. The evolution of the nuclear receptors is considered to have

happened in two waves of duplications[61]. Nuclear receptors are absent in plants and fungi but are present in animals, and the first NR is considered to have appeared around the emergence of the kingdom Animalia. In the first wave of duplication, the ancient NR gave rise to precursors of the seven large subgroups of NRs. The second duplication wave, that gave rise to the multiple variants of each subgroup and is considered to have taken place with the emergence of vertebrates.

The AR belongs to NR group 3, the *Estrogen receptor-like receptors*, based on phylogenetic resemblance[62] (**Table 1**). Within this group resides three estrogen receptor related orphan receptors that are not activated by estrogen but bind to estrogen response elements, and six steroid hormone receptors divided into two groups; the estrogen receptors (or 3-hydroxysteroid receptors) and the 3-Ketosteroid receptors. The steroid receptors are believed to be the products of the same ancestral estrogen activated steroid receptor that after a duplication event gave rise to an estrogen receptor and a 3-Ketosteroid receptor. Further duplication events later formed the six, now existing nuclear steroid hormone receptors.

Like the other nuclear receptors, the steroid hormone receptors share a similar functional structure, organized by domains (**Figure 5**).

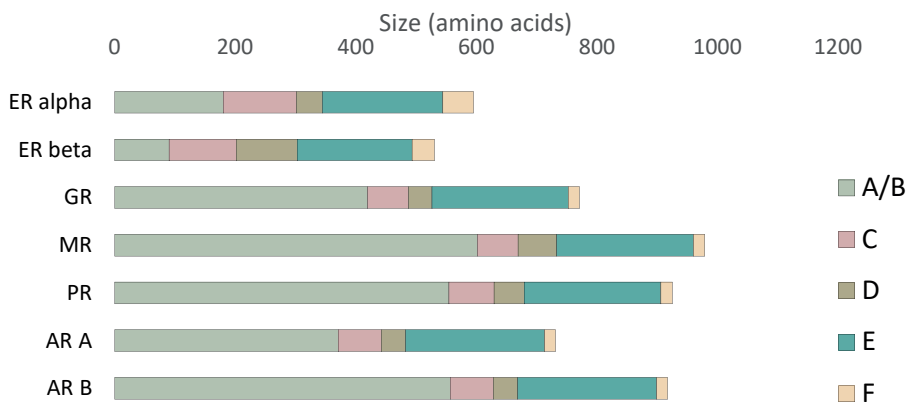


Figure 5 The difference in length of protein domains, the N-terminal (A/B), the DNA-binding (C), the hinge region (D), the ligand binding (E) and the C-terminal (F) domain in estrogen receptor alpha and beta (ER alpha, ER beta), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), progesterone receptor (PR) and androgen receptor transcript variant A (AR A) and B (AR B).

Table 1 The phylogenetic relationship between the members of the nuclear receptor family, as well as their ligands if known.

Gene	Group	Alt name	**Ligand**	**Size**
NR0B1	Miscellaneous	DAX1	*Orphan*	470
NR0B2	Miscellaneous	SHP	*Orphan*	257
NR2C1	Retinoid X Receptor-like	TR2	*Orphan*	603
NR2C2	Retinoid X Receptor-like	TR4	*Orphan*	596
NR2F1	Retinoid X Receptor-like	COUP-TFI	*Orphan*	423
NR2F2	Retinoid X Receptor-like	COUP-TFII	*Orphan*	414
NR2F6	Retinoid X Receptor-like	EAR-2	*Orphan*	404
NR2E1	Retinoid X Receptor-like	TLX	*Orphan*	385
NR2E3	Retinoid X Receptor-like	PNR	*Orphan*	410
NR2A1	Retinoid X Receptor-like	HNF4α	*Orphan*	474
NR2A2	Retinoid X Receptor-like	HNF4γ	*Orphan*	408
NR2B2	Retinoid X Receptor-like	RXRβ	Retinoic acid	533
NR2B1	Retinoid X Receptor-like	RXRα	Retinoic acid	462
NR2B3	Retinoid X Receptor-like	RXRγ	Retinoic acid	463
NR5A1	Steroidogenic Factor-like	SF1	*Orphan*	461
NR5A2	Steroidogenic Factor-like	LRH-1	*Orphan*	541
NR6A1	Germ Cell Nuclear Factor-like	GCNF	*Orphan*	480
NR1A1	Thyroid Hormone Receptor-like	TRα	Thyroid hormone	490
NR1A2	Thyroid Hormone Receptor-like	TRβ	Thyroid hormone	461
NR1B3	Thyroid Hormone Receptor-like	RARγ	Retinoic acid	454
NR1B1	Thyroid Hormone Receptor-like	RARα	Retinoic acid	462
NR1B2	Thyroid Hormone Receptor-like	RARβ	Retinoic acid	455
PPARG	Thyroid Hormone Receptor-like	PPARγ	Fatty acids, prostaglandin J2	505
NR1C1	Thyroid Hormone Receptor-like	PPARα	Fatty acids, leukotriene B4, fibrates	468
NR1C2	Thyroid Hormone Receptor-like	PPAR-β/δ	Fatty acids	441
NR1D1	Thyroid Hormone Receptor-like	Rev-ErbAα	*Orphan*	614
NR1D2	Thyroid Hormone Receptor-like	Rev-ErbAβ	*Orphan*	579
NR1F3	Thyroid Hormone Receptor-like	RORγ	Retinoic acid	518
NR1F1	Thyroid Hormone Receptor-like	RORα	Cholesterol. Cholesteryl sulphate	523
NR1F2	Thyroid Hormone Receptor-like	RORβ	Retinoic acid	470
NR1I1	Thyroid Hormone Receptor-like	VDR	1,25-dihydroxy vitamin D_3, litocholic acid	427
NR1I2	Thyroid Hormone Receptor-like	PXR	Xenobiotics, PCN	434
NR1I3	Thyroid Hormone Receptor-like	CAR	Xenobiotics, phenobarbital	352
NR1H4	Thyroid Hormone Receptor-like	FXR	Bile acids, Fexaramine	486
NR1H3	Thyroid Hormone Receptor-like	LXRα	Oxysterols, T0901317, GW3965	447
NR1H2	Thyroid Hormone Receptor-like	LXRβ	Oxysterols, T0901317, GW3965	460
NR4A1	Nerve Growth Factor IB-like	NGFIB	*Orphan*	598
NR4A2	Nerve Growth Factor IB-like	NURR1	*Orphan*	598
NR4A3	Nerve Growth Factor IB-like	NOR1	*Orphan*	626
NR3C4	Estrogen Receptor-like	AR	Testosterone, flutamide	920
NR3C3	Estrogen Receptor-like	PR	Progesterone, RU486, medroxyprogesterone acetate	933
NR3C1	Estrogen Receptor-like	GR	Cortisol, dexamethasone, RU486	777
NR3C2	Estrogen Receptor-like	MR	Aldosterone, spirolactone	984
NR3B1	Estrogen Receptor-like	ERRα	*Orphan*	423
NR3B2	Estrogen Receptor-like	ERRβ	DES, 4-OH tamoxifen	433
NR3B3	Estrogen Receptor-like	ERRγ	DES, 4-OH tamoxifen	458
NR3A1	Estrogen Receptor-like	ERα	Oestradiol-17β, tamoxifen, raloxifene	595
NR3A2	Estrogen Receptor-like	ERβ	Oestradiol-17β, various synthetic compounds	530

The androgen receptor gene

The androgen receptor gene

The AR gene consists of eight exons spanning over more than 90 kb[54,63] (**Figure 6**). It is the only steroid receptor located on the X-chromosome (Xq11-12) leaving the typical male karyotype (46XY) hemizygous for the AR. The genomic region where the AR resides is highly conserved between species[64] and has also been found to be the most divergent genomic segment between African and East-Asian populations[65] with a high frequency of derived alleles found in the African populations[66]. In humans, the X-chromosome has been suggested to have experienced an accelerated genetic drift post dispersal from Africa[67] with strong signals of positive selection in the genomic segment where the AR is located[68,69].

Several mutations in the AR have been described in AIS patients[57]. While the female carriers of these mutations have unaltered phenotype, the male carriers, who only carry one X-chromosome will be affected. These mutations alter the protein structure and function and thereby deem the male patients unresponsive to androgenic signaling, leading to symptoms ranging from mild AIS, with patients presenting with somewhat impaired spermatogenesis and reduced development of secondary sexual characteristics, to complete AIS, with patients presenting with female habitus but with absent ovaries[70].

The large first exon of the AR encodes the entirety of the N-terminal domain[63] and holds three common polymorphisms, the CAG-repeat, the GGN-repeat and the single nucleotide polymorphism (SNP) rs6152.

The CAG repeat

The CAG-repeat is as the name suggests, a long stretch of the bases C, A, and G repeated tandemly, encoding for a chain of glutamines, and the number of repeats is variable in humans[71], but also in other primates[72-74]. In humans, the average lengths and the range of the CAG-alleles differ between populations (**Figure 7**)[71,75,76] with a normal range between 10 and 30 repeats.

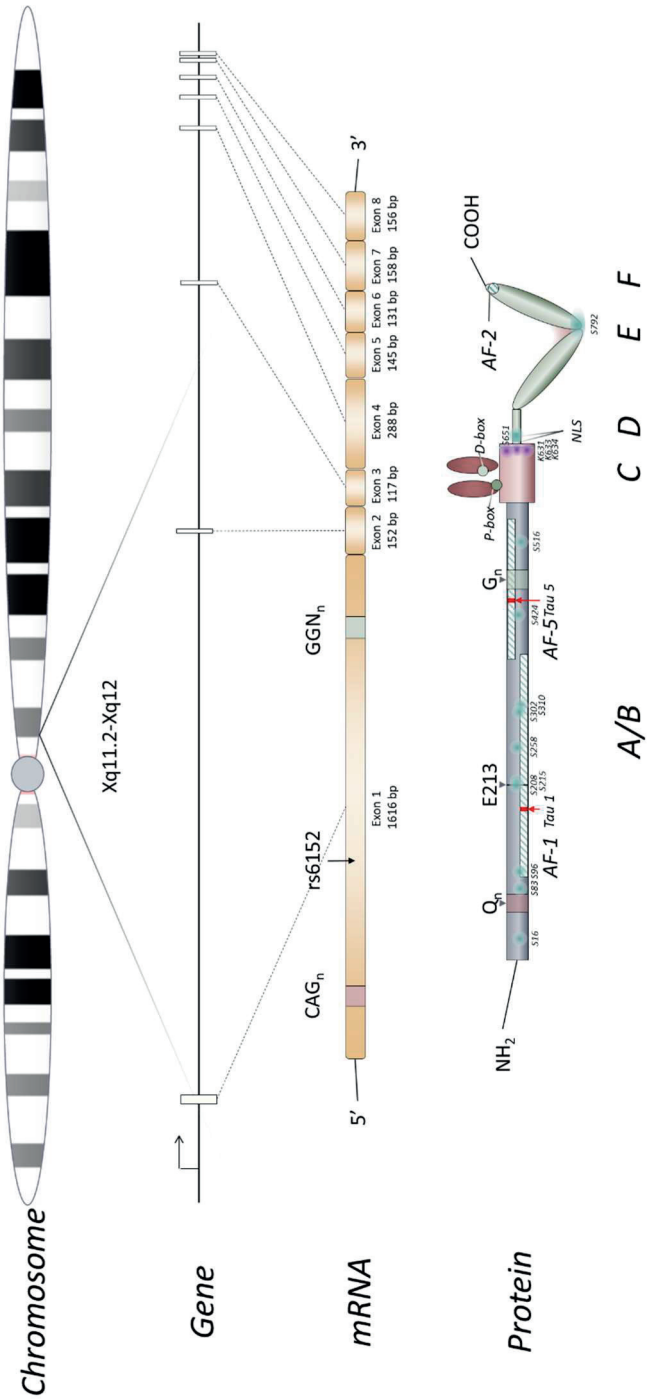


Figure 6 the structure of the AR at gene level, mRNA level and protein level. A/B at protein level corresponds to the NTD, C corresponds to the DBD, D corresponds to the hinge region and E+F corresponds to the LBD.

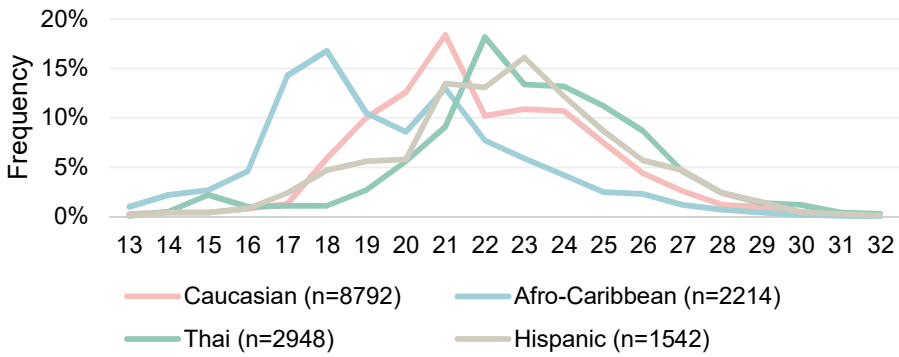


Figure 7 CAG-repeat length distribution in different populations (Adapted from Ackerman, C. M. et al 2012)

Longer repeats (>40) is known to cause the rare late-onset progressive motor-neuron disease Spinal and Bulbar Muscular Atrophy (SBMA) also called Kennedy's disease[77,78]. The length of the CAG-repeat is correlated with the severity of the disease and negatively correlated with the age of onset[79]. SBMA patients often also display endocrine symptoms of mild AIS such as subfertility, erectile dysfunction and gynecomastia[80].

Although the neurotoxic function of the AR harboring extreme CAG-lengths appears to be due to AR aggregation the AIS-symptoms of SBMA patients led to the theory that CAG-repeats length also in normal ranges is inversely correlated to androgen sensitivity, with the main focus often being aimed at the transactivation capacity of the AR[81-83] and association studies regarding CAG-repeat length and risk of a long range of conditions[84] such as infertility[85-88]and testicular cancer[89,90] but also sex hormone concentrations[91,92].

Also, as African populations display shorter CAG-repeat alleles and African-American men have the highest PCa incidence and mortality, in combination with the assumed inverse linear association between CAG length and androgen sensitivity, the CAG-repeat length in relation to PCa risk has been studied thoroughly with various results (for details regarding PCa risk and AR-variants, see page 29).

The GGN repeat

The second repeat polymorphism in the N-terminal domain is located downstream of the CAG-repeat and is called the GGN-repeat (**Figure 6**). This repeat is more complex in its structure with the consensus sequence $(GGT)_3GGG(GGT)_2(GGC)_n$ encoding for a glycine stretch[93]. As with the CAG-repeat, also the GGN-repeat length differs between human populations although the normal alleles are of much less variable lengths, with the most common alleles (23 and 24) covering 84% in

Sweden[76,94]. Association studies regarding the GGN-repeat are not at all as many as for the CAG-repeat but also these studies have associated the different lengths with reproductive parameters[95-97]

rs6152

Located in between the two repeat polymorphisms in exon 1 is the SNP rs6152, a synonymous variant encoding for glutamic acid (E213). The frequency of the two alleles of rs6152 differs significantly between world populations, where 100% of the members of the East-Asian (EAS) populations in 1000 genomes are carriers of the G-allele, while only 35% of African (AFR) populations are G-allele carriers (**Figure 8**). European (EUR) populations place in between with a G-allele frequency of 86%. Although this variant does not change the protein (**Figure 9**), the variant has been associated with differences in sex-hormone levels[98,99] and the G-allele is one of the strongest genetic risk markers for androgenetic alopecia in Caucasians[69,100-103].

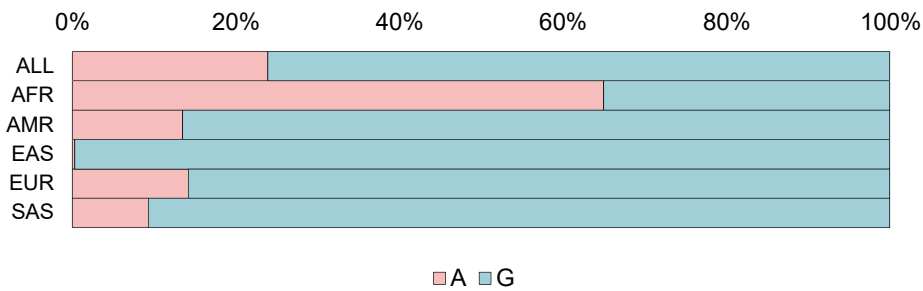


Figure 8 The proportion or rs6152 A- and G-allele carriers in the different 1000 genomes superpopulations

Figure 9 The synonymous SNP rs6152 does not affect the amino acid sequence of the translated protein

Genetic variants of the androgen receptor and prostate cancer

All the three genetic variants of the AR have been investigated in relation to risk of PCa with conflicting results.

The CAG repeat is most often studied with the presumption of linear relationship, with shorter alleles being tested against longer alleles, where the cut-off differs between studies. A compilation of odds ratios (OR) and 95% confidence intervals (CI) from Caucasian studies with the CAG-repeats defined as short <22 and long ≥22[104-126] derived from a large recent meta study[127] can be seen in (**Figure 10**). This meta study concluded, upon inclusion of studies also including non-Caucasian subjects that men carrying CAG-repeat lengths <22 had an elevated risk of developing PCa. The same meta study also investigated the PCa risk in carriers of short ≤16 GGC-repeats, corresponding to ≤22 GGN-repeats in relation to carriers of long (>22 GGN) and the pooled analysis revealed an increased risk for carriers of short GGN-repeats. The ORs for shorter GGN-repeat lengths in relation to PCa risk for the Caucasian populations[106-109,113,121,123,128] derived from the meta study can be seen in **Figure 11**.

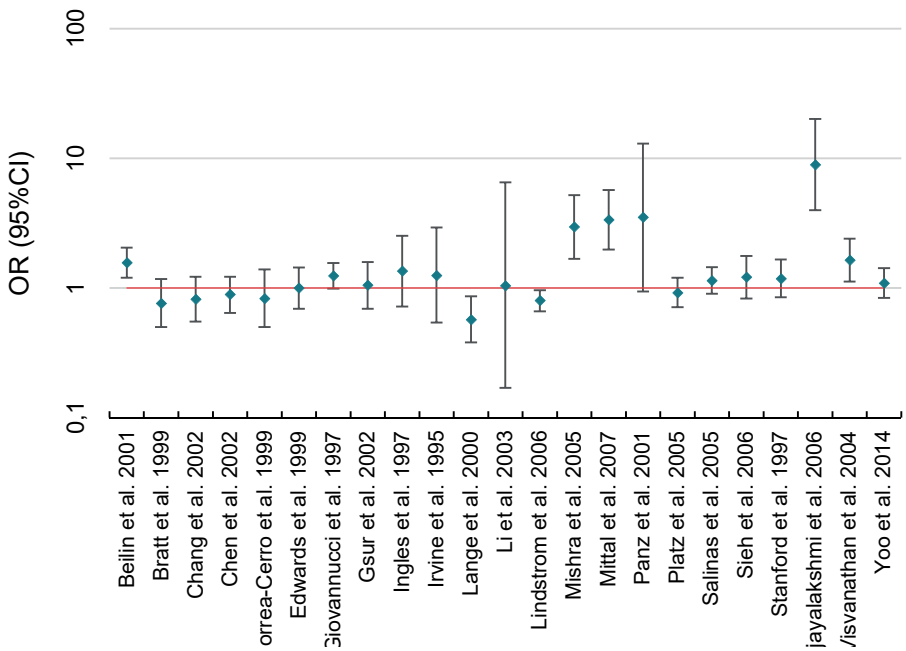


Figure 10. Odds ratios and 95% confidence intervals for PCa for carriers of short (<22) CAG repeats compared with carriers of CAG repeat lengths ≥22 (as indicated by the red line). Data derived from Weng, H. et. al. 2017

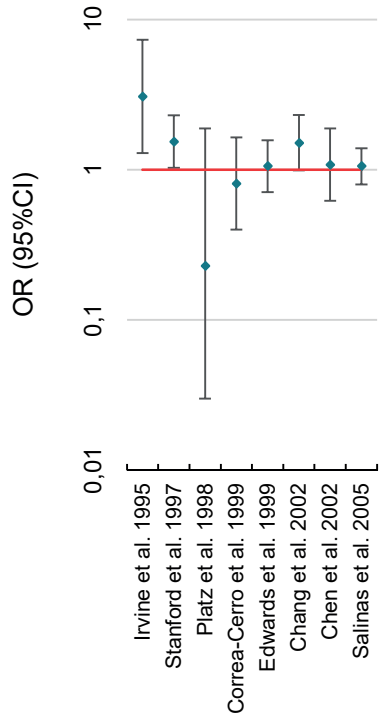


Figure 11 Odds ratios and 95% confidence intervals for PCa for carriers of short (≤22) GGN repeats compared with carriers of GGN repeat lengths >22 (as indicated by the red line). Data derived from Weng, H. et. al. 2017

Reports on AR SNPs have also been conflicting. For rs6152 some studies have found an increased risk or PCa for the G-allele[129,130] whereas others have not[131,132].

Transcription of the AR

The AR mRNA is expressed in several different tissues, but has highest expression in the liver, prostate and testes (https://www.ncbi.nlm.nih.gov/gene/367).

The AR promoter region lacks TATA and CCAAT boxes that are generally recognized by the transcription machinery and instead appears to attract the transcription factor specificity factor 1 (SP1) to GC-rich regions in at least two transcription initiation sites located ~1.1 kb upstream of the transcription initiation codon[133-136]. The expression is regulated by androgens, but the complexity of the regulation is not fully understood as both up- and down-regulation has been reported in different cell-types in the presence of androgens[137]. In addition, the AR protein translation can also be regulated by the additional factor of mRNA-stability. In

LNCaP cells for instance, the presence of androgens suppressed the mRNA transcription, but increased the half-life of the mRNA which subsequently led to a stable AR protein abundance[138].

The AR mRNA has long 3'- and 5'-untranslated region (UTR) sequences, harboring sequence elements that appear to interact with RNA-binding proteins and affect the turn over-rate of the mRNA[139]. Part of the differences ascribed to the CAG and GGN repeat lengths has also been hypothesized to be due to changes in mRNA stability, as for instance the stability of the CAG-hairpin formation is altered by the number of repeats[139-144].

The androgen receptor protein

The AR gene encodes an approximately 2760 bp open reading frame which is translated into a ~920aa long protein with an approximate mass of 100 kDa[54,63]. The length of the protein varies slightly depending on the two variable repetitive regions in exon 1. The protein consists of four functional domains, NTD, the DBD, the hinge region and the LBD (**Figure 6**).

The AR in its unliganded state is primarily located in the cytoplasm in a monomeric form, where it is associated with heat-shock proteins (HSP) and other chaperone proteins, which stabilize the protein in a conformation that allows for ligand binding[145,146]. Several post-translational modification sites are present in the AR (**Figure 6**), and regulation occurs by the action of for instance phosphorylation, methylation, acetylation and sumoylation[147].

The amino terminal domain

The NTD, or the transactivating domain, is the largest of the protein domains of the AR, and the domain that differs most from the NTDs of the other steroid receptors[148,149]. This domain harbors two areas responsible for transactivation function, activation function (AF) 1 and AF-5, where AF-1 has the strongest transactivation properties and is the main factor of ligand-dependent transcription activity[150-152].

A motif within the NTD, ^{23}FQNLF^{27}, is conserved across several species, and is thought to be essential in the activation of the AR upon ligand binding, as it interacts with the LBD but the exact mechanics of the interaction is not completely elucidated[149,150,153-155].

The DNA-binding domain

The DBD of the AR has the organization of two cysteine rich zinc-finger motifs, both of which consist of four cysteines binding a zinc-ion, and is encoded by exon 2 and 3 of the AR gene[54,156]. The most N-terminal zinc-finger includes a sequence element that is identical to the corresponding element in the other 3-ketosteroid receptors, called the proximal box (P-box). The P-box consists of a sequence of five amino acids and enables interaction between the AR and specific DNA-segments, androgen response elements (ARE), in promoters or enhancers of genes. The other zinc-finger has a distal box (D-box) which recognizes the D-boxes of another monomeric AR and enables dimerization of the two proteins. This dimerization reconfigures the protein so that the P-box is able to bind to AREs in the DNA of target genes (**Figure 12**).

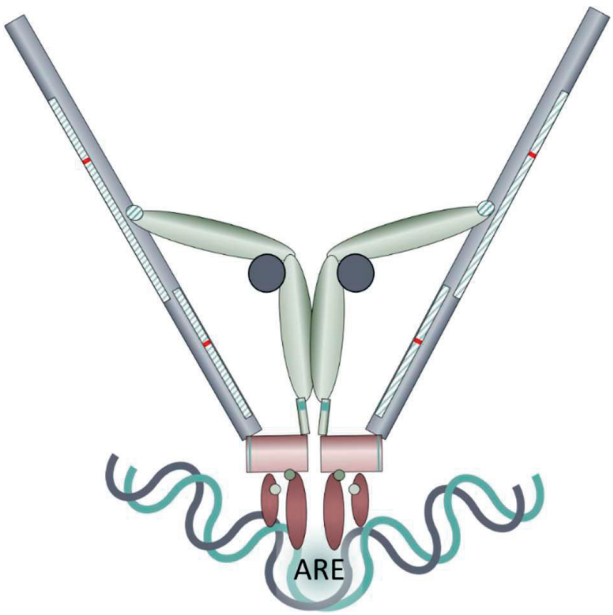


Figure 12 The binding of the androgen receptor to an androgen responsive element (ARE) is enabled by the dimerization of two ARs

The hinge region

The hinge region of the AR is a short stretch that separates the DBD and the LBD. It holds the nuclear localization signal (NLS) that is necessary for binding to the importin alpha, which mediates the transportation of the AR into the nucleus from the cytoplasm[154,157] (**Figure 13**)

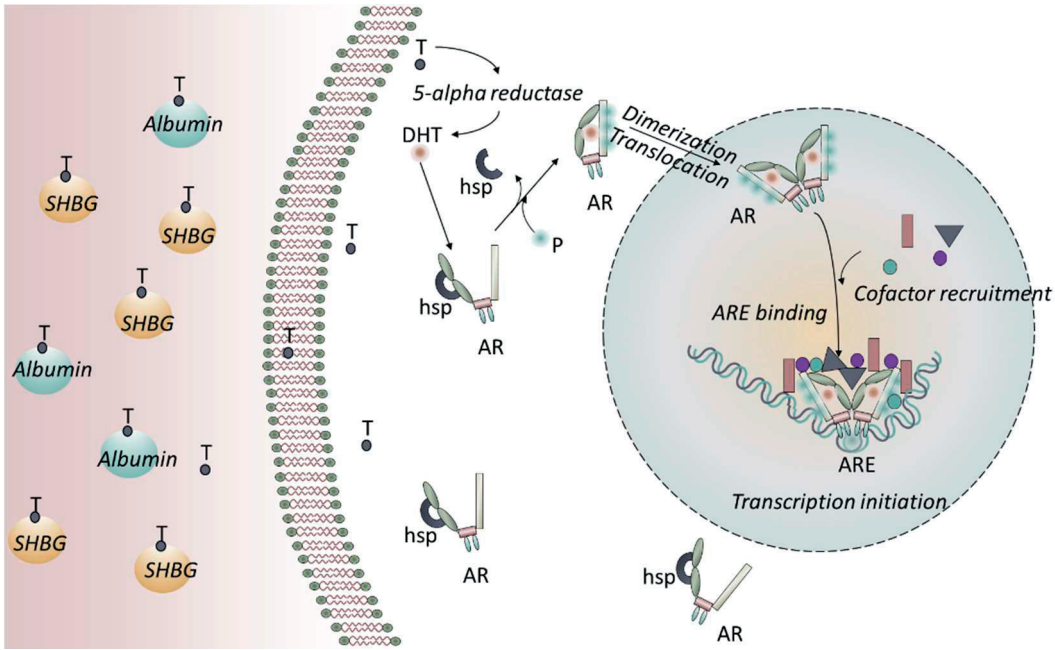


Figure 13 Schematic view of the ligand activation of AR, followed by dimerization and translocation into the nucleus, where binding to AREs of androgen response genes takes place, cofactors are recruited and transcription is initiated.

The ligand binding domain

The LBD of the AR consists of 12 alpha-helices which form a ligand-binding pocket by folding into an alpha helical sandwich[148,153]. When a ligand binds to the LBD of the AR, helix 12 folds over the pocket and encloses the ligand, resulting in a conformational change in the protein, which exposes a hydrophobic cleft, the AF-2, allowing AF-2 to bind to amino acid sequences, such as the previously mentioned ^{23}FQNLF^{27} motif of the NTD[158,159].

In summary, upon ligand binding the AR protein changes conformation, allowing dimerization, and thereby releases the chaperone proteins that kept the structure open for ligand interaction. In the process, the protein gets phosphorylated, dimerized and thereafter transported into the nucleus, where it attaches to an ARE. Upon binding to the ARE, several cofactors are recruited and transcription is initiated (**Figure 13**).

Part III: Prostate cancer

Short history of prostate cancer

While PCa today is the most common cancer in Swedish men and the leading cause of cancer related death in these men[160], it was historically considered a very rare disease and first described by the English physician J. Adams in 1853[161]. The absence of PCa descriptions in historical records and the rapid increase of PCa incidence in the past century has led to a common belief that PCa is a modern phenomenon caused by modern dietary and lifestyle factors.

While evidence of PCa in ancient remains are hard to detect, unless the tumor had advanced into skeletal metastases, paleopathologists have identified signs of PCa in 2000 years old skeletal remains of a cremated man[162], of a Scythian king from Siberia[163] and in a man from the Roman period found in Hungary[164] but also in an Egyptian mummy indicating that PCa is not a modern man-made disease[165].

The absence of descriptions of the disease before 1853 is partly explained by the use of indistinct terminology[166]. PCa was rarely distinguished from BPH as both conditions presented with the same symptoms, namely difficulties to urinate as the enlarged prostate pressed against the bladder or the urethra. The treatments of the patients diagnosed with prostate hypertrophy were mainly symptom relieving regardless of the malignancy of the hypertrophy.

PCa is known to be an old man's disease, affecting mainly men aged 55 or older. The rapid increase of PCa incidence in the Western world can therefore probably to a large extent be explained by the increased life expectancy, but also by increased use of diagnostic markers.

The first diagnostic marker for PCa was presented in 1938, when it was reported that elevated concentrations of acid phosphatase could be detected in the serum of patients with metastasized PCa[167]. A few years later, in 1941 , Huggins and Hodges proved that androgen suppressing therapy by orchiectomy or estrogen injections led to PCa regression, validated by measurements of the serum levels of acid phosphatase[168]. The report by Huggins led to one of the most important methods to treat PCa patients, the androgen deprivation therapy, and in 1966 Huggins was awarded the Nobel Prize in physiology.

The hunt for diagnostic markers with both high sensitivity and specificity for PCa was, however, not over with the clinical introduction of acid phosphatase. Elevated concentrations of acid phosphatase with prostatic origin is often seen also in benign prostatic conditions[169] and at first the test was not specific for prostatic acid phosphatase (PAP) but with a few modifications to the measuring methods to

increase specificity, PAP was the best available marker of PCa for several decades to come.

With the emergence of new biochemical methods, antigens were discovered at a high rate, and several different research groups independently discovered an antigen specific for the prostate in the late 1960's and 1970's. When a correlation between the antigen amount in blood and the concentration in the prostate was described, a new PCa marker was born [170,171]

The healthy prostate

The prostate is an exocrine gland which in humans have the shape and size of a walnut. It is located below the bladder and surrounds the urethra, acting as the junction between the urethra and the ejaculatory ducts (**Figure 14**). The development, growth and function of the prostate is highly dependent on androgens, most importantly DHT, and the absence of androgens quickly leads to a reduction in prostate size[172].

The prostate is a reproductive organ with the main function to secrete prostatic fluid into the seminal fluid upon ejaculation. The prostatic fluid makes up around 20% of the ejaculate and contains several compounds, for instance citric acid, zinc, prostatic acid phosphatase, electrolytes such as K^+ and Na^+, and PSA[173].

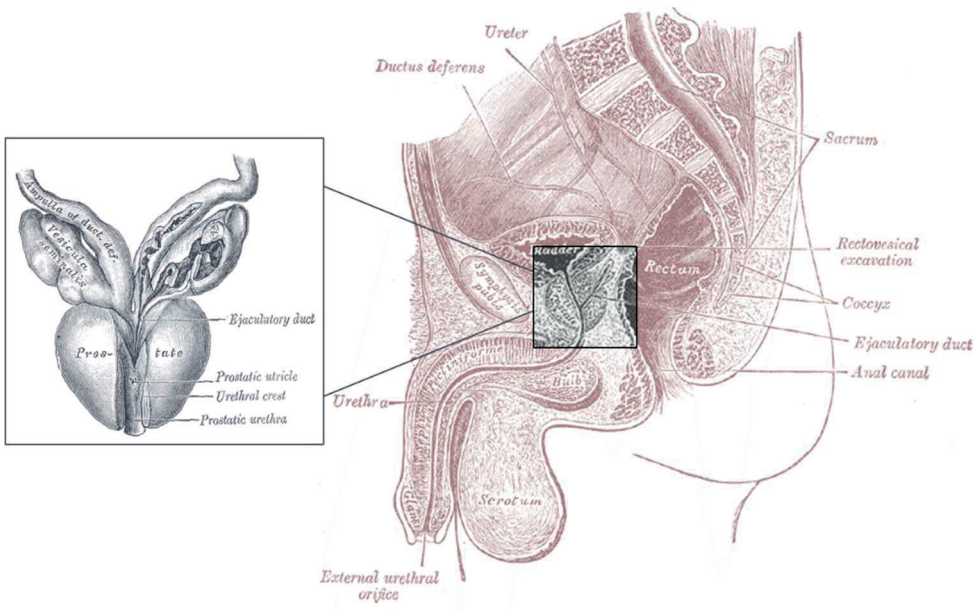


Figure 14. The location and appearance of the prostate gland

The prostatic fluid is important for *in vivo* fertilization[174]. The seminal vesicles produce coagulation factors which allow the semen to become gelatinous soon after ejaculation, restricting the movement of spermatozoa. The enzymatic properties of the prostatic fluid allows the seminal coagulate to slowly liquefy, and allow for optimized spermatozoa exposure to factors enhancing their motility and survival within the female genital react, increasing the chance of fertilization. The smooth muscle contraction of the prostate also helps eject the semen.

The prostate is made up of ~70% glandular tissue and ~30% stromal tissue and can be divided into zones, based on morphology, pathology and function (**Figure 15**)[175,176]. The peripheral zone makes up ~70% of the glandular volume and is located in the base of the gland, allowing it to be felt though digital rectal examination. The embryologic origin of this zone is the urogenital sinus. Approximately 75% of prostate tumors originate from the peripheral zone.

The transition zone is also derived from the urogenital sinus. It surrounds the urethra as it enters the prostate from the bladder and makes up only ~5% of the prostate volume. With age, this part of the prostate often enlarges and makes the passage of urine from the bladder through the prostate harder, a condition called BPH[177]. However, roughly 20% of PCa originate in the transition zone.

The central zone surrounds the ejaculatory ducts leading from the two seminal vesicles into the urethra and makes up ~25% of the prostate volume. This zone differs from the transitional and peripheral zones in its embryonic origin as it is derived from the Wollfian ducts, but also in its lower proportion (~5%) of PCa originating in this zone.

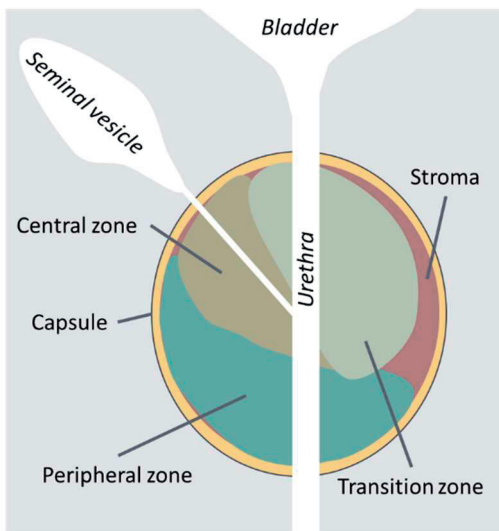


Figure 15. Schematic view of the zones of the prostate

Prostate cancer

Androgens and the AR are not only essential for normal prostate growth but also to some extent for the growth of PCa. Since the report by Huggins and Hodges, where androgen ablation was found to shrink the PCa, it has been one of the main methods to treat PCa that cannot be cured by radiation or surgery alone although the pathway that the medications target have been refined since the 1940's[168,178]. While estrogens and orchiectomy previously were the most common way to block the androgen production, today a large variety of GnRH analogs, GnRH antagonists and antiandrogens can be found on the market and can be used in different combinations to shrink the tumor and treat PCa[140].

However, PCa often becomes castration resistant (CRPC), deeming it unresponsive to androgen ablation and able to progress into lethal disease, illustrating the complexity of PCa progression and the difficulties in PCa treatment. The mechanisms behind CRPC is not fully understood, although some insights into the progression has revealed a continued AR activity albeit ligand independent, and several truncated AR variants, missing parts of the LBD, have been found in PCa[179].

The importance of AR and androgens in PCa has sometimes been interpreted as if higher testosterone concentrations could be a risk factor for PCa.

Testosterone concentration and prostate cancer

The relationship between testosterone concentrations and PCa risk is a complicated field of research with many studies reporting contradictory results.

While an activated AR *in vitro* appears to enhance proliferation in stromal cells, and promote PCa progression in epithelial luminar cells it appears to inhibit metastasis in basal cells[180]. Many PCa cell-lines respond to androgens and androgen ablation send them into programmed cell death[181-183]. It has been suggested that a subpopulation of androgen independent tumor cells, for instance malignant epithelial stem-cells, are present in the prostate at early disease, and that they upon androgen withdrawal and subsequent cell death of other PCa cells, are able to grow into CRPC[184-186].

Although it is scientifically established that PCa progression initially appears androgen dependent, the concentration of testosterone as a risk factor for PCa appears difficult to elucidate, as only few longitudinal studies on the subject exist. Population based studies have shown an increased PCa risk both in men with higher testosterone[187-189] and in men with lower testosterone[190]. Other studies report no correlation between testosterone concentrations and PCa risk[191-193].

However, autopsy studies have shown a large number of clinically indolent PCa in men who had died of unrelated causes[194-196], and although testosterone concentrations seem irrelevant in the *de novo* tumorigenesis it is possible that a certain amount of androgens are needed for the tumor to grow. The ”saturation model” has been suggested to describe the androgen sensitivity of the PCa, where extremely low androgen levels are enought to saturate the prostatic AR, and androgen concentrations above the saturation level will not lead to any additional growth[197].

Risk factors, incidence and mortality

PCa incidence and mortality increased steadily during the 20th century, and Sweden is no exception (**Figure 17** and **Figure 16**). A large part can be attributed to the increased life expectancy in humans. Age is one of the strongest risk factors for PCa, where the disease is rare in men younger than 50, after which the risk will increase quickly[198]. Additionally, the use of transurethral resection of the prostate (TURP) for men with BPH increased the number of spontaneous PCa discoveries and the largest incidence increase occurred after the introduction of the PSA-test[198].

Family history has also been strongly implicated in individual risk of PCa with the highest risk in men with relatives suffering from early-onset disease and men with more than one affected relative[199,200]. Part of this increased risk is due to one of several mutations identified in genes such as BRCA1 and 2[201] but the rarity of identified high-penetrance gene variants point to other factors or an additive effect of several susceptibility loci being more important in the family history risk increase[202].

Another important factor is ethnicity, where African-American men have both higher incidence and higher mortality rate compared to other American populations (https://www.cdc.gov/cancer/prostate/statistics/race.htm), and the lowest risk seen in Asian countries[203].

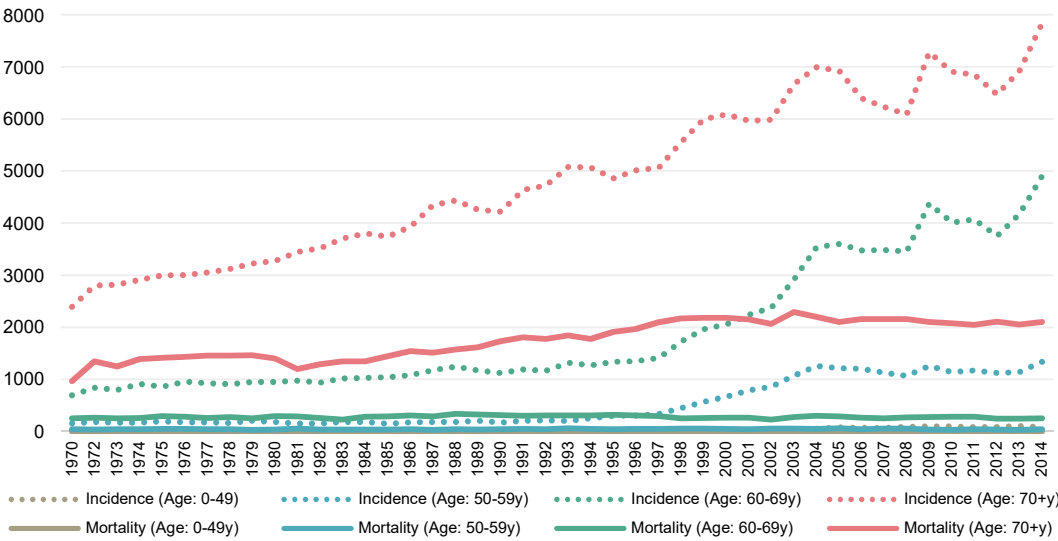


Figure 17 Crude PCa incidence and PCa mortality in Sweden divided into age categories (Data source: Cause of death & cancer registy. Stockholm: Socialstyrelsen. [16 december 2016]. http://www.socialstyrelsen.se/Statistik/statistikdatabas/)

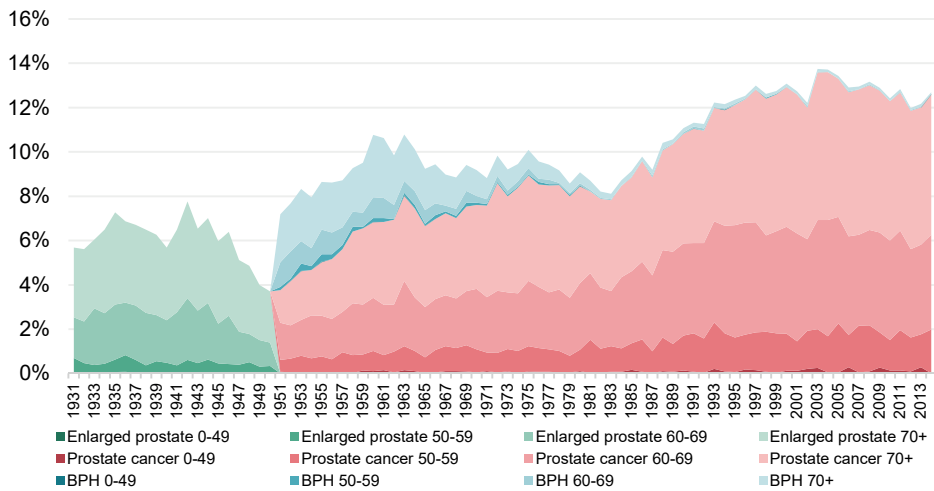


Figure 16 The proportion of prostate related causes of deaths, within all cause of death diagnoses, separated into age-categories (Data source: Dödsorsaker. Stockholm: Socialstyrelsen. [16 december 2016]. http://www.socialstyrelsen.se/Statistik/statistikdatabas/)

Prostate specific antigen

The PSA, also known as kallikrein-3 (KLK3), is an enzyme mainly secreted by the epithelial cells of the prostate gland, but low concentrations of the enzyme have also been detected in other tissues[204]. At least three AREs have been identified in the KLK3 gene promoter[205-208], and the expression of PSA is stimulated by the presence of androgens and the AR, making PSA a useful tool for monitoring of advanced PCa cases treated with androgen ablation. Initially, as the androgen concentration drops, so does the PSA concentration. However, if the PCa goes into CRPC, the PSA concentration often starts to rise again, as the AR activity no longer is ligand dependent[209,210].

The PSA is not in itself a tumor-specific enzyme. In the prostate it is present in an inactive form, but upon ejaculation it is cleaved into its active form by KLK2, another member of the same family. The activated PSA in turn liquefies the gelatinous semen matrix by cleaving the matrix upholding seminogelin proteins into smaller peptides, slowly releases the spermatozoa[211,212].

When the epithelial cells of the prostate are disrupted, due to inflammation or PCa, PSA leaks into the blood stream, making serum PSA concentrations a valuable tool in the diagnosis of prostatic disease[213-215].

Although PSA has been found to be a better prognostic tool for PCa than the previously utilized PAP-test[216-218], it is not a PCa-specific marker and large screening programs have been criticized as not all men with PCa are found based on their serum PSA while many insignificant tumors are discovered[219-221]. The diagnosis and treatment of these tumors, which might have never grown enough to become a problem for the patient, is a large problem[222,223] but should be weighed against the PCa mortality reduction in patients with aggressive non symptomatic tumors discovered through PSA-screening[224].

In a twin-study, 45% of the total PSA variability could be explained by inherited factors[225] and intra-individual PSA fluctuations have been reported[226,227] which has led to several studies regarding screening with other PSA thresholds[228] based on for instance age[229] or incorporation of other markers to increase the specificity of the test[230].

Aims

Although the role of testosterone in the growth of the PCa was well-established at the start of this thesis project, the androgen hypothesis; stating a role of testosterone in the etiology of the cancer, had begun to be questioned. Most studies on the subject were however conducted on older men with a short time to follow-up. At the same time, several reports regarding the association between low testosterone and mortality had been published but as the time to follow-up also in these studies were relatively short, the direction of the association was not known. Is severe illness a risk factor for low testosterone, or is low testosterone a risk factor for severe illness?

Genetic markers of the AR had also been studied extensively in relation to PCa, but the study designs often differed and the results often were inconclusive or contradicting each other making researchers questioning the role of genetic variants of the AR in the role of PCa[231] and suggesting a paradigm shift in the interpretation of the results[232]. Also, although genetic variants of the AR is believed to have different transactivation capacities, the role of AR in the concentrations of PSA, an androgen induced gene, in the serum of men without PCa was not elucidated.

Our general object of this work was to elucidate the combinatory inheritance of genetic variants of the AR, and to investigate how these markers could modulate the androgenic response to testosterone, or the risk of PCa in European men. Our hypothesis was that as genetic variants of the AR appear to modulate the transcription of PSA *in vitro*, they could also modulate the expression *in vivo* (**Figure 18**). Additionally, we wanted to investigate possible associations between genetic variants of the AR in relation to PCa with the combinatory effect of the genetic markers in mind. Finally, as testosterone, which act through the AR, has been suggested to have a role in the etiology of PCa, we wanted to investigate whether high testosterone concentrations were a risk factor of PCa, and if low testosterone was a risk factor for all-cause mortality.

In summary, our specific aims were to investigate:

1) The association between genetic variants of the AR and PSA concentrations in men without PCa

2) The association between genetic variants of the AR and PCa risk

3) The association between lifetime exposure to different testosterone concentrations and risk of PCa in later life

4) The association between subnormal and supranormal testosterone concentrations and risk of all-cause mortality

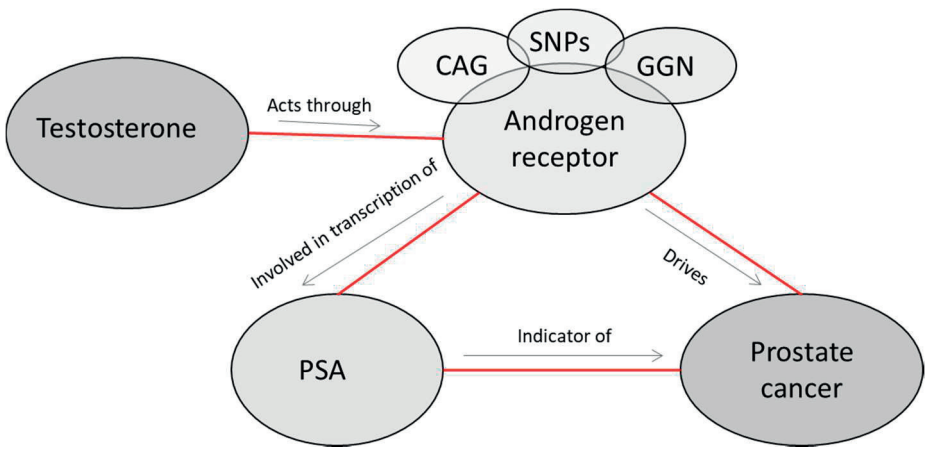


Figure 18 Mindmap describing how the testosterone concentrations and variants of the AR could modulate the risk of PCa

Materials and Methods

Subjects for genetic associations

To investigate genetic variants of the AR in relation to risk of PCa and concentrations of PSA in men without PCa, the following cohorts were used. (A more detailed introduction of these men can be found in the materials and methods section of study I, II and V).

European Male Ageing Study (EMAS)

The European Male Ageing Study was initiated in 2002 with the aim to study the ageing process in men by documenting hormonal status and symptoms related to ageing[233]. In the period between 2003 and 2005, eight European centers collected baseline information from 3369 men, aged 40-79 years, belonging to the general population. After a median of 4.3 years after enrollment, a postal questionnaire was sent out where the men would self-assess their health. Blood samples were collected both at baseline and at follow-up and amongst other clinical compounds, PSA was measured. As one of the aims of the study where EMAS was used was to investigate whether AR-variants were associated with concentrations of PSA in men without PCa, the presence of baseline PSA measured regardless of any symptoms of prostatic disease was the main motive for the selection of this cohort. All self-reported PCa cases, both prevalent and incident, were excluded for PSA calculations, as it is likely that they have undergone treatments affecting their testosterone concentrations and subsequently their PSA concentrations. Thereafter the risk of having a PSA concentration above the clinically utilized thresholds of 3 or 4 ng/ml for men carrying AR gene variants, was calculated. The information regarding prostatic disease was collected at follow-up, assuring the absence of disease influencing the PSA concentrations at baseline.

The PCa cases were thereafter included again, and used to investigate the association between AR-variants and risk of PCa (**Figure 19**).

Malmö Diet and Cancer Study (MDCS) nested case-control subset

The Malmö Diet and Cancer Study was initiated by the Swedish Cancer Society (Cancerfonden) in the early 1990's, as a part of the European Prospective Investigation into Diet and Cancer (EPIC), with the aim to study dietary and lifestyle factors in relation to cancer. In the period 1991-1995, n=42 624 women born between 1923 and 1950 and n= 31 514 men born between 1923 and 1945, all residents of Malmö, Sweden were invited to participate, and of these ~40 % were recruited to the study[234].

In 2006, a nested case-control study regarding subfertility and PCa risk was performed on the MDCS material[235]. In this study, all prevalent PCa cases still alive as of 2006 (n=661) were identified and matched to controls (n=661) based on age (±90 days) and date of enrollment (±90 days) resulting in a subset of n=1322 eligible men. These men were contacted, and those who agreed to participate (n=975), were asked to fill out a questionnaire regarding their fertility.

In order to validate the findings from study I, regarding AR-variants and PCa risk, this subset of the MDCS, consisting of all living PCa cases as of 2006 identified through the Swedish National Cancer Registry was utilized. The inclusion criteria for the nested case-control study, including only men alive at PCa data retrieval was similar to that of EMAS (**Figure 19**).

Swedish Osteoporotic Fractures in Men Study

The Swedish Osteoporotic Fractures in Men Study (MrOS) is part of a collaborative effort including Hong Kong and USA, with the objective to study the epidemiology of bone mineral density, osteoporosis and fractures in ageing men[236]. The Swedish MrOS consists of n=3014 men, aged 69 to 81 years, who were randomly identified using national population registries and enrolled from centers in Malmö, Uppsala and Gothenburg between October 2001 and December 2004.

In this thesis, these men were combined with the MDCS cohort, in order to increase the power of the analysis by increasing the sample size. However, the follow-up procedure for this study differed from the two other cohorts, as PCa diagnoses were registered also for men who had died at follow-up. In order to ensure similar cohort structures, the deceased men were excluded for the work in this thesis (**Figure 19**).

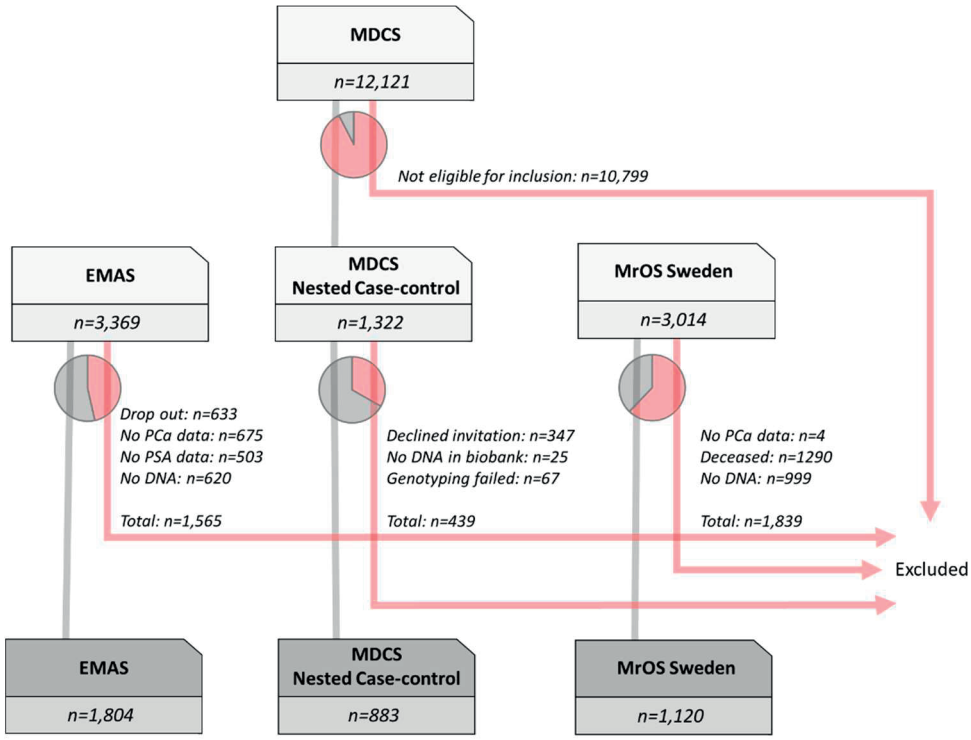


Figure 19 Flowchart describing the inclusion procedure in each genotyped cohort

Scandinavian Testicular Cancer Patients (TC) & Swedish Military Conscripts (M)

All men younger than 50 years with a testicular germ cell cancer (TGCC) diagnosed at the Department of Oncology, Lund University Hospital between March 1996 and October 2006, and at the Department of Oncology, Radiumhemmet and Södersjukhuset, Karolinska University Hospital between November 1998 and October 2006 (n=460) were invited to participate in a study regarding genetic variants and TGCC risk[237].

In addition, from a cohort of 300 men from a Danish TGCC study collected at the Department of Oncology and the Department of Growth and Reproduction, 100 samples were randomly selected and added to the Swedish TGCC cohort to create a Scandinavian TGCC cohort.

The controls for the TC cohort was made up of Swedish Military Conscripts, aged 18-20 years, initially recruited between 2000 and 2001 for a study on reproductive function in young Swedish men[238].

Genotypes

All the above mentioned cohorts had previously been genotyped for SNPs in the AR (**Figure 20, Table 2**). The genotyping methods differed between the different studies and more detailed information can be found in previously published works regarding EMAS[239], MrOS[240], TC & M[241] and MDCS[242]

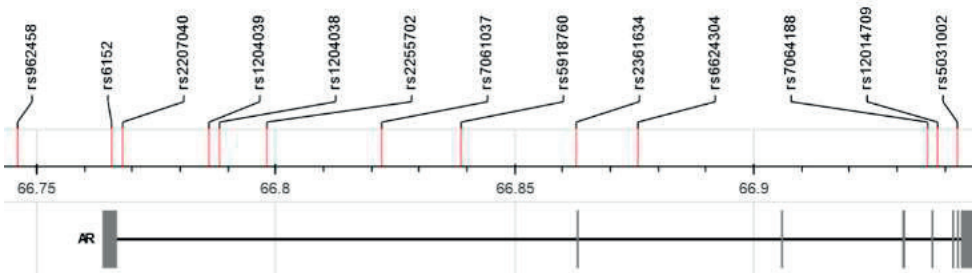


Figure 20 Location of the genotyped SNPs in relation to the AR gene

In short, the SNPs in EMAS, TC & M and MDCS were genotyped using MALDI-TOF mass spectrometry using Sequenom MassARRAY technology (Sequenom Inc., San Diego, CA, USA), while the Gothenburg part of MrOS was genotyped and imputed using the Illumina HumanOmni1 Quad v1.0 BeadChip and the Malmö part of MrOS was genotyped using the Illumina HumanOmniExpress-12 v1 BeadChip (Illumina, Inc., San Diego, CA).
The repetititive regions were amplified by PCR and analysed by sequencing on an ABIPRISM 3100 Genetic Analyser and genotyped using Genescan (ABI, Foster City, CA) for the CAG-repeat in the EMAS cohort, and on a Beckman Coulter CEQ 2000XL sequencer (Beckman Coulter, Bromma, Sweden) for the CAG and GGN-repeat in the TC & M cohorts.

The genotyped SNPs varied in each cohort and therefore, in order to combine the datasets, the linkage disequilibrium (LD) between the SNPs was investigated, to identify proxies for each SNP.

The pairwise LD (D' and r2) was calculated using the EUR population of 1000 genomes, through the webpage LDlink (https://analysistools.nci.nih.gov/LDlink/).

The r2 value is a stricter measurement of LD, and a high r2 value indicates that the alleles are more or less in complete linkage, so that variant A is always inherited with variant B and vice versa, while a high D' but a low r2-value usually occurs in the case of rare alleles, where the rare allele is commonly inherited with a more common allele of another SNP, but the more common allele does not always come with the rare allele.

Several of the SNPs were in high LD (r2>0.9) and could be considered to be representational of each other (**Figure 21**). These SNPs could be divided into two SNP-proxies, where the first one consisted of the SNPs rs6152, rs1204039, rs1204038, rs2255702, rs7061037, rs5918760 and rs6624304 and the second consisted of the SNPs rs7064188 and rs12014709. The other SNPs were not co-inherited with each other.

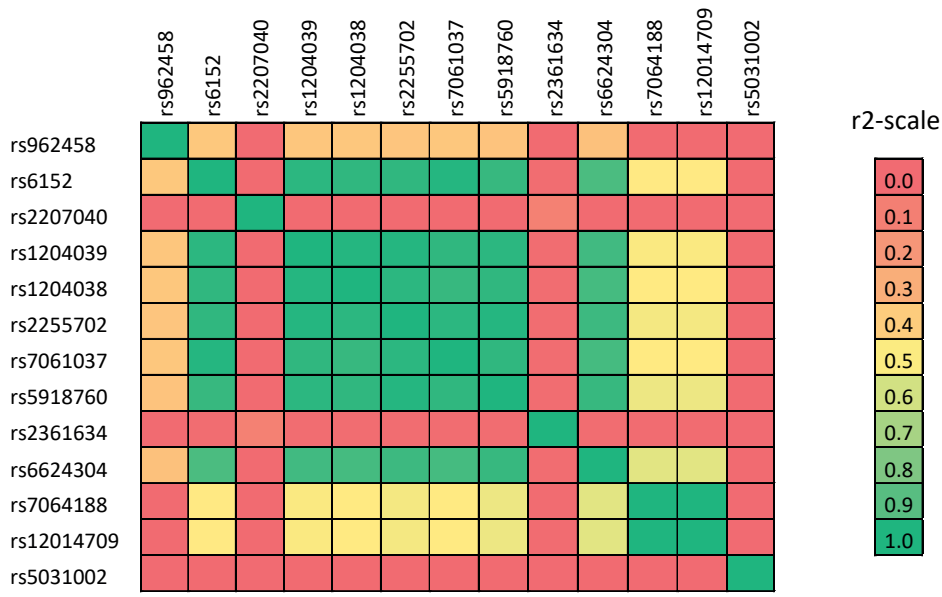


Figure 21 Heatmap of r2 values for the pairwise comparison between SNPs genotyped in at least one of the datasets used in this thesis

Table 2 Genotyping frequencies for each SNP in each of the cohorts

AR Variant	EMAS	MrOS	MDCS	TC	M
rs962458	-	1115 (99.6)	876 (99.2)	379 (99.2)	212 (99.1)
rs6152	1732 (96.0)	1110 (99.1)	-	347 (90.8)	210 (98.1)
rs2207040	-	595 (53.1)	874 (99.0)	378 (99.0)	213 (99.5)
rs1204039	-	1111 (99.2)	-	361 (94.5)	210 (98.1)
rs1204038	1797 (99.6)	1111 (99.2)	-	379 (99.2)	212 (99.1)
rs2255702	1759 (97.5)	1110 (99.1)	-	-	-
rs7061037	1765 (97.8)	1110 (99.1)	-	361 (94.5)	210 (98.1)
rs5918760	1796 (99.6)	1113 (99.4)	-	-	-
rs2361634	-	1038 (92.7)	-	380 (99.5)	212 (99.1)
rs6624304	-	1120 (100.0)	883 (100.0)	-	-
rs7064188	-	1120 (100.0)	871 (98.6)	-	-
rs12014709	-	1120 (100.0)	-	378 (99.0)	212 (99.1)
rs5031002	-	1103 (98.5)	-	380 (99.5)	212 (99.1)
CAG	1804 (100.0)	-	-	379 (99.2)	214 (100.0)
GGN	-	-	-	355 (92.9)	213 (99.5)

Haplotype construction

Based on the LD between the SNPs, five haplotype designating nodes were defined (**Figure 22a**). The seven SNPs with co-inherited alleles were used as the first node, and could divide the datasets into two level 1 haplotypes; H1 and H2, with frequencies of ~85% and ~15%, respectively, in the EUR dataset. The other nodes were used to divide the H1 and H2 haplotypes into sub-haplotypes (**Figure 22**b).

As all SNPs genotyped in the EMAS cohort were in linkage and belonged to node 1, this haplotype level was the only one that this dataset could be divided into (**Table 3**). The MrOS dataset had been genotyped for the SNPs used for division into level 2 and 3 haplotypes, but as the quality of imputation was too low (R^2-quality quality metric < 0.3) the data for the SNPs rs2207040, rs2361634 and rs5031002 were excluded from the analyses, leaving MrOS available only for division into haplotypes H1 and H2, as well as subdivision of H2 into H2a and H2b (**Table 3**). The MDCS SNP data allowed for division into all level 2 haplotypes, while the TC and M datasets, which also included CAG and GGN-data, allowed for complete division into level 3 haplotypes.

Using these haplotypes, the linkage between AR-haplotype and CAG-repeat length could be investigated in the EMAS, TC and M cohorts and the linkage between AR-haplotype and GGN-repeat length could be investigated in the TC and M cohorts.

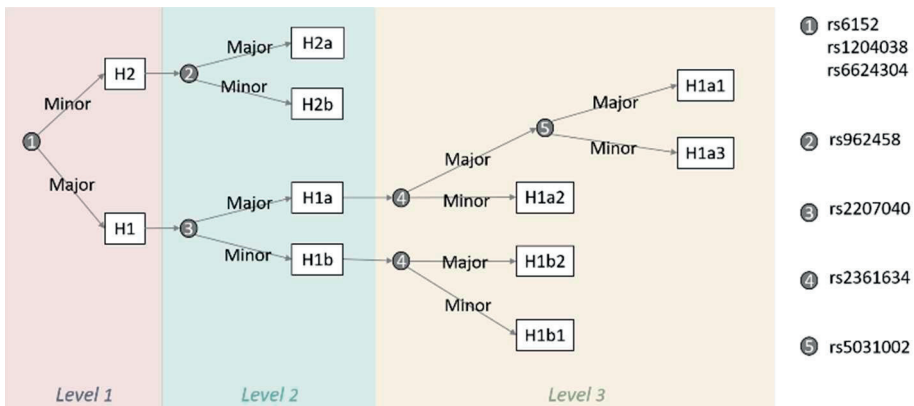
Thereafter, the association between the AR-haplotypes and risk of PCa was investigated in the EMAS, MDCS and MrOS cohorts and in EMAS, the association between AR-variants and serum PSA was also investigated.

The haplotype frequency was also compared to the five super-populations of 1000 genomes; AFR, Ad mixed American (AMR), EAS, EUR and South Asian (SAS), as the CAG-repeat length previously has been described to differ between populations, as has the PCa risk.

Further on, in order to find a potential cause for the differences in PCa risk, the LD between the risk haplotype and other genetic variants located in exonic regions were investigated using SNAP and LDlink and the results were compared with results from studies regarding androgenetic alopecia, another androgen driven condition.

Finally, using publicly available fold prediction software, RNAstructure, other potential causes for the differences were investigated.

a



b

Lvl1	Lvl2	Lvl3	rs962458	rs6152	rs2207040	rs1204039	rs1204038	rs2255702	rs7061037	rs5918760	rs2361634	rs6624304	rs7064188	rs12014709	rs5031002
H1	H1a	H1a1	A	G	A	G	G	C	A	C	A	C	A	T	G
H2	H2a		A	A	A	A	A	T	G	T	A	T	G	G	G
H2	H2b		G	A	A	A	A	T	G	T	A	T	A	T	G
H1	H1a	H1a2	A	G	A	G	G	C	A	C	G	C	A	T	G
H1	H1a	H1a3	A	G	A	G	G	C	A	C	A	C	A	T	A
H1	H1b	H1b1	A	G	G	G	G	C	A	C	G	C	A	T	G
H1	H1b	H1b2	A	G	G	G	G	C	A	C	A	C	A	T	G

Figure 22 a) Flowchart describing the division into level 1, level 2 and level 3 haplotypes, and **b)** the haplotypes constructed in a European dataset based on the SNPs in the flowchart. Minor alleles are represented by lighter background color.

Table 3. The number of subjects in each dataset successfully genotyped for each of the SNPs used to divide the men into haplotype carriers, as well as the haplotype level into which the men in each dataset could be divided into based on the genotyped SNPs.

Node		EMAS	MrOS	MDCS	TC	M
1		1797 (99.6)[a]	1120 (100.0)[b]	883 (100.0)[b]	379 (99.2)[a]	212 (99.1)[a]
2	rs962458		1115 (99.6)	876 (99.2)	379 (99.2)	212 (99.1)
3	rs2207040		[c]	874 (99.0)	378 (99.0)	213 (99.5)
4	rs2361634		[c]		380 (99.5)	212 (99.1)
5	rs5031002		[c]		380 (99.5)	212 (99.1)
	CAG	1804 (100.0)			379 (99.2)	214 (100.0)
	GGN				355 (92.9)	213 (99.5)
Haplotype level		1	1	2	3	3

a) rs1204038 b)rs6624304 c)Excluded due to low quality imputation

Subjects without genotype information

Testosterone studies

In order to investigate the association between testosterone concentrations in younger men, and long term risk of PCa and all-cause mortality, all testosterone measurements performed at the Department of Clinical Chemistry, Skåne University Hospital Malmö, Sweden between November 23, 1987 and July 29, 1992 were collected and using the personal identification number, the data was linked to the Swedish Cancer Registry, the Swedish Cause of Death Registry, and the Population Registry as of December 2013.

Between 1987 and 1992, serum levels of total testosterone were measured using the same radioimmunoassay with within and between assay variations of 6% and 9%, respectively[243].

In total, 10 540 testosterone measurements from 7 277 individuals were registered (**Figure 23**). All female patients (n=2 999) were excluded from the statistical analysis as were men who were diagnosed with PCa prior to their first registered testosterone measurement as it is likely that these men were under androgen suppressive therapy. In addition, all children and men younger than 20 years at the time of measurement were excluded since they were less likely to have completed pubertal development at the time of measuring.

Men who did not have a valid personal identification number or had disappeared from the population registry for unknown reasons were excluded, as were men who had had more than two testosterone measurements performed within two years, as there was high probability that those men were on androgen replacement therapy or had a chronic disease affecting their hormone levels. Finally, in order to assure that testosterone levels were not an effect of general morbidity, men who died within a year from their measurement were excluded resulting in a final dataset of n=3084 men with a mean and median age of 47 and 48 years, respectively (range: 20-87 years).

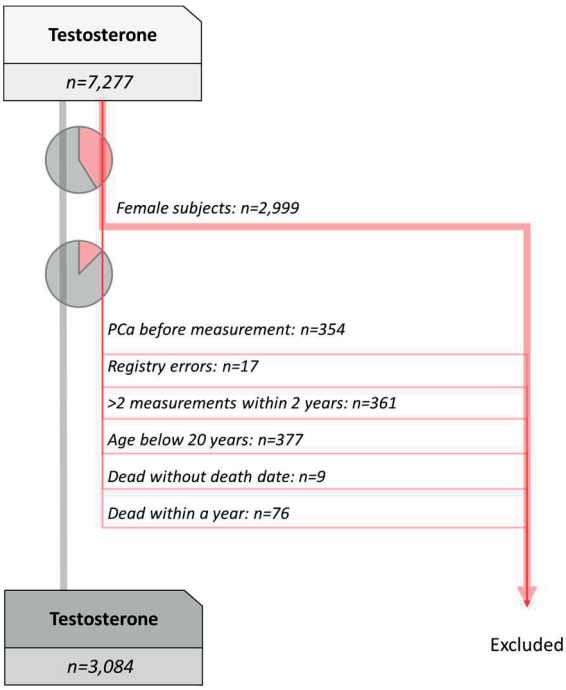


Figure 23 Flowchart describing the inclusion process of the testosterone cohort

As there is no consensus regarding the threshold for low and high testosterone, respectively, and as testosterone concentration is known to be negatively associated with age and the age-range in the dataset was large, age-normalized z-score was used for stratifying the cohort into low, middle and high testosterone concentration.

The first registered testosterone measurement in the database was used for the analyses. Twenty roughly equally sized groups were generated based on age and within each age group, the mean and standard deviation for testosterone was calculated. The age-normalized z-score was created by calculating the difference between the individual testosterone value and the mean testosterone concentration

within the specific age group to which the man belonged and then dividing this value by the standard deviation within the specific group.

The distribution of testosterone by age and z-score category is presented in **Figure 25**.

As this dataset consisted of patients with unknown medical history, in order to estimate whether this cohort was roughly representative for the general age matched population, the dataset was compared with the reference values of the laboratory; (8-30 nmol/L for men ≤50 years and 5-30 nmol/L for men >50 years; http://www.analysportalen-labmedicin.skane.se/viewAnalys.asp?Nr=2423) under the assumption that these values represented the 5th and 95th percentile in a normal healthy population and the concordance between the low and high groups defined by the thresholds used at the Department of Clinical Chemistry and the 5% lowest and highest z-score groups in our cohort was 98% after exclusion (**Figure 24**).

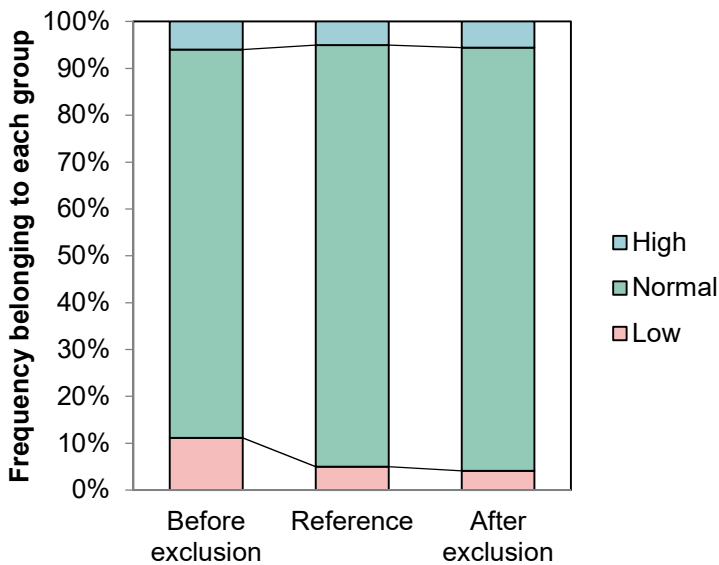


Figure 24 The proportion of men having testosterone concentrations in high, low and normal range based on the reference thresholds used at the Department of Clinical Chemistry at Skåne University Hospital before and after exclusion criteria were executed.

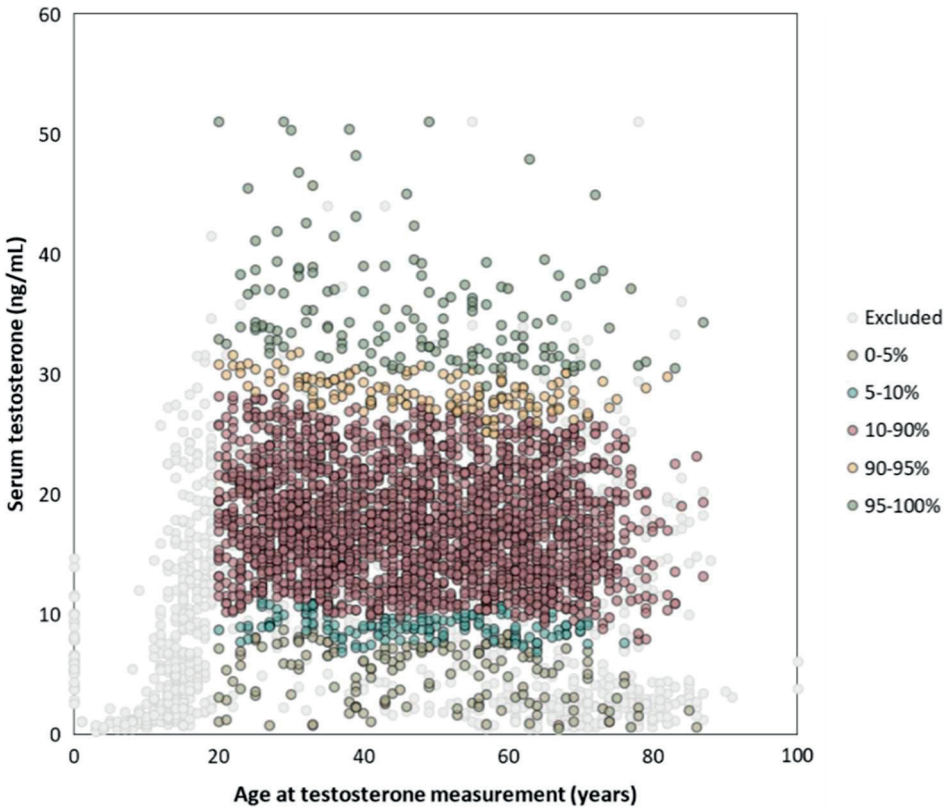


Figure 25 Testosterone concentration by age for each age-adjusted z-score category, as well as the men excluded from the analysis.

Men with a registered PCa-diagnosis (ICD7: 177; ICD9: 185; ICD10: C61) were considered PCa-cases. For men who were deceased at the time of data requisition, the underlying cause of death was categorized based on the ICD9 and ICD10 diagnoses. To gain power for the statistical analyses, general chapter categories, e.g. *chapter IX: diseases of the circulatory system* comprising all cardiovascular diseases (ICD9-codes 390-459 and ICD10-codes I00-I99 were used for the calculations regarding underlying cause of death in relation to testosterone.

Statistics

Study I: The CAG length was trichotomized into groups of similar sizes containing short (≤20CAG, n=520), average (21-23CAG, n=592) and long (≥24CAG, n=575) alleles. The CAG-groups and the haplotype tagging SNP alleles were tested independently, but they were also combined into 6 groups, rs1204038G or rs1204038A combined with short, average or long CAG-groups.

The difference in CAG repeat length for the two haplotypes was tested using the T-test, The association between ln-transformed PSA concentration and genotype was investigated using a univariate linear regression model. All analyses regarding PSA included center as covariate, as the different centers measured PSA using different measuring methods.

The risk of having PSA above clinically used thresholds for referring patients for urological examination, 3ng/mL or 4ng/mL, was analysed using logistic regression both with and without age as a covariate.

Finally, using the information from the follow-up questionnaire, the ORs for ever having been diagnosed with PCa for different genotypes was calculated by means of logistic regression analysis with center and age as covariates. This was done to deduce if carriers of the variant with increased PSA levels were more prone to be diagnosed with PCa. Among the 2736 men with follow-up data available, 932 were excluded due to lack of DNA for genetic analysis or lack of baseline PSA data.

In order to test the robustness of the association between genotype and PSA levels in relation to previous risk of being diagnosed with PCa, thereby ensuring that the difference in PSA levels was not due to a larger number of men with PCa being excluded from one group, the analysis of the association between genotype and PSA was repeated for men younger than 50 years (n=440), since no men in this age span had been diagnosed with PCa.

Study II: Logistic regression was used to calculate the OR and 95%CI for having PCa in relation to the rs6624304 genotype. All calculations were performed both with and without age-adjustment. In the MDCS study, calculations were done for the whole cohort as well as for those born in Sweden only.
Only men alive in December 2013 were used in the calculations on PCa risk in the MrOS cohort. For this cohort, additionally, the genotype-related OR for dying during the period between study initiation and December 2013, was calculated. As cause of death was not available, to rule out that carriers of one haplotype had a more aggressive PCa than the other one, the number of days between PCa diagnosis and death was investigated using linear regression.In addition, for the included cases, tumor data (Gleason, PSA, Metastasis, Nodes, Tumor stage) at the

time of diagnosis for the two haplotypes was compared using Mann Whitney U-test and Pearson's chi-square test.
Comparison of haplotype frequencies in the different cohorts to the European subpopulation from 1000 genomes was done using Fisher's exact test.

Study III and IV: Cox proportional hazards regression was performed to calculate the hazard ratio (HR) and 95% CI for PCa and all-cause mortality. For the deceased men, HR was also calculated for the cause of death-categories that had a frequency of ≥3%.

Since, in the older age group, there is a higher probability that low testosterone levels might be caused by an underlying condition, the analyses for associations between testosterone levels and risk of, as well as cause of, mortality were also performed separately for men ≥50 years and men <50 years at measurement, respectively. For the calculations regarding PCa and mortality, the groups used were based on following z-score percentile intervals: 0-5%, 5-10%, 10-90%, 90-95% and 95-100%. For the calculations of underlying cause of death the same groups were used, but as the number of cases in each cause of death was comparably low, to increase the statistical power tests were also performed with the z-score percentile intervals 0-10%, 10-90% and 90-100%.

Study V: Logistic regression was used to calculate the OR and 95% CI for having PCa for each of the haplotypes, using the most prevalent haplotype as reference. To further investigate the extent of the LD in the genomic region surrounding the AR, the web-based LDlink application [244] was used to identify SNPs with a high D' (>0.95) both up and downstream of the AR.

Thereafter, the genomic region with a high D' value identified via LDlink was investigated more thoroughly using the human genome browser [245] to identify common SNPs in exons, promoters or UTRs of nearby genes. Haplotype data regarding these SNPs and the AR haplotype-tagging SNPs were then downloaded for the five different 1000 genomes superpopulations using LDlink.

Statistical analyses were conducted using the SPSS v. 20-v.22 software (SPSS, Inc., Chicago).

Results and discussion

Testosterone concentrations and risk of PCa

In the investigation regarding serum testosterone and risk of PCa after 20-years, with the 10-90% group as the reference, no statistically significant difference in PCa occurrence was observed in any of the z-score groups (**Table 4**). However, while the 5-10% and 90-95% groups both had HRs for PCa similar to the reference, the highest and lowest 5% were consistently lower, although not statistically significantly so. High testosterone was not a risk factor for PCa. On the contrary, men with the 5% highest testosterone concentrations, as well as men with the 5% lowest concentrations, seemed to be at a slightly reduced risk compared to those in the middle group, although the difference did not reach statistical significance.

These results were in accordance with a previous report by Muller et al. who in the placebo arm of the Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trial found that testosterone levels were unrelated to PCa detection or grade after four years of follow-up[193]. However, in a secondary analysis in the same article of the data from the REDUCE trial, those with low baseline concentrations (<10nmol/L) had the lowest PCa risk, as had men with very high serum testosterone, although not statistically significantly so.

Low testosterone is associated with reduced fertility, and in two large national registry-based studies on more than 50 000 men, using fertility as a proxy for androgenicity, childless men had a significantly reduced risk of PCa compared to men who had fathered a child[246-248]. In a subsequent population based nested case control study, infertile men were at 50% reduced risk for PCa compared to fertile men[235], and a previous study found impaired Leydig cell function with lower testosterone concentrations in a larger proportion of infertile men compared to fertile men[249].

Taken together, it seems possible that men with particularly high or low testosterone levels could have slightly reduced PCa risk, but that the effect is too weak to detect even in a cohort of this size and even when two decades have passed between hormone measurement and PCa analysis.

Table 4 Testosterone concentrations and age, as well as the hazard ratio and 95% confidence intervals for prostate cancer for the different age normalized z-score groups.

			Testosterone	Age	Prostate cancer				
		n (%)	median (range)	mean (SD)	n (%)	HR(95%CI)	*p*	HR(95%CI)	p^{adj}
All ages	5% lowest	156 (5)	5.4 (0.5-8.7)	49 (16)	6 (4)	0.54 (0.24-1.23)	*0.142*	0.53 (0.23-1.19)	*0.122*
	6-10% lowest	153 (5)	9.0 (6.8-11.2)	47 (14)	12 (8)	0.98 (0.55-1.76)	*0.958*	1.03 (0.57-1.84)	*0.931*
	Middle 11-90 %	2466 (80)	17.1 (7.8-28.3)	47 (15)	189 (8)	Ref	*Ref*	Ref	*Ref*
	6-10% highest	156 (5)	28.2 (25.0-31.8)	49 (14)	12 (8)	1.05 (0.59-1.88)	*0.874*	1.01 (0.56-1.80)	*0.984*
	5% highest	153 (5)	33.9 (29.0-51.0)	47 (16)	8 (5)	0.70 (0.35-1.42)	*0.324*	0.68 (0.33-1.38)	*0.282*
Age <50	5% lowest	83 (5)	5.9 (0.7-8.7)	36 (9)	1 (1)	0.42 (0.06-3.03)	*0.387*	0.47 (0.07-3.43)	*0.458*
	6-10% lowest	87 (5)	9.0 (6.8-11.2)	37 (8)	4 (5)	1.43 (0.51-3.98)	*0.495*	1.19 (0.43-3.32)	*0.740*
	Middle 11-90 %	1358 (80)	17.6 (9.7-28.3)	35 (8)	44 (3)	Ref	*Ref*	Ref	*Ref*
	6-10% highest	73 (4)	29.0 (26.9-31.8)	36 (8)	0 (0)	-	-	-	-
	5% highest	86 (5)	34.3 (30.2-51.0)	36 (9)	1 (1)	0.37 (0.05-2.69)	*0.326*	0.33 (0.05-2.36)	*0.266*
Age ≥50	5% lowest	73 (5)	4.8 (0.5-8.3)	63 (8)	5 (7)	0.56 (0.23-1.36)	*0.200*	0.56 (0.23-1.35)	*0.195*
	6-10% lowest	66 (5)	8.8 (6.8-10.6)	60 (7)	8 (12)	0.83 (0.41-1.70)	*0.614*	0.90 (0.44-1.83)	*0.767*
	Middle 11-90 %	1108 (79)	16.6 (7.8-26.9)	62 (8)	145 (13)	Ref	*Ref*	Ref	*Ref*
	6-10% highest	83 (6)	27.6 (25.0-30.3)	61 (7)	12 (14)	1.15 (0.64-2.07)	*0.650*	1.22 (0.67-2.19)	*0.517*
	5% highest	67 (5)	32.2 (29-0-47.9)	62 (8)	7 (10)	0.85 (0.40-1.81)	*0.668*	0.82 (0.38-1.75)	*0.602*

Adj) adjusted for age

Testosterone concentrations and risk of mortality

Testosterone concentrations and risk of all-cause mortality

The risk of all-cause mortality was significantly higher for men in the 0-5% group (

Table 5). However, when the men were divided into younger and older than 50 years of age, the significant difference was seen only in younger men with the 5% lowest testosterone, no difference was seen in older men.

This finding is in accordance with several previous studies, reviewed in[250]. The results however, indicates that this increased risk was restricted to younger men, which suggests that low testosterone might be the cause of increased mortality rather than a consequence of a concomitant life-threating disease.

Testosterone concentrations and risk of mortality: chapter categories

Out of the 14 cause of death chapter categories present in the database, six occurred with a frequency above 3%, the most frequent being chapter IX; diseases of the circulatory system, to which n=475 (40%) of the underlying cause of deaths belonged (**Table 5**). No significantly increased risks were found for the two groups of men with low testosterone when all deceased men were analyzed. However, when divided into two groups, <50 and ≥50 years at time of testosterone

measurement, the younger men in the 6-10% group had a statistically significantly increased risk of chapter IX; diseases of the circulatory system.

Using the 10% lowest and highest groups, no statistically significant differences were detected when all men were analyzed, but when split into older and younger men, a statistically significantly increased risk was detected for chapter IX; diseases of the circulatory system in the younger men from the 10% lowest group.

While the dataset consisted of testosterone measurements ordained by a medical professional, it is likely that the results from this study would not be representable for a normal population without the strict inclusion criteria. As is seen in **Figure 24**, a large part of the initial cohort consisted of men with testosterone levels below the clinical criteria for normal concentrations. As a decrease in testosterone concentrations can be a consequence of medication or severe illness, the exclusion of all men dead within a year and men with a previous PCa-diagnosis made the cohort more comparable to a general population. Additionally, even though the younger men with low testosterone concentrations had a larger proportion of mortality, the age at death and the time between testosterone measurement and death was similar in this group when compared to the other groups, indicating that severe illness at the time of testosterone measurement is not the underlying cause of the increased proportion of deaths (**Figure 26**).

Together, the results from the testosterone study indicate that although androgen replacement therapy could push men in the lowest testosterone range into a PCa risk similar to those seen in normal men, many low testosterone associated deaths could be avoided[251].

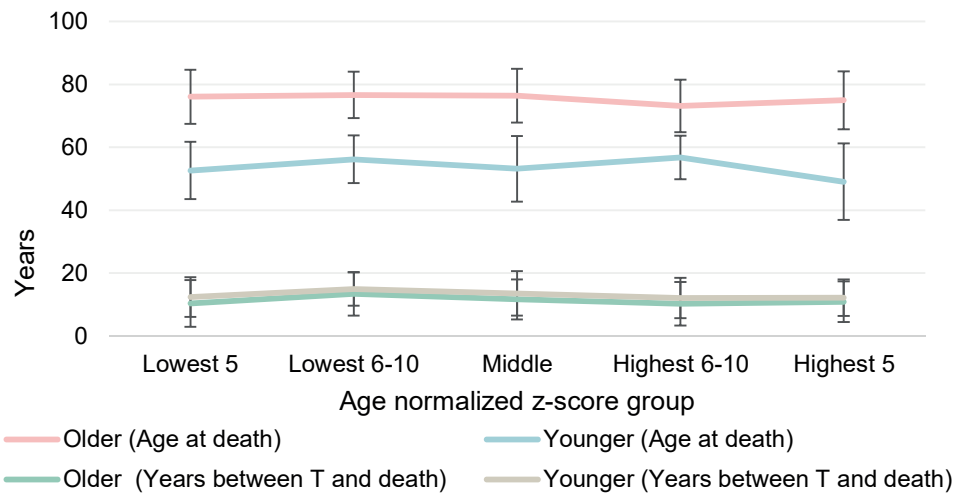


Figure 26. Average age at death with standard deviations for each z-score group for older (Age ≥ 50 at testosterone measurement) and younger (Age < 50 at testosterone measurement), as well as

average time between testosterone measurement and death date with standard deviations for each age and z-score group.

Table 5. Hazard ratios and 95% confidence intervals for all-cause mortality and the ICD chapter IX: Diseases of the circulatory system for the men in the different z-score groups

				Age normalized Z-score group						
			Total	Lowest 5	Lowest 6-10	Lowest 10	Middle	Highest 10	Highest 6-10	Highest 5
All-cause mortality	All ages	Deceased, n(%)	1200 (39)	76 (49)	59 (39)	135 (44)	937 (38)	128 (41)	64 (41)	64 (42)
		HR (95% CI)		**1.41 (1.11-1.78)**	0.94 (0.73-1.23)	1.16 (0.87-1.39)	*ref*	1.14 (0.95-1.37)	1.11 (0.86-1.44)	1.17 (0.91-1.50)
		p		***0.004****	*0.667*	0.112	*ref*	0.167	*0.406*	*0.234*
		HR (95% CI)		**1.39 (1.10-1.75)**	0.98 (0.75-1.27)	1.17 (0.98-1.41)	*ref*	1.09 (0.91-1.31)	1.06 (0.82-1.37)	1.13 (0.88-1.45)
		p[a]		***0.006****	*0.875*	0.084	*ref*	0.354	*0.663*	*0.355*
	Age <50	Deceased, n(%)	243 (14)	22 (27)	17 (20)	39 (23)	182 (13)	22 (14)	9 (12)	13 (15)
		HR (95% CI)		**2.22 (1.43-3.46)**	1.42 (0.86-2.33)	1.78 (1.26-2.52)	*ref*	1.06 (0.68-1.65)	0.91 (0.47-1.78)	1.19 (0.68-2.09)
		p		***<0.001****	*0.169*	0.001*	*ref*	0.803	*0.783*	*0.540*
		HR (95% CI)		**2.31 (1.48-3.60)**	1.27 (0.77-2.10)	1.71 (1.21-2.41)	*ref*	1.04 (0.67-1.62)	0.92 (0.47-1.80)	1.14 (0.65-2.01)
		p[a]		***<0.001****	*0.343*	0.003*	*ref*	0.860	*0.813*	*0.643*
	Age ≥50	Deceased, n(%)	957 (69)	54 (74)	42 (64)	96 (69)	755 (68)	106 (71)	55 (66)	51 (76)
		HR (95% CI)		1.19 (0.90-1.57)	0.80 (0.59-1.10)	0.98 (0.79-1.22)	*ref*	1.10 (0.90-1.35)	0.99 (0.76-1.31)	1.26 (0.95-1.67)
		p		*0.225*	*0.171*	0.983	*ref*	0.340	*0.962*	*0.117*
		HR (95% CI)		1.19 (0.90-1.57)	0.91 (0.67-1.25)	1.05 (0.85-1.30)	*ref*	1.10 (0.90-1.35)	1.09 (0.83-1.44)	1.12 (0.84-1.48)
		p[a]		*0.219*	*0.562*	0.653	*ref*	0.347	*0.539*	*0.446*
Chapter IX: Diseases of the circulatory system[b]	All ages	Yes, n (%)	475 (40)	31 (41)	24 (41)	55 (41)	372 (40)	48 (38)	26 (41)	22 (34)
		No, n (%)	725 (60)	45 (59)	35 (59)	80 (59)	565 (60)	80 (62)	38 (59)	42 (66)
		HR (95% CI)		1.12 (0.77-1.61)	0.82 (0.54-1.24)	0.97 (0.73-1.28)	*ref*	1.11 (0.82-1.50)	1.23 (0.82-1.83)	1.00 (0.65-1.54)
		p		*0.557*	*0.351*	*0.805*	*ref*	*0.490*	*0.318*	*0.988*
		HR (95% CI)		1.21 (0.84-1.75)	0.90 (0.59-1.36)	*1.05 (0.79-1.40)*	*ref*	*1.13 (0.84-1.53)*	1.24 (0.83-1.85)	1.03 (0.67-1.58)
		p[a]		*0.307*	*0.607*	*0.735*	*ref*	*0.417*	*0.291*	*0.895*
	Age <50	Yes, n (%)	53 (22)	6 (27)	**8 (47)**	**14 (36)**	35 (19)	4 (18)	2 (22)	2 (15)
		No, n (%)	189 (78)	16 (73)	**9 (53)**	**25 (94)**	146 (81)	18 (82)	7 (78)	11 (85)
		HR (95% CI)		1.94 (0.81-4.64)	**2.52 (1.16-5.47)**	**2.23 (1.19-4.19)**	*ref*	1.45 (0.51 (4.13)	1.73 (0.41-7.30)	1.24 (0.30-5.21)
		p		0.138	**0.020***	***0.012****	*ref*	*0.491*	0.453	0.769
		HR (95% CI)		2.03 (0.84-4.89)	**2.49 (1.15-5.42)**	**2.27 (1.21-4.27)**	*ref*	1.41 (0.49-4.02)	1.61 (0.38-6.83)	1.25 (0.30-5.26)
		p[a]		0.144	**0.021***	***0.011****	*ref*	*0.522*	0.517	0.757
	Age ≥50	Yes, n (%)	422 (44)	25 (46)	16 (38)	41 (43)	337 (45)	44 (42)	24 (44)	20 (39)
		No, n (%)	536 (56)	29 (54)	26 (62)	55 (57)	419 (55)	62 (58)	31 (56)	31 (61)
		HR (95% CI)		1.09 (0.72-1.63)	0.64 (0.39-1.06)	0.86 (0.62-1.19)	*ref*	1.04 (0.76-1.43)	1.11 (0.73-1.69)	0.97 (0.62-1.53)
		p		*0.693*	*0.084*	*0.347*	*ref*	*0.790*	*0.614*	*0.902*
		HR (95% CI)		1.11 (0.74-1.67)	0.69 (0.42-1.15)	0.90 (0.65-1.25)	*ref*	1.12 (0.82-1.53)	1.23 (0.81-1.86)	1.01 (0.65-1.59)
		p[a]		*0.622*	*0.155*	*0.521*	*ref*	*0.486*	*0.340*	*0.952*

Androgen receptor haplotypes

Frequencies in European populations

When the haplotype frequencies in TC+M were compared to the frequencies in the EUR subset of 1000 genomes to ensure that the frequencies were representable of a European cohort, the frequencies were similar (**Figure 27**). However, a slightly higher proportion of the H2b haplotype was present in the TC+M cohort.

Frequencies in world populations

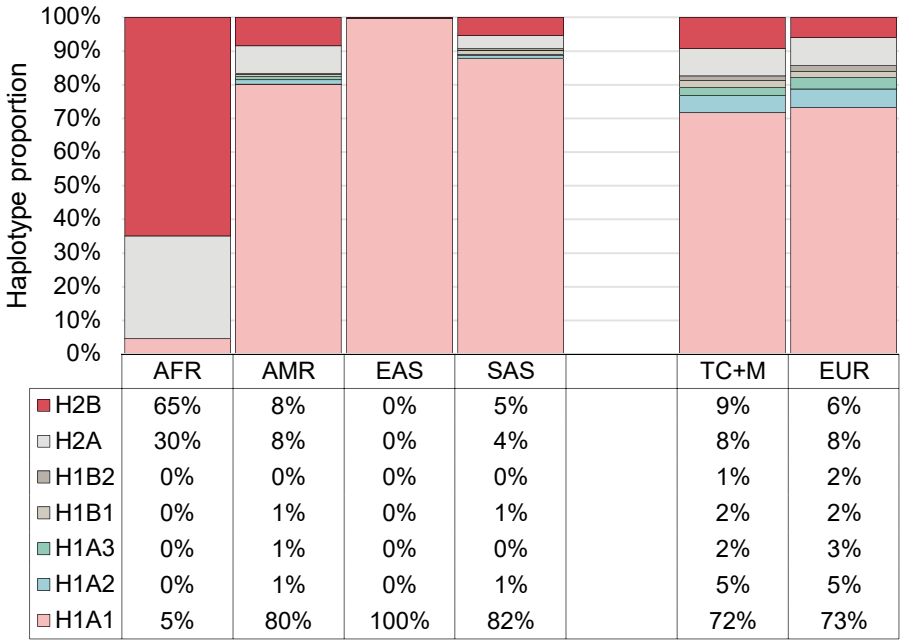
The largest difference in haplotype proportion was between the EAS and the AFR populations, where the EAS population almost exclusively presented with the H1a1 haplotype to only 5% of the AFR population.

Haplotypes and CAG/GGN repeat lengths

The CAG and GGN repeat lengths in the TC+M cohort can be seen in **Figure 28a**, and **Figure 29a**, respectively.

A significant difference in CAG-repeat lengths was observed between the H1 and H2 haplotype in the EMAS cohort, and the same was observed in the TC+M cohort where the H1 and H2 haplotypes had an average CAG of 22.2 and 20.1, respectively (**Table 6**, **Figure 28b**). Additionally, GGN-length differed with 85% of the H2-carriers having 24 repeats and 64% of the H1-carriers having 23 repeats (**Table 6**, **Figure 29a-f**). When the haplotypes were divided further into level 3 haplotypes, the CAG-allele spread in the H1 sub-haplotypes was large, with the H1b-haplotypes standing out with 63% of the H1b2 haplotype carriers presenting with 18 CAG-repeats and the H1b1 haplotype having its mode at 24 CAG-repeats (**Table 6**, **Figure 28c, e &f**). The same bimodal CAG distribution seen in the EMAS cohort was observed in the TC+M.

The H2-haplotypes were not as variable but the two allele peaks seen in **Figure 28b** were found to represent the two sub-haplotypes as 70% of H2a haplotype carriers had 21 CAG repeats and 46% of H2b carriers had 18 CAG repeats (**Table 6**, **Figure 28d**). As association studies regarding repeat variants in the AR are conducted with long alleles being compared to short alleles in order to gain statistical power, and the repeat cut-offs differ between studies it is possible that adding the haplotype tagging SNPs into the study could add more statistical power, and also reveal certain sub-groups with an increased risk, as for instance the GGN24-allele consists of ~50% of H1 and ~50% H2, and the H2-alleles differ in their CAG-repeat lengths.



	AFR	AMR	EAS	SAS		TC+M	EUR
H2B	65%	8%	0%	5%		9%	6%
H2A	30%	8%	0%	4%		8%	8%
H1B2	0%	0%	0%	0%		1%	2%
H1B1	0%	1%	0%	1%		2%	2%
H1A3	0%	1%	0%	0%		2%	3%
H1A2	0%	1%	0%	1%		5%	5%
H1A1	5%	80%	100%	82%		72%	73%

Figure 27 Haplotype proportion in each superpopulation of 1000 genomes as well as in the TC+M dataset. Haplotypes were constructed based on the SNPs rs962458, rs2207040, rs1204038, rs2361634 and rs5031002.

Table 6. The average CAG and GGN repeat length, as well as the most common allele (mode) for all haplotypes in the TC+M cohort.

Haplotype level			CAG			GGN		
1	2	3	Genotyped n (%)	Average (SD)	Mode (n; %)	Genotyped n (%)	Average (SD)	Mode (n; %)
H1			479 (99)	22.2 (3.0)	21 (70; 15)	456 (95)	22.7 (2.2)	23 (291;64)
	H1a		459(99)	22.3 (3.0)	21 (9; 15)	437 (95)	22.6 (2.3)	23 (272; 62)
		H1a1	415 (99)	22.2 (2.9)	20 (60; 14)	395 (94)	22.8 (1.7)	23 (241; 61)
		H1a2	30 (100)	23.1 (3.3)	21 (9; 30)	29 (97)	20.9 (5.5)	23 (22; 76)
		H1a3	14 (100)	23.2 (3.1)	21 (6; 43)	13 (93)	21.5 (3.0)	23 (9; 69)
	H1b		20(100)	21.2 (3.6)	18 (5; 25)	19 (95)	23.0 (0.0)	23 (19; 100)
		H1b1	12 (100)	23.8 (1.6)	24 (4; 33)	11 (92)	23.0 (0.0)	23 (11; 100)
		H1b2	8 (100)	17.1 (1.3)	18 (5; 63)	8 (100)	23.0 (0.0)	23 (8; 100)
H2			101 (100)	20.1 (2.1)	21 (34; 34)	99 (98)	23.7 (1.3)	24 (84; 85)
	H2a		47(100)	21.4 (1.7)	21 (33; 70)	46 (98)	23.6 (1.5)	24 (39; 85)
	H2b		54(100)	19.0 (1.7)	18 (25; 46)	53 (98)	23.7 (1.0)	24 (45; 85)

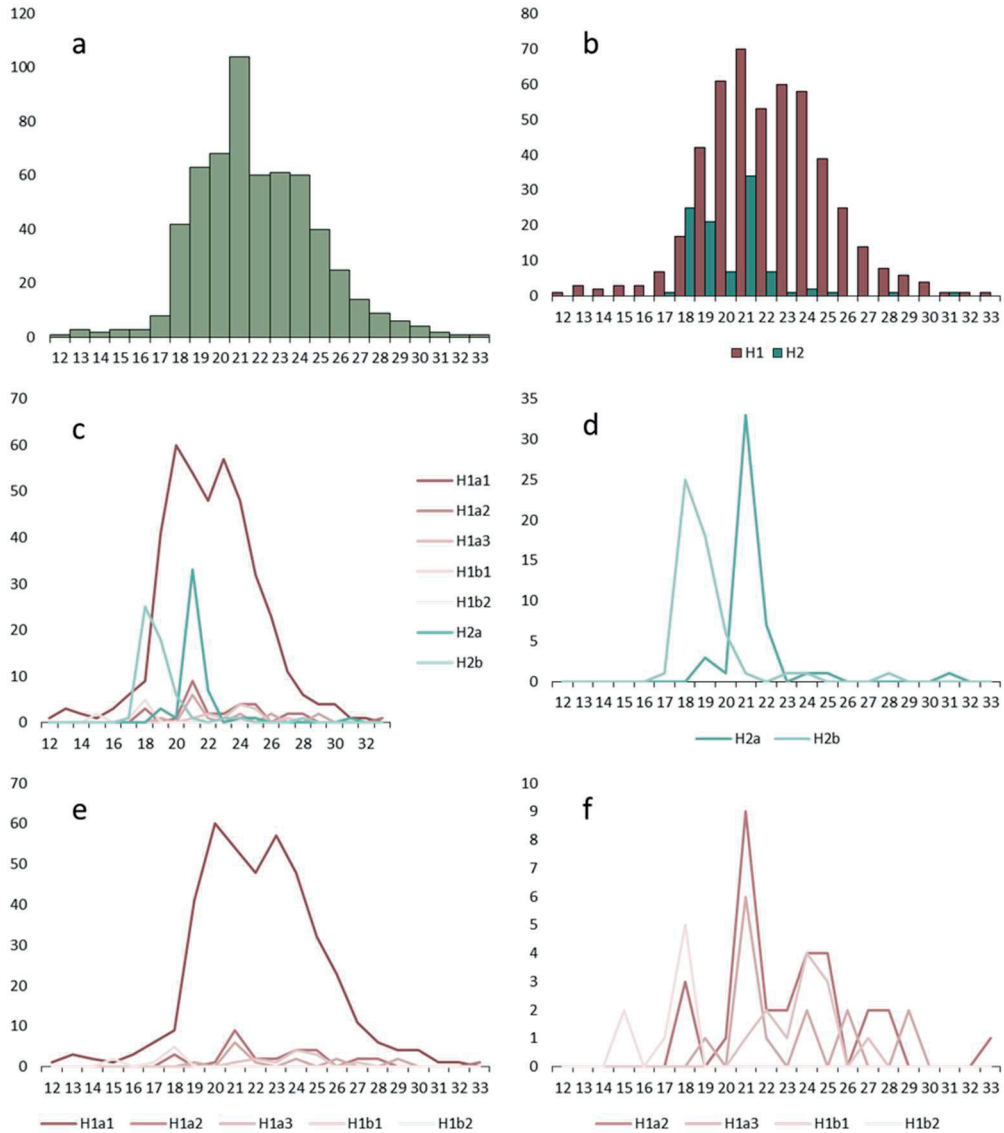


Figure 28 CAG repeat allele frequency distribution in **a)** All men, **b)** the H1 and H2 haplotypes, **c)** all haplotypes. Furthermore, the distribution in **d)** the two H2 haplotypes **e)** all the level 3 H1 haplotypes and **f)** in all level 3 H1 haplotypes excluding the most common H1a1.

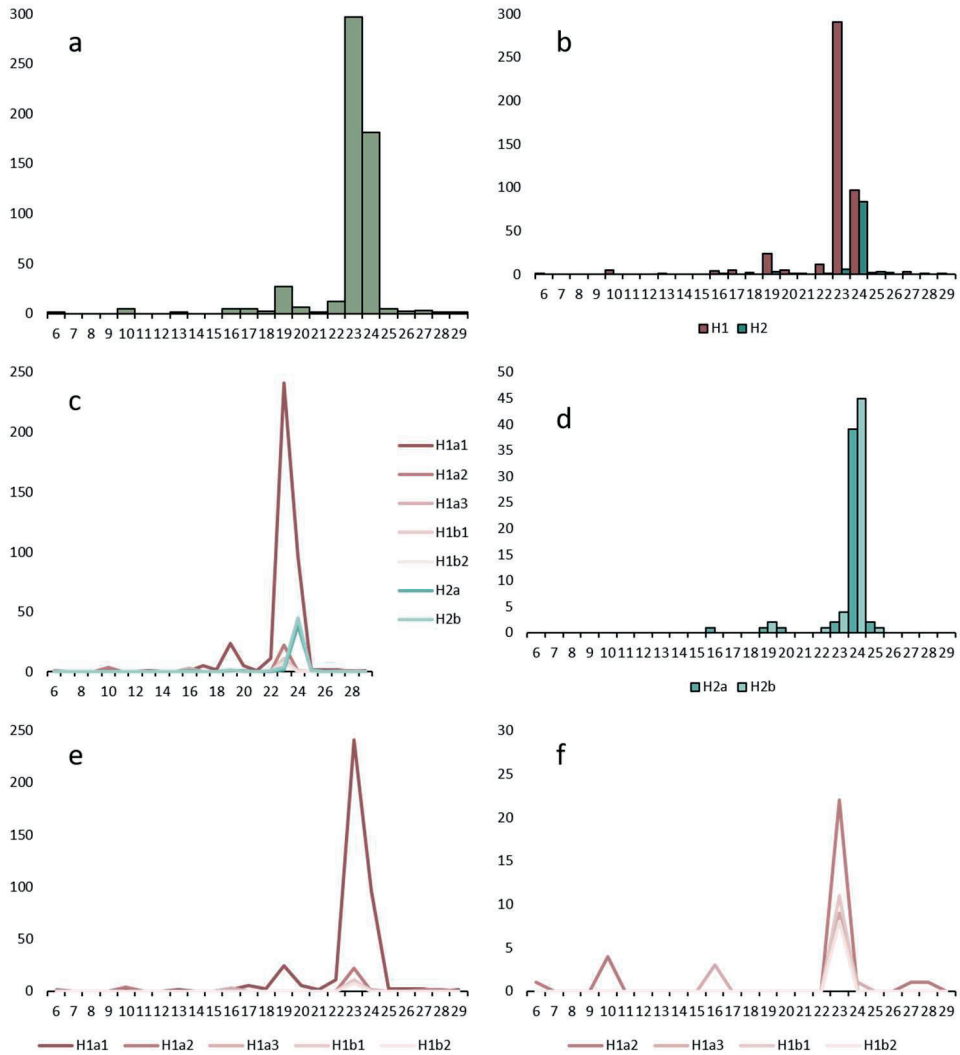


Figure 29 GGN repeat allele frequency distribution in **a)** All men, **b)** the H1 and H2 haplotypes, **c)** all haplotypes. Furthermore, the distribution in **d)** the two H2 haplotypes **e)** all the level 3 H1 haplotypes and **f)** in all level 3 H1 haplotypes excluding the most common H1a1.

AR-haplotypes, CAG and serum PSA concentrations

The transcription of PSA is androgen regulated and in study I the PSA concentration in men without PCa was tested against CAG-repeat length and against haplotype.

Regarding CAG and PSA no statistically significant differences were observed but when the H2 haplotype was tested against the H2 haplotype, the A-allele had 14% higher PSA concentrations (*p=0.045* with outliers included and *p=0.007* with outliers removed, both adjusted for center). The proportion of H2 and H1 carriers with PSA >3 ng/ml was 14% and 9%, respectively, and the proportion of H2 and H1 carriers with PSA >4 ng/ml was 9% and 5%, respectively (**Table 7**).

Table 7 Odds ratio and 95% confidence intervals for PSA above clinically utilized thresholds for carriers of H2 with H1 carriers as reference.

	PSA≤3 ng/ml n (%)	PSA>3 ng/ml n (%)	OR (95%CI)	*p*	PSA≤4 ng/ml n (%)	PSA>4 ng/ml n (%)	OR (95%CI)	*p*
H1	1309 (91)	128 (9)	*ref*	*ref*	1371 (95)	66 (5)	*ref*	*ref*
H2	234 (86)	38 (14)	1.69 (1.13-2.52)	*0.011*	248 (91)	24 (9)	1.99 (1.21-3.29)	*0.007*

All analyses adjusted for age at first visit and center with outliers excluded.

Studies regarding AR haplotype and PSA concentration in men without PCa are few, but in association studies regarding CAG-repeat length and PSA concentrations, shorter CAG-repeat lengths have been found to be associated with high serum PSA in older men[252] and with higher seminal PSA in younger men[253]. Since the H2-haplotype more often have short CAG-alleles it is possible that the correlation between PSA and CAG is capturing the same population as the H2 haplotype in our studies and that we would have seen the same correlation had more men been elderly in the EMAS cohort.

Another study in elderly men without PCa, but with urinary tract symptoms did not find an association between CAG and PSA[254].

Interestingly, another study of men without histologic evidence of PCa but with PSA above 4 ng/mL, African American men had significantly higher PSA levels when compared with White men[255], and the H2 allele is the most common haplotype in African and African American populations.

However, more longitudinal studies regarding PSA concentration in men and AR variants are needed, as it could be a valuable tool for adjustment in PSA screening,

AR-haplotypes and PCa risk

In study I the association between AR-haplotype and PCa-risk was also tested, and to some surprise, the H2 variant that presented with higher PSA had a lower risk of PCa. In a follow up study regarding PCa risk and haplotype, the same association was statistically significant in the MDCD dataset, but not in the MrOS dataset, possibly due to the smaller number of PCa cases in this cohort. However, the tendency in the MrOS dataset was the same and when the two datasets were pooled the results showed a statistically significant risk reduction for the H2-haplotype (**Table 8, Figure 30**).

In the MrOS and MDCS datasets additional SNPs had been genotyped, allowing for sub-group analyses. As the CAG-repeat length differs between the two H2 haplotypes a difference in risk reduction could be possible. However, while both H2 haplotypes displayed the same tendency towards decreased PCa risk, the results were not statistically significant in MDCS or MrOS alone, or when the two datasets were pooled (**Table 8, Figure 30**).

In addition, as all three datasets (EMAS, MDCS and MrOS) had level 1 haplotype information, all three datasets could be pooled, allowing for the PCa risk in the H2 haplotype to be approximated to an OR (95%CI) of 0.70 (0.54-0.91), with a p-value of 0.007, based on a cohort of n=689 PCa cases and n=3214 men without PCa.

A limitation to the PCa association is that only men alive at inclusion were used, introducing a risk of survivor bias to the results. It is possible that carriers of the H2-haplotype are more likely to suffer from more aggressive PCa with an earlier age-of-onset. While we did not find a difference in the cancer staging of the PCa patients, the reported data was not sufficient to carry out dependable statistical analysis.
The haplotype proportion in the datasets did not differ between our datasets and the EUR dataset of 1000 genomes, and in the MrOS-dataset, we could not see a genotype difference between the men who had died and the ones left in the study. However, we did not have cause of death for the deceased men and it is possible that a larger proportion of the deaths would be PCa in one of the haplotype carriers.

Based on our results, with the reservation that more longitudinal studies have to be performed taking into account information regarding for instance medication, smoking and drinking habits and diet, the 15% of European men who are carriers of the H2 haplotype appear to have a reduced risk of being diagnosed with PCa, but at the same time are more likely to present with a PSA above clinically utilized thresholds for the suspicion of PCa, and it is possible that the genotype information should be taken into account when PSA is being tested.

Table 8 OR and 95% CI for PCa in carriers of the H2-haplotypes in each individual cohort, as well as the pooled cohorts where possible

Dataset		Haplotype							
		H1		H2		H2a		H2b	
		Case	Control	Case	Control	Case	Control	Case	Control
EMAS	n (%)	57 (4)	1459 (96)	3 (1)	278 (99)				
	OR (95%CI)	*ref*		0.28 (0.09-0.89)					
	p	*ref*		*0.031**					
	Age Mean (sd)	68 (8)	59 (11)	79 (7)	59 (11)				
	OR (95%CI)^{adj}	*ref*		0.27 (0.08-0.88)					
	p^{adj}	*ref*		*0.030**					
MDCS	n (%)	400 (54)	347 (46)	58 (43)	78 (57)	30 (43)	39 (57)	28 (42)	38 (58)
	OR (95%CI)	*ref*		0.65 (0.45-0.93)		0.67 (0.41-1.10)		0.64 (0.38-1.06)	
	p	*ref*		*0.020**		*0.111*		*0.085*	
	Age Mean (sd)	74 (6)	74 (6)	74 (6)	73 (6)	75 (5)	72 (6)	73 (6)	73 (6)
	OR (95%CI)^{adj}	*ref*		0.65 (0.45-0.94)		0.67 (0.41-1.10)		0.64 (0.39-1.07)	
	p^{adj}	*ref*		*0.021**		*0.113*		*0.087*	
MrOS	n (%)	150 (16)	801 (84)	21 (12)	148 (88)	9 (10)	78 (90)	12 (15)	67 (85)
	OR (95%CI)	*ref*		0.76 (0.47-1.24)		0.62 (0.30-1.26)		0.96 (0.51-1.81)	
	p	*ref*		*0.226*		*0.182*		*0.891*	
	Age Mean (sd)	74 (3)	74 (3)	73 (3)	75 (3)	72 (3)	75 (3)	73 (3)	74 (3)
	OR (95%CI)^{adj}	*ref*		0.76 (0.47-1.24)		0.62 (0.30-1.26)		0.95 (0.50-1.81)	
	p^{adj}	*ref*		*0.275*		*0.189*		*0.884*	
Merged (all)	n (%)	607 (19)	2607 (81)	82 (14)	504 (86)				
	OR (95%CI)	*ref*		0.70 (0.54-0.90)					
	p	*ref*		0.005*					
	Age Mean (sd)	74 (6)	66 (11)	74 (5)	66 (11)				
	OR (95%CI)^{adj}	*ref*		0.70 (0.54-0.91)					
	p^{adj}	*ref*		*0.007**					
Merged (MrOS & MDCS)	n (%)	550 (32)	1148 (68)			39 (25)	117 (75)	40 (28)	105 (72)
	OR (95%CI)	*ref*				0.70 (0.48-1.01)		0.80 (0.55-1.16)	
	p	*ref*				*0.059*		*0.235*	
	Age Mean (sd)	74 (5)	74 (4)			75 (5)	74 (5)	73 (5)	74 (5)
	OR (95%CI)^{adj}	*ref*				0.70 (0.48-1.01)		0.79 (0.54-1.16)	
	p^{adj}	*ref*				*0.058*		*0.229*	

Adj) Adjusted for age

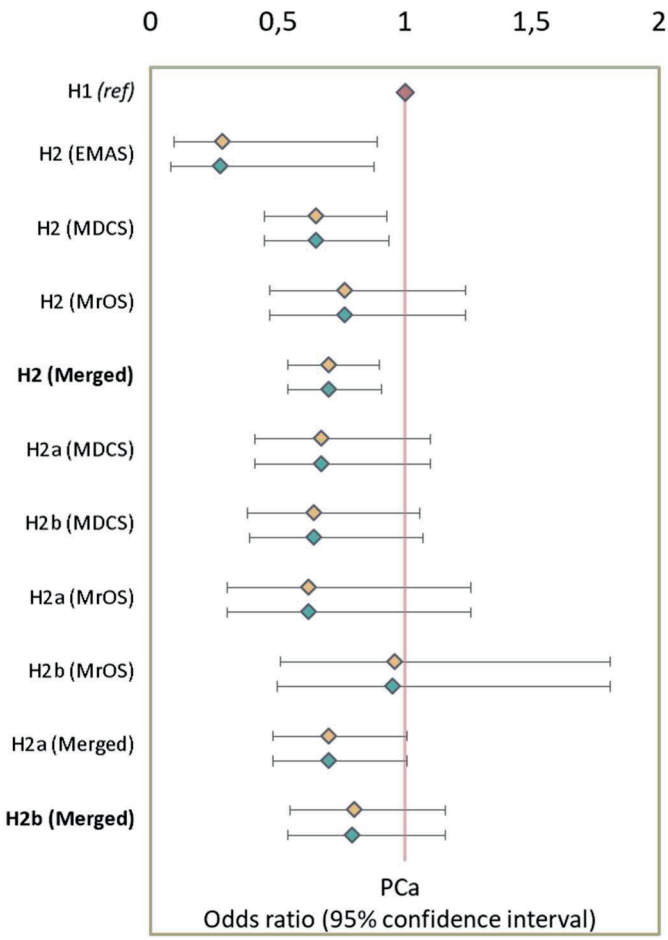


Figure 30. OR (95%CI) for PCa before (in pink) and after (in green) adjustment for age for the H2 haplotype in each dataset alone, as well as the combined datasets where division into the different subhaplotypes was possible after combination (for subanalysis of the H2a and H2b-haplotypes, the EMAS dataset was not included).

Potential causes for haplotype risk differences

The AR haplotypes identified in this thesis contains a combination of intronic and a synonymous SNP and would therefore not alter the amino acid sequence of the protein. However, other factors could still be at work.

As the SNP haplotypes appear to be linked to differences in CAG and GGN-repeat length, it is possible that the risk associations are due to specific combinations of repeat lengths.

mRNA stability

The CAG-repeat has been found to form a semi-stable hairpin in the AR mRNA, which is further stabilized by its flanking regions[140]. The difference in free energy changes depending on whether the CAG-repeat is an even or un-even number as the hairpin loop of the AR mRNA will consist of four or seven nucleotides in uneven and even repeats, respectively[141,142]. Additionally, each increment will grow the hairpin formation longer, which increases the stability of the GC-rich hairpin, and could affect for instance the translational speed and efficiency, but also the interaction between RNA-binding proteins and the mRNA[140]. In a study by Ding et. al., the calculated free energy for the GGN repeat was investigated, as they found that GGN length was inversely correlated with protein amount, and found that longer GGN repeats had a lower negative free energy that shorter GGN repeats[144].

Using the RNA-fold predictor software RNAstructure[256], the combined effect of rs6152, CAG and GGN-repeat length could be tested (**Figure 31**) and both an increased number of CAG-repeats and GGN-repeats lowered the negative free energy, meaning they increased the mRNA stability. I addition, the presence of the rs6152 G allele, present in the H1 haplotype further increased the stability.

As the most common combination of alleles were for H1; rs6152G, GGN23 and CAG24, or CAG21, and for H2; rs6152A, GGN24 and CAG21, the combinatory effect of these variants was tested specifically (**Figure 32**). This experiment suggested that the A-allele would be less stable in all combinations, except for when combined with long CAG and long GGN, where instead the G-allele in combination with short CAG and short GGN was slightly less stable. The stability of mRNA and its effect on protein translation is a fairly recent research area, and one of the factors thought to influence the stability is codon optimization[257], where two codons, although they are translated into the same amino acid, appear to have effect on the translation efficiency. These synonymous variants were previously considered inconsequential but studies on for instance synonymous variants in the dopamine receptor[258], psoriasis susceptibility variants[259], and the combinatory effect of synonymous and non-synonymous variants in the COMT-gene[260] have opened up

new possibilities to understand the dynamics of mRNA-folding and protein translation[261].

Therefore, it seems possible that the synonymous variant rs6152 could affect mRNA stability and thereby AR protein concentrations in the cells. However, as the mRNA stability is calculated *in silico*, *in vitro* experiments are needed to confirm the results.

rs6152	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G
GGN	17	17	18	18	19	19	20	20	21	21	22	22	23	23	24	24
6	-1065,2	-1068,7	-1068,0	-1071,5	-1070,2	-1073,7	-1073,0	-1075,1	-1075,3	-1078,3	-1078,5	-1079,9	-1080,1	-1083,1	-1081,7	-1084,7
7	-1067,9	-1071,6	-1070,7	-1074,4	-1072,9	-1076,6	-1075,7	-1078,0	-1078,0	-1081,2	-1081,2	-1082,8	-1082,8	-1086,0	-1084,4	-1087,6
8	-1067,7	-1071,2	-1070,5	-1074,0	-1072,7	-1076,2	-1075,5	-1077,6	-1077,8	-1080,8	-1081,0	-1082,4	-1082,6	-1085,6	-1084,2	-1087,2
9	-1070,4	-1074,1	-1073,2	-1076,9	-1075,4	-1079,1	-1078,2	-1080,5	-1080,5	-1083,7	-1083,7	-1085,3	-1085,3	-1088,5	-1086,9	-1090,1
10	-1070,2	-1073,7	-1073,0	-1076,5	-1075,2	-1078,7	-1078,0	-1080,1	-1080,3	-1083,3	-1083,5	-1084,9	-1085,1	-1088,1	-1086,7	-1089,7
11	-1072,9	-1076,6	-1075,7	-1079,4	-1077,9	-1081,6	-1080,7	-1083,0	-1083,0	-1086,2	-1086,2	-1087,8	-1087,8	-1091,0	-1089,4	-1092,6
12	-1072,7	-1076,2	-1075,5	-1079,0	-1077,7	-1081,2	-1080,5	-1082,6	-1082,8	-1085,8	-1086,0	-1087,4	-1087,6	-1090,6	-1089,2	-1092,2
13	-1075,4	-1079,1	-1078,2	-1081,9	-1080,4	-1084,1	-1083,2	-1085,5	-1085,5	-1088,7	-1088,7	-1090,3	-1090,3	-1093,5	-1091,9	-1095,1
14	-1075,2	-1078,7	-1078,0	-1081,5	-1080,2	-1083,7	-1083,0	-1085,1	-1085,3	-1088,3	-1088,5	-1089,9	-1090,1	-1093,1	-1091,7	-1094,7
15	-1077,9	-1081,6	-1080,7	-1084,4	-1082,9	-1086,6	-1085,7	-1088,0	-1088,0	-1091,2	-1091,2	-1092,8	-1092,8	-1096,0	-1094,4	-1097,6
16	-1077,7	-1081,2	-1080,5	-1084,0	-1082,7	-1086,2	-1085,5	-1087,6	-1087,8	-1090,8	-1091,0	-1092,4	-1092,6	-1095,6	-1094,2	-1097,2
17	-1080,4	-1084,1	-1083,2	-1086,9	-1085,4	-1089,1	-1088,2	-1090,5	-1090,5	-1093,7	-1093,7	-1095,3	-1095,3	-1098,5	-1096,9	-1100,1
18	-1080,2	-1083,7	-1083,0	-1086,5	-1085,2	-1088,7	-1088,0	-1090,1	-1090,3	-1093,3	-1093,5	-1094,9	-1095,1	-1098,1	-1096,7	-1099,7
19	-1082,9	-1086,6	-1085,7	-1089,4	-1087,9	-1091,6	-1090,7	-1093,0	-1093,0	-1096,2	-1096,2	-1097,8	-1097,8	-1101,0	-1099,4	-1102,6
20	-1082,7	-1086,2	-1085,5	-1089,0	-1087,7	-1091,2	-1090,5	-1092,6	-1092,8	-1095,8	-1096,0	-1097,4	-1097,6	-1100,6	-1099,2	-1102,2
21	-1085,4	-1089,1	-1088,2	-1091,9	-1090,4	-1094,1	-1093,2	-1095,5	-1095,5	-1098,7	-1098,7	-1100,3	-1100,3	-1103,5	-1101,9	-1105,1
22	-1085,2	-1088,7	-1088,0	-1091,5	-1090,2	-1093,7	-1093,0	-1095,1	-1095,3	-1098,3	-1098,5	-1099,9	-1100,1	-1103,1	-1101,7	-1104,7
23	-1087,9	-1091,6	-1090,7	-1094,4	-1092,9	-1096,6	-1095,7	-1098,0	-1098,0	-1101,2	-1101,2	-1102,8	-1102,8	-1106,0	-1104,4	-1107,6
24	-1087,7	-1091,2	-1090,5	-1094,0	-1092,7	-1096,2	-1095,5	-1097,6	-1097,8	-1100,8	-1101,0	-1102,4	-1102,6	-1105,6	-1104,2	-1107,2
25	-1090,4	-1094,1	-1093,2	-1096,9	-1095,4	-1099,1	-1098,2	-1100,5	-1100,5	-1103,7	-1103,7	-1105,3	-1105,3	-1108,5	-1106,9	-1110,1
26	-1090,2	-1093,7	-1093,0	-1096,5	-1095,2	-1098,7	-1098,0	-1100,1	-1100,3	-1103,3	-1103,5	-1104,9	-1105,1	-1108,1	-1106,7	-1109,7
27	-1092,9	-1096,6	-1095,7	-1099,4	-1097,9	-1101,6	-1100,7	-1103,0	-1103,0	-1106,2	-1106,2	-1107,8	-1107,8	-1111,0	-1109,4	-1112,6
28	-1092,7	-1096,2	-1095,5	-1099,0	-1097,7	-1101,2	-1100,5	-1102,6	-1102,8	-1105,8	-1106,0	-1107,4	-1107,6	-1110,6	-1109,2	-1112,2
29	-1095,4	-1099,1	-1098,2	-1101,9	-1100,4	-1104,1	-1103,2	-1105,5	-1105,5	-1108,7	-1108,7	-1110,3	-1110,3	-1113,5	-1111,9	-1115,1
30	-1095,2	-1098,7	-1098,0	-1101,5	-1100,2	-1103,7	-1103,0	-1105,1	-1105,3	-1108,3	-1108,5	-1109,9	-1110,1	-1113,1	-1111,7	-1114,7

CAG-repeat number

Figure 31 The calculated negative free energy for different CAG- and GGN-repeat lengths and the rs6152 SNP. Both longer CAG- and GGN repeats appears to increase the stability of the mRNA, as do the G-allele of rs6152. Yellow indicates lower mRNA stability and purple indicates higher mRNA stability.

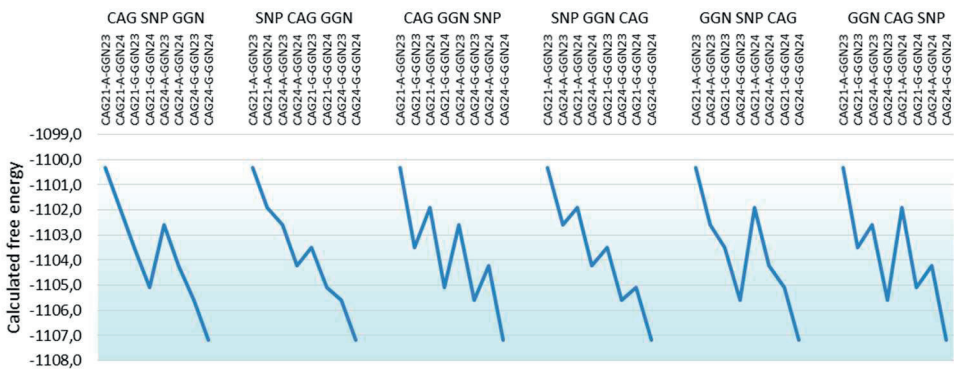


Figure 32 The combinatory effect of the most common CAG-repeat lengths, the most common GGN-repeat lengths and the two rs6152 variants. Each line shows the decrease in negative free energy sorted in order on one of the variants followed by a second and then third variant. The leftmost is for instance sorted first by CAG-repeat length, then by the SNP allele and then by GGN-repeat.

Linkage to genes further upstream and downstream

The genomic location where the AR resides has been suggested to have experienced an accelerated genetic drift after the human dispersal out of Africa[67,68], with a large number of derived allelic variants in the chromosomal region in European and Asian populations[69]. As the linkage distribution is somewhat high across the region, with several SNPs located upstreams and downstreams of the AR having high D', exonic SNPs in the genes surrounding AR were investigated in relation to the AR-haplotypes in the EUR superpopulation of 1000 genomes.

The upstream genes are HEPH, encoding for hephaestin, a copper dependent ferroxidase involved in the absorption of iron by the small intestines[262] and EDA2R, encoding for ectodysplasin A2 receptor, which bind the EDA-A2 isoform of transcripts of the anhidrotic ectodermal dysplasia (EDA) gene, involved in the development of skin, hair and teeth. This gene together with the EDAR gene were found to have undergone a recent natural selection in Asia[68]. The OPHN1 gene located downstream of the AR, encodes oligophrenin-1, which stimulates GTP hydrolysis of members of the Rho family. The frequency of the major allele of exonic SNPs were found to differ between the different haplotypes (**Figure 33**).

The EDA2R SNP rs1385699C>T (R57K) which has been suggested to be the main driver of the positive selection seen in the chromosomal region[69] is found at a lower frequency in carriers of the H2a, H2b and H1a3 haplotypes. Although a direct association to PCa is not evident in these genes, further studies into the genomic location might reveal clues to why the H2 haplotype is associated with a reduced risk of PCa.

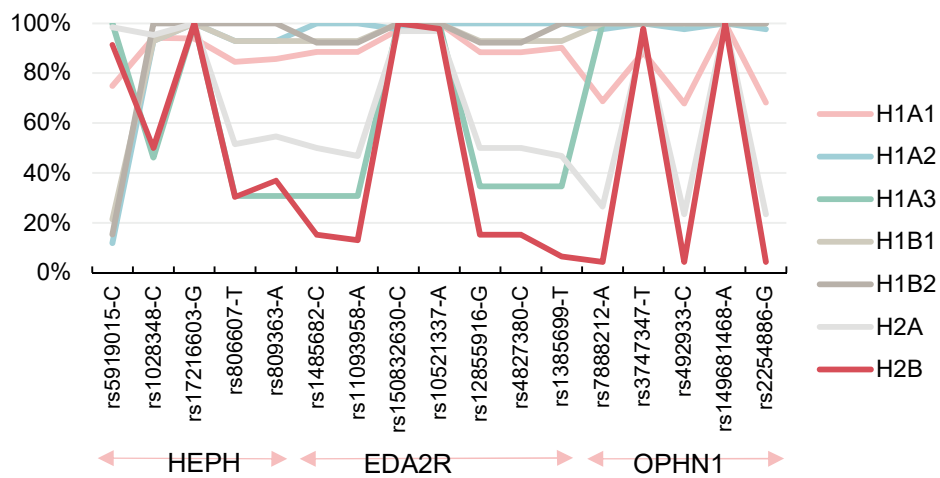


Figure 33 The major allele frequency of exonic SNPs in genes located near the AR gene divided by each haplotype.

Populärvetenskaplig sammanfattning

Prostatacancer är den vanligaste cancertypen hos män i Sverige, och ungefär var åttonde man diagnosticeras med sjukdomen innan sin 75-årsdag. Efter Charles B. Huggins nobelprisbelönade upptäckt att kastration, som markant sänker mängden testosteron i kroppen, kan förhindra prostatacancerns tillväxt har testosteron räknats som en bidragande faktor till cancerns bildande.

Testosteron, och en testosteronvariant som bildas i vissa delar av kroppen, 5-alfadihydrotestosteron (DHT), fungerar genom att koppla sig samman med androgenreceptorn (AR), som aktiveras av testosteronet och därefter slår igång processer som ger de egenskaper som vi oftast förknippar med testosteron, exempelvis skäggväxt, ökad muskelmassa, sexualdrift och mörkare röst.

Prostatan är en testosteronberoende körtel som tillverkar och utsöndrar prostatasekret som blandas med sädescellerna vid utlösning. Det innehåller flera viktiga beståndsdelar, bland annat enzymet prostataspecifikt antigen (PSA), som bidrar till att spermierna lättare kan nå kvinnans ägg.

Ofta förknippas PSA med PSA-testet, det blodprov som läkare tar när de misstänker prostatacancer. Vanligtvis finns bara väldigt låga nivåer av PSA i blodet men vissa prostatasjukdomar, däribland prostatacancer, kan göra så att PSA sipprar ut från prostatan och ökar koncentrationen av PSA i blodet. En hög PSA-koncentration i blodet kan alltså ibland vara ett tecken på prostatacancer, men tyvärr är inte testet perfekt. Många män diagnosticeras i onödan medan några mäns prostatacancer inte upptäcks förrän den har utvecklat metastaser och blivit så svårbehandlad att den ofta leder till döden.

För att PSA ska bildas krävs det att testosteron eller DHT binder till AR så att den kan sätta igång produktionen av PSA. Många studier har dock visat att det finns skillnader i hur effektivt AR fungerar, beroende på genetiska varianter i AR-genen. Dessutom har de olika genetiska varianterna i flera studier kopplats till olika risker för att utveckla prostatacancer, på samma sätt som olika testosteronnivåer har kopplats till prostatacancerrisk.

Trots att många studier gjorts i ämnet har forskare hittills inte kommit fram till något riktigt svar på om, och i så fall hur, testosteron och AR-varianter påverkar prostatacancerrisken.

För att vi skulle kunna se om testosteron i unga år påverkade den framtida risken för prostatacancer använde vi oss av de ca 4000 testosteronmätningar som gjordes på män av okänd anledning vid Skånes universitetssjukhus i Malmö mellan 1987 och 1992. Efter att vi tagit bort de som vi misstänkte hade prostatacancer eller blev behandlade med testosteron hade vi 3000 män kvar, som vi matchade med cancer- och dödsorsaksdatabaser för att ta reda på om låga eller höga testosteronnivåer utgjorde en risk för prostatacancer. Vi hittade ingen ökad risk för prostatacancerdiagnos hos de män som hade extremt höga eller låga testosteronkoncentrationer efter att de följts i ungefär 25 år men vi hittade istället en ökad risk för dödsfall hos de män som var yngre än 50 år när de mätte sitt testosteron och som visat sig ha låga testosteronnivåer.

Vi undersökte också genetiska varianter i AR hos ca 1800 Europeiska medelålders män som hade mätt sitt PSA. När vi tagit bort de män som hade haft prostatacancer ur beräkningen hittade vi där att risken att ha en PSA-koncentration som var högre än 4 ng/ml, ett gränsvärde som används kliniskt vid misstanke om prostatacancer, var dubbelt så hög hos de ca 15% av männen som hade ett A istället för ett G på plats 639 i den genetiska koden för AR, trots att de inte hade prostatacancer. Dessutom, när vi lade till ytterligare män, totalt 689 män med prostatacancer och 3214 män utan prostatacancer, såg vi att de män som hade ett A på plats 639 hade ungefär 40% lägre risk för prostatacancer.

Eftersom de 15% av Europeiska män som har ett A på plats 639 i AR-genen verkar ha en lägre prostatacancerrisk men samtidigt en högre risk att ha PSA-nivåer över kliniska gränsvärden så är det möjligt att de oftare blir upptäckta vid PSA-tester trots att de inte är sjuka. Samtidigt verkar män med G-varianten ha en ökad risk för prostatacancer men lägre PSA-nivåer vilket kan bidra till att deras cancer inte upptäcks i tid. Även om ytterligare studier krävs för att validera resultaten och undersöka orsaken bakom sambanden så är vår förhoppning att den genetiska varianten kan användas för att förbättra och precisera PSA-testet så att de män som behöver behandling upptäcks i tid.

Samtidigt indikerar våra resultat att höga testosteronnivåer inte innebär någon ökad risk för prostatacancer, men att män med låga testosteronnivåer verkar löpa en ökad risk att avlida i förtid. Det innebär att läkare i framtiden bör vara mer uppmärksamma på symptom på lågt testosteron även hos yngre män för att förbättra livskvaliten men också livslängden hos dessa män.

Acknowledgements

Först av allt vill jag tacka Yvonne Lundberg Giwercman, min huvudhandledare, som såg potential i mig och litade på min kompetens från dag ett. Utan din hjälp hade jag fortfarande skakat av nervositet inför presentationer och jag hade antagligen fortfarande suttit och petat i detaljer i det första datasetet jag fått tag i. Du har lärt mig att inte vara rädd och hur man ser skogen trots att det är en massa träd i vägen (även om jag fortfarande kan fastna vid ett träd eller två om jag tycker att de är extra intressanta ☺).

Jag vill också tacka mina bihandledare, Anders Bjartell som alltid visat intresse för mitt arbete, och Leon Brokken, som alltid funnits där för mig för att svara på alla mina frågor eller bjuda mig på öl när det behövts.

Aleksander Giwercman, som alltid ställt smarta och kritiska frågor som fått mig att tänka steget längre. Du har hjälpt mig att se patienterna bakom siffrorna, och ditt extrema informationstörstande har inspirerat mig.

Johan Malm och Lars Rylander, mina medförfattare på testosteronmanuskripten, som varit fantastiska bollplank när jag känt mig osäker över vilka metoder som skulle användas.

Min fantastiska forskargrupp som gjort det till ett nöje att komma till CRC. Hamideh, som alltid tagit hand om oss i gruppen och funnits där när något känts jobbigt. Ida L som direkt gjorde mig till en del av gruppen och som varit en fantastisk rumskamrat. Ditt engagemang och din förmåga att skapa struktur har alltid inspirerat mig (även om det kanske inte syns på mitt skrivbord).
Linus och Calle, som alltid varit bra med- och motspelare i alla pranks på labbet, vars mancave inne på labbet alltid hade någon knäpp historia att bjuda på om labbandet känts tråkigt och som alltid varit pigga på Azalee-AW.

Dorota och Louise, som varit fantastiska vänner på labbet och också fantastiskt resesällskap. Maria, som alltid varit full av kluriga idéer och Matilda, som trots att det ibland krävdes hetsinstagrammande stod ut med allt dumt vi hittade på (och lärde mig att man kan äta efterrätt till lunch ;)). Alexandra, som lärt mig allt jag kan om hästavel. Nishtman, som alltid livat upp konferenser och möten och nu

livar upp i gruppen. Lena, tack för allt ditt slit med genotypningen! Yahia, som hela tiden pendlar mellan störig och skitskoj men ändå oftast är en uppiggande diskussionspartner på kontoret. Christel, som introducerade metros horoskop för mig. Hannah, ett uppslagsverk av metoder och cellinjer och som alltid ställt upp och svarat på frågor även de gånger du inte varit en del av vår grupp.

Dessutom alla läkare i gruppen, som sorgligt nog bara varit där på märkligt schemalagda dagar men som alltid varit en frisk fläkt och ofta en *voice of reason* när våra samtalsämnen spårat för mycket på fikorna. Johannes, som alltid lugnt och metodiskt svarat på alla mina märkliga frågor om medicin. Jonatan, jag blir fortfarande rörd när jag tänker på stenen som du tog med för att din son tyckte att den var så fin att du skulle visa upp på jobbet. Sigrid, tack för alla stresshanteringssamtal! Carolina, som verkligen gjort allt för att hjälpa mig med bland annat mina frågor om cancerdiagnoser och ICD-koder. Amelie, du är en solstråle och en virvelvind.

Gunilla och Ida B, ni räddade mitt liv och jag är er evigt tacksam.

Anna, vars outtröttliga engagemang och positiva inställning alltid gjort fikorna roligare och jobbet lättare.

Min familj, som aldrig ställt några krav på mig men som alltid stöttat mig. Pappa som alltid har uppmuntrat mig när jag tagit sönder saker för att förstå hur de är uppbyggda och som alltid hjälpt mig att sätta ihop sakerna igen. Mamma som alltid varit på min sida och uppmuntrat min kreativitet. Ni tvingade mig aldrig att studera men indirekt har ni uppfostrat mig till att bli en sökare och en forskare.

Love, min bästa vän, som jag vuxit upp med och som alltid hjälpt mig när livet känts tungt.

Nina, också min bästa vän, som i 24 år stått vid min sida och backat upp mig när jag behövt stöd, och som inspirerat mig med sin orädda attityd och outsinliga kreativitet.

David, som från första stund förstått mig och stöttat mig och som låtit mig sitta uppe hela nätterna och skriva. Nu är jag klar!

References

1. Nieschlag, E. & Nieschlag, S. Testosterone deficiency: a historical perspective. *Asian journal of andrology* **16**, 161-168 (2014).
2. Klein, M. [Berthold's article: Transplantation of the testes (1849)]. *Archives d'anatomie, d'histologie et d'embryologie normales et experimentales* **51**, 379-386 (1968).
3. Brown, S. NOTE ON THE EFFECTS PRODUCED ON MAN BY SUBCUTANEOUS INJECTIONS OF A LIQUID OBTAINED FROM THE TESTICLES OF ANIMALS. *The Lancet* **134**, 105-107.
4. Cussons, A.J., Bhagat, C.I., Fletcher, S.J. & Walsh, J.P. Brown-Sequard revisited: a lesson from history on the placebo effect of androgen treatment. *Med J Aust* **177**, 678-679 (2002).
5. Gallagher, T. & Koch, F.C. The testicular hormone. *Journal of Biological Chemistry* **84**, 495-500 (1929).
6. Butenandt, A. Über die chemische Untersuchung der Sexualhormone. *Angewandte Chemie* **44**, 905-908 (1931).
7. David, K., Dingemanse, E., Freud, J. & Laqueur, E. Über krystallinisches männliches Hormon aus Hoden (Testosteron), wirksamer als aus Harn oder aus Cholesterin bereitetes Androsteron. in *Hoppe-Seyler´s Zeitschrift für physiologische Chemie*, Vol. 233 281 (1935).
8. Butenandt, A., Hanisch, G. Testosterone. The transformation of dehydroandrosterone into androstendiol and testosterone; a method for producing testosterone from cholesterin. *Hoppe-Seyler's Z Physiol Chem* **1935**, 89-98 (1935).
9. Ruzicka, L., Wettstein, A. Synthesis of the testicular hormone (testosterone) (androstene 3-on-17-ol) *Helv chim Acta.* **1935**, 1264-1275 (1935).
10. Hamilton, J.B. TREATMENT OF SEXUAL UNDERDEVELOPMENT WITH SYNTHETIC MALE HORMONE SUBSTANCE1. *Endocrinology* **21**, 649-654 (1937).
11. Stocco, D.M., Wang, X., Jo, Y. & Manna, P.R. Multiple signaling pathways regulating steroidogenesis and steroidogenic acute regulatory protein expression: more complicated than we thought. *Mol Endocrinol* **19**, 2647-2659 (2005).

12. Krsmanovic, L.Z., Hu, L., Leung, P.K., Feng, H. & Catt, K.J. The hypothalamic GnRH pulse generator: multiple regulatory mechanisms. *Trends Endocrinol Metab* **20**, 402-408 (2009).
13. Askew, E.B., Gampe, R.T., Jr., Stanley, T.B., Faggart, J.L. & Wilson, E.M. Modulation of androgen receptor activation function 2 by testosterone and dihydrotestosterone. *The Journal of biological chemistry* **282**, 25801-25816 (2007).
14. de Ronde, W., Pols, H.A., van Leeuwen, J.P. & de Jong, F.H. The importance of oestrogens in males. *Clin Endocrinol (Oxf)* **58**, 529-542 (2003).
15. Caronia, L.M.*, et al.* Abrupt decrease in serum testosterone levels after an oral glucose load in men: implications for screening for hypogonadism. *Clin Endocrinol (Oxf)* **78**, 291-296 (2013).
16. Dabbs, J.M., Jr. Salivary testosterone measurements: reliability across hours, days, and weeks. *Physiol Behav* **48**, 83-86 (1990).
17. Gupta, S.K., Lindemulder, E.A. & Sathyan, G. Modeling of circadian testosterone in healthy men and hypogonadal men. *J Clin Pharmacol* **40**, 731-738 (2000).
18. Cooke, R.R., McIntosh, J.E. & McIntosh, R.P. Circadian variation in serum free and non-SHBG-bound testosterone in normal men: measurements, and simulation using a mass action model. *Clin Endocrinol (Oxf)* **39**, 163-171 (1993).
19. Huhtaniemi, I. Late-onset hypogonadism: current concepts and controversies of pathogenesis, diagnosis and treatment. *Asian journal of andrology* **16**, 192-202 (2014).
20. Harman, S.M.*, et al.* Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. *The Journal of clinical endocrinology and metabolism* **86**, 724-731 (2001).
21. Kaufman, J.M. & Vermeulen, A. The decline of androgen levels in elderly men and its clinical and therapeutic implications. *Endocrine reviews* **26**, 833-876 (2005).
22. Layton, J.B.*, et al.* Testosterone Lab Testing and Initiation in the United Kingdom and the United States, 2000 to 2011. *The Journal of Clinical Endocrinology & Metabolism* **99**, 835-842 (2014).
23. Malik, R.D., Lapin, B., Wang, C.E., Lakeman, J.C. & Helfand, B.T. Are we testing appropriately for low testosterone?: Characterization of tested men and compliance with current guidelines. *The journal of sexual medicine* **12**, 66-75 (2015).
24. Tajar, A.*, et al.* Characteristics of Secondary, Primary, and Compensated Hypogonadism in Aging Men: Evidence from the European Male Ageing Study. *The Journal of Clinical Endocrinology & Metabolism* **95**, 1810-1818 (2010).
25. Bhasin, S. & Basaria, S. Diagnosis and treatment of hypogonadism in men. *Best Pract Res Clin Endocrinol Metab* **25**, 251-270 (2011).

26. Kumar, P., Kumar, N., Thakur, D.S. & Patidar, A. Male hypogonadism: Symptoms and treatment. *J Adv Pharm Technol Res* **1**, 297-301 (2010).
27. Carnegie, C. Diagnosis of hypogonadism: clinical assessments and laboratory tests. *Rev Urol* **6 Suppl 6**, S3-8 (2004).
28. Makhsida, N., Shah, J., Yan, G., Fisch, H. & Shabsigh, R. Hypogonadism and metabolic syndrome: implications for testosterone therapy. *The Journal of urology* **174**, 827-834 (2005).
29. Naifar, M.*, et al.* Male hypogonadism and metabolic syndrome. *Andrologia* **47**, 579-586 (2015).
30. Corona, G.*, et al.* Hypogonadism and metabolic syndrome. *Journal of endocrinological investigation* **34**, 557-567 (2011).
31. Ding, E.L., Song, Y., Malik, V.S. & Liu, S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *Jama* **295**, 1288-1299 (2006).
32. Kupelian, V.*, et al.* Low sex hormone-binding globulin, total testosterone, and symptomatic androgen deficiency are associated with development of the metabolic syndrome in nonobese men. *The Journal of clinical endocrinology and metabolism* **91**, 843-850 (2006).
33. Traish, A.M., Saad, F., Feeley, R.J. & Guay, A. The dark side of testosterone deficiency: III. Cardiovascular disease. *J Androl* **30**, 477-494 (2009).
34. Shores, M.M., Matsumoto, A.M., Sloan, K.L. & Kivlahan, D.R. Low serum testosterone and mortality in male veterans. *Archives of internal medicine* **166**, 1660-1665 (2006).
35. Laughlin, G.A., Barrett-Connor, E. & Bergstrom, J. Low serum testosterone and mortality in older men. *The Journal of clinical endocrinology and metabolism* **93**, 68-75 (2008).
36. Gilbert, R. *Early Modern Hermaphrodites: Sex and Other Stories*, (Palgrave Macmillan UK, 2002).
37. Steinberg, A. & Rosner, F. *Encyclopedia of Jewish Medical Ethics: A Compilation of Jewish Medical Law on All Topics of Medical Interest*, (Feldheim Publishers, 2003).
38. Creighton, S.M. & Liao, L.M. Changing attitudes to sex assignment in intersex. *BJU international* **93**, 659-664 (2004).
39. Sinding, C. Clinical research and basic science: the development of the concept of end-organ resistance to a hormone. *J Hist Med Allied Sci* **45**, 198-232 (1990).
40. Reifenstein, E.C., Jr. Hereditary familial hypogonadism. *Proc Am Fed Clin Res* **3**, 86 (1947).
41. Gilbert, D., Savoie, Sebaoun, Alexandre, C. & Belaisch, J. [Case of familial androgynism with severe hypospadias, gynecomastia and excessive estrogen production]. *Ann Endocrinol (Paris)* **18**, 93-101 (1957).
42. Lubs, H.A., Jr., Vilar, O. & Bergenstal, D.M. Familial male pseudohermaphrodism with labial testes and partial feminization:

endocrine studies and genetic aspects. *The Journal of clinical endocrinology and metabolism* **19**, 1110-1120 (1959).
43. Rosewater, S., Gwinup, G. & Hamwi, G.J. Familial Gynecomastia. *Ann Intern Med* **63**, 377-385 (1965).
44. Lyon, M.F. & Hawkes, S.G. X-linked gene for testicular feminization in the mouse. *Nature* **227**, 1217-1219 (1970).
45. Bremner, W.J., Ott, J., Moore, D.J. & Paulsen, C.A. Reifenstein's syndrome: investigation of linkage to X-chromosomal loci. *Clin Genet* **6**, 216-220 (1974).
46. Wilson, J.D., Harrod, M.J., Goldstein, J.L., Hemsell, D.L. & MacDonald, P.C. Familial incomplete male pseudohermaphroditism, type 1. Evidence for androgen resistance and variable clinical manifestations in a family with the Reifenstein syndrome. *The New England journal of medicine* **290**, 1097-1103 (1974).
47. Amrhein, J.A., Klingensmith, G.J., Walsh, P.C., McKusick, V.A. & Migeon, C.J. Partial androgen insensitivity: the Reifenstein syndrome revisited. *The New England journal of medicine* **297**, 350-356 (1977).
48. Aiman, J., Griffin, J.E., Gazak, J.M., Wilson, J.D. & MacDonald, P.C. Androgen insensitivity as a cause of infertility in otherwise normal men. *The New England journal of medicine* **300**, 223-227 (1979).
49. Migeon, B.R., Brown, T.R., Axelman, J. & Migeon, C.J. Studies of the locus for androgen receptor: localization on the human X chromosome and evidence for homology with the Tfm locus in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 6339-6343 (1981).
50. Lubahn, D.B.*, et al.* Cloning of human androgen receptor complementary DNA and localization to the X chromosome. *Science* **240**, 327-330 (1988).
51. Chang, C.S., Kokontis, J. & Liao, S.T. Molecular cloning of human and rat complementary DNA encoding androgen receptors. *Science* **240**, 324-326 (1988).
52. Brown, T.R.*, et al.* Deletion of the steroid-binding domain of the human androgen receptor gene in one family with complete androgen insensitivity syndrome: evidence for further genetic heterogeneity in this syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 8151-8155 (1988).
53. Brown, C.J.*, et al.* Androgen receptor locus on the human X chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. *American journal of human genetics* **44**, 264-269 (1989).
54. Lubahn, D.B.*, et al.* Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 9534-9538 (1989).

55. Laspada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E. & Fischbeck, K.H. Androgen Receptor Gene-Mutations in X-Linked Spinal and Bulbar Muscular-Atrophy. *Nature* **352**, 77-79 (1991).
56. Patterson, M.N., Hughes, I.A., Gottlieb, B. & Pinsky, L. The androgen receptor gene mutations database. *Nucleic acids research* **22**, 3560-3562 (1994).
57. Gottlieb, B., Beitel, L.K., Nadarajah, A., Paliouras, M. & Trifiro, M. The androgen receptor gene mutations database: 2012 update. *Hum Mutat* **33**, 887-894 (2012).
58. Huang, P., Chandra, V. & Rastinejad, F. Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics. *Annual review of physiology* **72**, 247-272 (2010).
59. Robinson-Rechavi, M., Carpentier, A.S., Duffraisse, M. & Laudet, V. How many nuclear hormone receptors are there in the human genome? *Trends in genetics : TIG* **17**, 554-556 (2001).
60. Zhang, Z.D., Cayting, P., Weinstock, G. & Gerstein, M. Analysis of nuclear receptor pseudogenes in vertebrates: how the silent tell their stories. *Molecular biology and evolution* **25**, 131-143 (2008).
61. Escriva, H., Bertrand, S. & Laudet, V. The evolution of the nuclear receptor superfamily. *Essays Biochem* **40**, 11-26 (2004).
62. Nuclear Receptors Nomenclature, C. A unified nomenclature system for the nuclear receptor superfamily. *Cell* **97**, 161-163 (1999).
63. Brinkmann, A.O.*, et al.* The human androgen receptor: domain structure, genomic organization and regulation of expression. *J Steroid Biochem* **34**, 307-310 (1989).
64. Spencer, J.A.*, et al.* The androgen receptor gene is located on a highly conserved region of the X chromosomes of marsupial and monotreme as well as eutherian mammals. *J Hered* **82**, 134-139 (1991).
65. Casto, A.M., Henn, B.M., Kidd, J.M., Bustamante, C.D. & Feldman, M.W. A Tale of Two Haplotypes: The EDA2R/AR Intergenic Region Is the Most Divergent Genomic Segment between Africans and East Asians in the Human Genome. *Hum Biol* **84**, 641-694 (2012).
66. Lambert, C.A.*, et al.* Highly Punctuated Patterns of Population Structure on the X Chromosome and Implications for African Evolutionary History. *American journal of human genetics* **86**, 34-44 (2010).
67. Keinan, A., Mullikin, J.C., Patterson, N. & Reich, D. Accelerated genetic drift on chromosome X during the human dispersal out of Africa. *Nature genetics* **41**, 66-70 (2009).
68. Sabeti, P.C.*, et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-U912 (2007).
69. Hillmer, A.M.*, et al.* Recent positive selection of a human androgen receptor/ectodysplasin A2 receptor haplotype and its relationship to male pattern baldness. *Hum Genet* **126**, 255-264 (2009).
70. Hughes, I.A.*, et al.* Androgen insensitivity syndrome. *Lancet* **380**, 1419-1428 (2012).

71. Edwards, A., Hammond, H.A., Jin, L., Caskey, C.T. & Chakraborty, R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**, 241-253 (1992).
72. Mubiru, J.N.*, et al.* Androgen receptor CAG repeat polymorphism in males of six non-human primate species. *Journal of medical primatology* **41**, 67-70 (2012).
73. Choong, C.S., Kemppainen, J.A. & Wilson, E.M. Evolution of the primate androgen receptor: a structural basis for disease. *Journal of molecular evolution* **47**, 334-342 (1998).
74. Hong, K.W.*, et al.* Comparison of androgen receptor CAG and GGN repeat length polymorphism in humans and apes. *Primates* **47**, 248-254 (2006).
75. Ackerman, C.M.*, et al.* Ethnic Variation in Allele Distribution of the Androgen Receptor (AR) (CAG)(n) Repeat. *J Androl* **33**, 210-215 (2012).
76. Kittles, R.A.*, et al.* Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: implications for prostate cancer risk. *Hum Genet* **109**, 253-261 (2001).
77. La Spada, A.R.*, et al.* Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Nature genetics* **2**, 301-304 (1992).
78. Kennedy, W.R., Alter, M. & Sung, J.H. Progressive proximal spinal and bulbar muscular atrophy of late onset. A sex-linked recessive trait. *Neurology* **18**, 671-680 (1968).
79. Fratta, P.*, et al.* Correlation of clinical and molecular features in spinal bulbar muscular atrophy. *Neurology* **82**, 2077-2084 (2014).
80. Rhodes, L.E.*, et al.* Clinical features of spinal and bulbar muscular atrophy. *Brain* **132**, 3242-3251 (2009).
81. Chamberlain, N.L., Driver, E.D. & Miesfeld, R.L. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic acids research* **22**, 3181-3186 (1994).
82. Tut, T.G., Ghadessy, F.J., Trifiro, M.A., Pinsky, L. & Yong, E.L. Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. *The Journal of clinical endocrinology and metabolism* **82**, 3777-3782 (1997).
83. Beilin, J., Ball, E.M., Favaloro, J.M. & Zajac, J.D. Effect of the androgen receptor CAG repeat polymorphism on transcriptional activity: specificity in prostate and non-prostate cell lines. *Journal of molecular endocrinology* **25**, 85-96 (2000).
84. Ryan, C.P. & Crespi, B.J. Androgen receptor polyglutamine repeat number: models of selection and disease susceptibility. *Evolutionary applications* **6**, 180-196 (2013).
85. Davis-Dao, C.A., Tuazon, E.D., Sokol, R.Z. & Cortessis, V.K. Male infertility and variation in CAG repeat length in the androgen receptor

gene: a meta-analysis. *The Journal of clinical endocrinology and metabolism* **92**, 4319-4326 (2007).
86. Asatiani, K., von Eckardstein, S., Simoni, M., Gromoll, J. & Nieschlag, E. CAG repeat length in the androgen receptor gene affects the risk of male infertility. *International journal of andrology* **26**, 255-261 (2003).
87. Dowsing, A.T.*, et al.* Linkage between male infertility and trinucleotide repeat expansion in the androgen-receptor gene. *Lancet* **354**, 640-643 (1999).
88. Giwercman, Y.L.*, et al.* No association between the androgen receptor gene CAG repeat and impaired sperm production in Swedish men. *Clin Genet* **54**, 435-436 (1998).
89. Garolla, A.*, et al.* Molecular analysis of the androgen receptor gene in testicular cancer. *Endocr Relat Cancer* **12**, 645-655 (2005).
90. Giwercman, A.*, et al.* Linkage between androgen receptor gene CAG trinucleotide repeat length and testicular germ cell cancer histological type and clinical stage. *Eur J Cancer* **40**, 2152-2158 (2004).
91. Huhtaniemi, I.T.*, et al.* Increased Estrogen Rather Than Decreased Androgen Action Is Associated with Longer Androgen Receptor CAG Repeats. *J Clin Endocr Metab* **94**, 277-284 (2009).
92. Crabbe, P.*, et al.* Part of the interindividual variation in serum testosterone levels in healthy men reflects differences in androgen sensitivity and feedback set point: contribution of the androgen receptor polyglutamine tract polymorphism. *The Journal of clinical endocrinology and metabolism* **92**, 3604-3610 (2007).
93. Lubahn, D.B.*, et al.* The human androgen receptor: complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate. *Mol Endocrinol* **2**, 1265-1275 (1988).
94. Lundin, K.B., Giwercman, A., Richthoff, J., Abrahamsson, P.A. & Giwercman, Y.L. No association between mutations in the human androgen receptor GGN repeat and inter-sex conditions. *Molecular human reproduction* **9**, 375-379 (2003).
95. Aschim, E.L.*, et al.* Linkage between cryptorchidism, hypospadias, and GGN repeat length in the androgen receptor gene. *The Journal of clinical endocrinology and metabolism* **89**, 5105-5109 (2004).
96. Brokken, L.J.*, et al.* Non-linear association between androgen receptor CAG and GGN repeat lengths and reproductive parameters in fertile European and Inuit men. *Molecular and cellular endocrinology* (2013).
97. Lundin, K.B., Giwercman, Y.L., Rylander, L., Hagmar, L. & Giwercman, A. Androgen receptor gene GGN repeat length and reproductive characteristics in young Swedish men. *European journal of endocrinology / European Federation of Endocrine Societies* **155**, 347-354 (2006).
98. Karlson, E.W.*, et al.* A prospective study of androgen levels, hormone-related genes and risk of rheumatoid arthritis. *Arthritis research & therapy* **11**, R97 (2009).

99. Huhtaniemi, I.T., *et al.* Effect of polymorphisms in selected genes involved in pituitary-testicular function on reproductive hormones and phenotype in aging men. *The Journal of clinical endocrinology and metabolism* **95**, 1898-1908 (2010).
100. Hillmer, A.M., *et al.* Genetic variation in the human androgen receptor gene is the major determinant of common early-onset androgenetic alopecia. *American journal of human genetics* **77**, 140-148 (2005).
101. Ellis, J.A., Stebbing, M. & Harrap, S.B. Polymorphism of the androgen receptor gene is associated with male pattern baldness. *J Invest Dermatol* **116**, 452-455 (2001).
102. Ellis, J.A., *et al.* Baldness and the androgen receptor: the AR polyglycine repeat polymorphism does not confer susceptibility to androgenetic alopecia. *Hum Genet* **121**, 451-457 (2007).
103. Zhuo, F.L., *et al.* Androgen receptor gene polymorphisms and risk for androgenetic alopecia: a meta-analysis. *Clinical and experimental dermatology* **37**, 104-111 (2012).
104. Beilin, J., *et al.* A case-control study of the androgen receptor gene CAG repeat polymorphism in Australian prostate carcinoma subjects. *Cancer* **92**, 941-949 (2001).
105. Bratt, O., *et al.* CAG repeat length in the androgen receptor gene is related to age at diagnosis of prostate cancer and response to endocrine therapy, but not to prostate cancer risk. *British journal of cancer* **81**, 672-676 (1999).
106. Chang, B.L. Polymorphic GGC repeats in the androgen receptor gene are associated with hereditary and sporadic prostate cancer risk. *Hum Genet* **110**, 122-129 (2002).
107. Chen, C. Androgen receptor polymorphisms and the incidence of prostate cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **11**, 1033-1040 (2002).
108. Correa-Cerro, L. (CAG)nCAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population. *Eur J Hum Genet* **7**, 357-362 (1999).
109. Edwards, S.M. Androgen receptor polymorphisms: association with prostate cancer risk, relapse and overall survival. *International journal of cancer. Journal international du cancer* **84**, 458-465 (1999).
110. Giovannucci, E. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci Usa* **94**, 3320-3323 (1997).
111. Gsur, A. Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis* **23**, 1647-1651 (2002).
112. Ingles, S.A. Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *Journal of the National Cancer Institute* **89**, 166-170 (1997).

113. Irvine, R.A., Yu, M.C., Ross, R.K. & Coetzee, G.A. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer research* **55**, 1937-1940 (1995).
114. Lange, E.M. The polymorphic exon 1 androgen receptor CAG repeat in men with a potential inherited predisposition to prostate cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **9**, 439-442 (2000).
115. Li, C. Difference between Swedish and Japanese men in the association between AR CAG repeats and prostate cancer suggesting a susceptibility-modifying locus overlapping the androgen receptor gene. *Int J Mol Med* **11**, 529-533 (2003).
116. Lindstrom, S. Systematic replication study of reported genetic associations in prostate cancer: Strong support for genetic variation in the androgen pathway. *Prostate* **66**, 1729-1743 (2006).
117. Mishra, D., Thangaraj, K., Mandhani, A., Kumar, A. & Mittal, R. Is reduced CAG repeat length in androgen receptor gene associated with risk of prostate cancer in Indian population? *Clin Genet* **68**, 55-60 (2005).
118. Mittal, R.D., Mishra, D.K., Thangaraj, K., Singh, R. & Mandhani, A. Is there an inter-relationship between prostate specific antigen, kallikrein-2 and androgen receptor gene polymorphisms with risk of prostate cancer in north Indian population? *Steroids* **72**, 335-341 (2007).
119. Panz, V.R. Tandem CAG repeats of the androgen receptor gene and prostate cancer risk in black and white men. *Endocrine* **15**, 213-216 (2001).
120. Platz, E.A. 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 : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **14**, 1262-1269 (2005).
121. Salinas, C.A., Austin, M.A., Ostrander, E.O. & Stanford, J.L. Polymorphisms in the androgen receptor and the prostate-specific antigen genes and prostate cancer risk. *Prostate* **65**, 58-65 (2005).
122. Sieh, W. Genetic Susceptibility to Prostate Cancer: Prostate-specific Antigen and its Interaction with the Androgen Receptor (United States). *Cancer causes & control : CCC* **17**, 187-197 (2006).
123. Stanford, J.L. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer research* **57**, 1194-1198 (1997).
124. Vijayalakshmi, K. GGN repeat length and GGN/CAG haplotype variations in the androgen receptor gene and prostate cancer risk in south Indian men. *Journal of human genetics* **51**, 998-1005 (2006).

125. Visvanathan, K. Association among an ornithine decarboxylase polymorphism, androgen receptor gene (CAG) repeat length and prostate cancer risk. *The Journal of urology* **171**, 652-655 (2004).
126. Yoo, S. Androgen receptor CAG repeat polymorphism and risk of TMPRSS2:ERG-positive prostate cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **23**, 2027-2031 (2014).
127. Weng, H.*, et al.* Androgen receptor gene polymorphisms and risk of prostate cancer: a meta-analysis. *Sci Rep* **7**, 40554 (2017).
128. Platz, E.A.*, et al.* The androgen receptor gene GGN microsatellite and prostate cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **7**, 379-384 (1998).
129. Lindstrom, S.*, et al.* Germ-line genetic variation in the key androgen-regulating genes androgen receptor, cytochrome P450, and steroid-5-alpha-reductase type 2 is important for prostate cancer development. *Cancer research* **66**, 11077-11083 (2006).
130. Hayes, V.M.*, et al.* The E211 G>A androgen receptor polymorphism is associated with a decreased risk of metastatic prostate cancer and androgenetic alopecia. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **14**, 993-996 (2005).
131. Freedman, M.L.*, et al.* Systematic evaluation of genetic variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study. *American journal of human genetics* **76**, 82-90 (2005).
132. Lindstrom, S.*, et al.* 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. *J Clin Endocr Metab* **95**, E121-E127 (2010).
133. Tilley, W.D., Marcelli, M. & McPhaul, M.J. Expression of the human androgen receptor gene utilizes a common promoter in diverse human tissues and cell lines. *The Journal of biological chemistry* **265**, 13776-13781 (1990).
134. Song, C.S.*, et al.* A distal activation domain is critical in the regulation of the rat androgen receptor gene promoter. *The Biochemical journal* **294 (Pt 3)**, 779-784 (1993).
135. Faber, P.W.*, et al.* Characterization of the human androgen receptor transcription unit. *The Journal of biological chemistry* **266**, 10743-10749 (1991).
136. Faber, P.W., van Rooij, H.C., Schipper, H.J., Brinkmann, A.O. & Trapman, J. Two different, overlapping pathways of transcription initiation are active on the TATA-less human androgen receptor promoter.

The role of Sp1. *The Journal of biological chemistry* **268**, 9296-9301 (1993).
137. Wolf, D.A., Herzinger, T., Hermeking, H., Blaschke, D. & Horz, W. Transcriptional and posttranscriptional regulation of human androgen receptor expression by androgen. *Mol Endocrinol* **7**, 924-936 (1993).
138. Krongrad, A., Wilson, C.M., Wilson, J.D., Allman, D.R. & McPhaul, M.J. Androgen increases androgen receptor protein while decreasing receptor mRNA in LNCaP cells. *Molecular and cellular endocrinology* **76**, 79-88 (1991).
139. Yeap, B.B., Wilce, J.A. & Leedman, P.J. The androgen receptor mRNA. *Bioessays* **26**, 672-682 (2004).
140. Jasinska, A*., et al.* Structures of trinucleotide repeats in human transcripts and their functional implications. *Nucleic acids research* **31**, 5463-5468 (2003).
141. de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K. & Krzyzosiak, W.J. Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. *Nucleic acids research* **39**, 3852-3863 (2011).
142. Nalavade, R., Griesche, N., Ryan, D.P., Hildebrand, S. & Krauss, S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death Dis* **4**, e752 (2013).
143. Ding, D., Xu, L., Menon, M., Reddy, G.P. & Barrack, E.R. Effect of a short CAG (glutamine) repeat on human androgen receptor function. *Prostate* **58**, 23-32 (2004).
144. Ding, D., Xu, L., Menon, M., Reddy, G.P. & Barrack, E.R. Effect of GGC (glycine) repeat length polymorphism in the human androgen receptor on androgen action. *Prostate* **62**, 133-139 (2005).
145. Pratt, W.B. & Toft, D.O. Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocrine reviews* **18**, 306-360 (1997).
146. Prescott, J. & Coetzee, G.A. Molecular chaperones throughout the life cycle of the androgen receptor. *Cancer letters* **231**, 12-19 (2006).
147. Gioeli, D. & Paschal, B.M. Post-translational modification of the androgen receptor. *Molecular and cellular endocrinology* **352**, 70-78 (2012).
148. Gelmann, E.P. Molecular biology of the androgen receptor. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **20**, 3001-3015 (2002).
149. Shen, H.C. & Coetzee, G.A. The androgen receptor: unlocking the secrets of its unique transactivation domain. *Vitamins and hormones* **71**, 301-319 (2005).
150. Jenster, G., van der Korput, H.A., Trapman, J. & Brinkmann, A.O. Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *The Journal of biological chemistry* **270**, 7341-7346 (1995).

151. Simental, J.A., Sar, M., Lane, M.V., French, F.S. & Wilson, E.M. Transcriptional activation and nuclear targeting signals of the human androgen receptor. *The Journal of biological chemistry* **266**, 510-518 (1991).
152. Chamberlain, N.L., Whitacre, D.C. & Miesfeld, R.L. Delineation of two distinct type 1 activation functions in the androgen receptor amino-terminal domain. *The Journal of biological chemistry* **271**, 26772-26778 (1996).
153. McEwan, I.J. Molecular mechanisms of androgen receptor-mediated gene regulation: structure-function analysis of the AF-1 domain. *Endocr Relat Cancer* **11**, 281-293 (2004).
154. Claessens, F.*, et al.* Diverse roles of androgen receptor (AR) domains in AR-mediated signaling. *Nucl Recept Signal* **6**, e008 (2008).
155. Steketee, K.*, et al.* Amino acids 3-13 and amino acids in and flanking the 23FxxLF27 motif modulate the interaction between the N-terminal and ligand-binding domain of the androgen receptor. *Eur J Biochem* **269**, 5780-5791 (2002).
156. Freedman, L.P.*, et al.* The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature* **334**, 543-546 (1988).
157. Bennett, N.C., Gardiner, R.A., Hooper, J.D., Johnson, D.W. & Gobe, G.C. Molecular cell biology of androgen receptor signalling. *Int J Biochem Cell Biol* **42**, 813-827 (2010).
158. Matias, P.M.*, et al.* Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *The Journal of biological chemistry* **275**, 26164-26171 (2000).
159. Schaufele, F.*, et al.* The structural basis of androgen receptor activation: intramolecular and intermolecular amino-carboxy interactions. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9802-9807 (2005).
160. Socialstyrelsen. Statistics on Cancer Incidence 2015. **2017-01-20**(2017).
161. Adams, J. The case of scirrhous of the prostate gland with corresponding affliction of the lymphatic glands in the lumbar region and in the pelvis. *Lancet* **1**(1853).
162. Grevin, G., Lagier, R. & Baud, C.A. Metastatic carcinoma of presumed prostatic origin in cremated bones from the first century A.D. *Virchows Arch* **431**, 211-214 (1997).
163. Schultz, M., Parzinger, H., Posdnjakov, D.V., Chikisheva, T.A. & Schmidt-Schultz, T.H. Oldest known case of metastasizing prostate carcinoma diagnosed in the skeleton of a 2,700-year-old Scythian king from Arzhan (Siberia, Russia). *International journal of cancer. Journal international du cancer* **121**, 2591-2595 (2007).
164. Merczi, M.*, et al.* Skeletal metastatic carcinomas from the Roman period (1st to 5th Century AD) in Hungary. *Pathobiology* **81**, 100-111 (2014).

165. Prates, C., Sousa, S., Oliveira, C. & Ikram, S. Prostate metastatic bone cancer in an Egyptian Ptolemaic mummy, a proposed radiological diagnosis. *International Journal of Paleopathology* **1**, 98-103 (2011).
166. Denmeade, S.R. & Isaacs, J.T. A history of prostate cancer treatment. *Nature reviews. Cancer* **2**, 389-396 (2002).
167. Gutman, A.B. & Gutman, E.B. AN " ACID " PHOSPHATASE OCCURRING IN THE SERUM OF PATIENTS WITH METASTASIZING CARCINOMA OF THE PROSTATE GLAND. *J Clin Invest* **17**, 473-478 (1938).
168. Huggins, C. & Hodges, C.V. Studies on prostatic cancer: I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. 1941. *The Journal of urology* **168**, 9-12 (2002).
169. Lowe, F.C. & Trauzzi, S.J. Prostatic acid phosphatase in 1993. Its limited clinical utility. *The Urologic clinics of North America* **20**, 589-595 (1993).
170. Stamey, T.A.*, et al.* Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *The New England journal of medicine* **317**, 909-916 (1987).
171. Nadji, M.*, et al.* Prostatic-specific antigen: an immunohistologic marker for prostatic neoplasms. *Cancer* **48**, 1229-1232 (1981).
172. Randall, V.A. Role of 5 alpha-reductase in health and disease. *Baillieres Clin Endocrinol Metab* **8**, 405-431 (1994).
173. Kavanagh, J.P. Sodium, potassium, calcium, magnesium, zinc, citrate and chloride content of human prostatic and seminal fluid. *J Reprod Fertil* **75**, 35-41 (1985).
174. Verze, P., Cai, T. & Lorenzetti, S. The role of the prostate in male fertility, health and disease. *Nature reviews. Urology* **13**, 379-386 (2016).
175. Lee, C.H., Akin-Olugbade, O. & Kirschenbaum, A. Overview of prostate anatomy, histology, and pathology. *Endocrinol Metab Clin North Am* **40**, 565-575, viii-ix (2011).
176. McNeal, J.E. The zonal anatomy of the prostate. *Prostate* **2**, 35-49 (1981).
177. Roehrborn, C.G. Benign Prostatic Hyperplasia: An Overview. *Reviews in Urology* **7**, S3-S14 (2005).
178. Patterson, S.G., Balducci, L. & Pow-Sang, J.M. Controversies surrounding androgen deprivation for prostate cancer. *Cancer Control* **9**, 315-325 (2002).
179. Uo, T., Plymate, S.R. & Sprenger, C.C. Allosteric alterations in the androgen receptor and activity in prostate cancer. *Endocrine-Related Cancer* **24**, R335-R348 (2017).
180. Niu, Y.*, et al.* Androgen receptor is a tumor suppressor and proliferator in prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 12182-12187 (2008).
181. Kyprianou, N., English, H.F. & Isaacs, J.T. Programmed cell death during regression of PC-82 human prostate cancer following androgen ablation. *Cancer research* **50**, 3748-3753 (1990).

182. Webber, M.M., Bello, D. & Quader, S. Immortalized and tumorigenic adult human prostatic epithelial cell lines: characteristics and applications. Part I. Cell markers and immortalized nontumorigenic cell lines. *Prostate* **29**, 386-394 (1996).
183. Schwab, T.S.*, et al.* Phenotypic characterization of immortalized normal and primary tumor-derived human prostate epithelial cell cultures. *Prostate* **44**, 164-171 (2000).
184. Isaacs, J.T. The biology of hormone refractory prostate cancer. Why does it develop? *The Urologic clinics of North America* **26**, 263-273 (1999).
185. Denmeade, S.R., Lin, X.S. & Isaacs, J.T. Role of programmed (apoptotic) cell death during the progression and therapy for prostate cancer. *Prostate* **28**, 251-265 (1996).
186. Feldman, B.J. & Feldman, D. The development of androgen-independent prostate cancer. *Nat Rev Cancer* **1**, 34-45 (2001).
187. Gann, P.H., Hennekens, C.H., Ma, J., Longcope, C. & Stampfer, M.J. Prospective study of sex hormone levels and risk of prostate cancer. *Journal of the National Cancer Institute* **88**, 1118-1126 (1996).
188. Pierorazio, P.M.*, et al.* Serum testosterone is associated with aggressive prostate cancer in older men: results from the Baltimore Longitudinal Study of Aging. *BJU international* **105**, 824-829 (2010).
189. Shaneyfelt, T., Husein, R., Bubley, G. & Mantzoros, C.S. Hormonal predictors of prostate cancer: a meta-analysis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **18**, 847-853 (2000).
190. Shin, B.S.*, et al.* Is a decreased serum testosterone level a risk factor for prostate cancer? A cohort study of korean men. *Korean J Urol* **51**, 819-823 (2010).
191. Isbarn, H.*, et al.* Testosterone and prostate cancer: revisiting old paradigms. *European urology* **56**, 48-56 (2009).
192. Endogenous, H.*, et al.* Endogenous sex hormones and prostate cancer: a collaborative analysis of 18 prospective studies. *Journal of the National Cancer Institute* **100**, 170-183 (2008).
193. Muller, R.L.*, et al.* Serum testosterone and dihydrotestosterone and prostate cancer risk in the placebo arm of the Reduction by Dutasteride of Prostate Cancer Events trial. *European urology* **62**, 757-764 (2012).
194. Zlotta, A.R.*, et al.* Prevalence of prostate cancer on autopsy: cross-sectional study on unscreened Caucasian and Asian men. *Journal of the National Cancer Institute* **105**, 1050-1058 (2013).
195. Breslow, N.*, et al.* Latent carcinoma of prostate at autopsy in seven areas. The International Agency for Research on Cancer, Lyons, France. *International journal of cancer. Journal international du cancer* **20**, 680-688 (1977).
196. Yatani, R.*, et al.* Geographic pathology of latent prostatic carcinoma. *International journal of cancer. Journal international du cancer* **29**, 611-616 (1982).

197. Morgentaler, A. & Traish, A.M. Shifting the paradigm of testosterone and prostate cancer: the saturation model and the limits of androgen-dependent growth. *European urology* **55**, 310-320 (2009).
198. Gann, P.H. Risk factors for prostate cancer. *Rev Urol* **4 Suppl 5**, S3-S10 (2002).
199. Gronberg, H., Damber, L. & Damber, J.E. Familial prostate cancer in Sweden. A nationwide register cohort study. *Cancer* **77**, 138-143 (1996).
200. Johns, L.E. & Houlston, R.S. A systematic review and meta-analysis of familial prostate cancer risk. *BJU international* **91**, 789-794 (2003).
201. Cavanagh, H. & Rogers, K.M.A. The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers. *Hereditary Cancer in Clinical Practice* **13**, 16 (2015).
202. Nwosu, V., Carpten, J., Trent, J.M. & Sheridan, R. Heterogeneity of genetic alterations in prostate cancer: evidence of the complex nature of the disease. *Human molecular genetics* **10**, 2313-2318 (2001).
203. Hsing, A.W., Tsao, L. & Devesa, S.S. International trends and patterns of prostate cancer incidence and mortality. *International journal of cancer. Journal international du cancer* **85**, 60-67 (2000).
204. Olsson, A.Y., Bjartell, A., Lilja, H. & Lundwall, A. Expression of prostate-specific antigen (PSA) and human glandular kallikrein 2 (hK2) in ileum and other extraprostatic tissues. *International journal of cancer. Journal international du cancer* **113**, 290-297 (2005).
205. Cleutjens, K.B., van Eekelen, C.C., van der Korput, H.A., Brinkmann, A.O. & Trapman, J. Two androgen response regions cooperate in steroid hormone regulated activity of the prostate-specific antigen promoter. *The Journal of biological chemistry* **271**, 6379-6388 (1996).
206. Cleutjens, K.B.*, et al.* An androgen response element in a far upstream enhancer region is essential for high, androgen-regulated activity of the prostate-specific antigen promoter. *Mol Endocrinol* **11**, 148-161 (1997).
207. Schuur, E.R.*, et al.* Prostate-specific antigen expression is regulated by an upstream enhancer. *The Journal of biological chemistry* **271**, 7043-7051 (1996).
208. Wolf, D.A., Schulz, P. & Fittler, F. Transcriptional regulation of prostate kallikrein-like genes by androgen. *Mol Endocrinol* **6**, 753-762 (1992).
209. Young, C.Y.*, et al.* Hormonal regulation of prostate-specific antigen messenger RNA in human prostatic adenocarcinoma cell line LNCaP. *Cancer research* **51**, 3748-3752 (1991).
210. Mizokami, A.*, et al.* Understanding prostate-specific antigen dynamics in monitoring metastatic castration-resistant prostate cancer: implications for clinical practice. *Asian journal of andrology* **19**, 143-148 (2017).
211. Lilja, H., Abrahamsson, P.A. & Lundwall, A. Semenogelin, the predominant protein in human semen. Primary structure and identification of closely related proteins in the male accessory sex glands and on the spermatozoa. *The Journal of biological chemistry* **264**, 1894-1900 (1989).

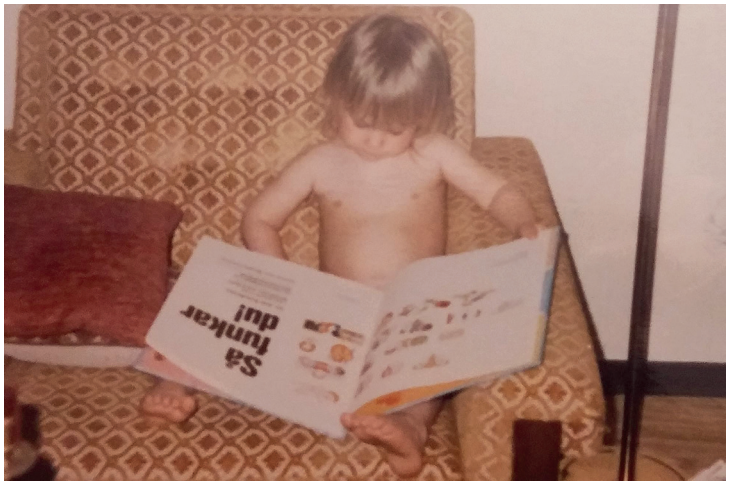
212. Malm, J. & Lilja, H. Biochemistry of prostate specific antigen, PSA. *Scandinavian journal of clinical and laboratory investigation. Supplementum* **221**, 15-22 (1995).
213. Irani, J.*, et al.* Inflammation in benign prostatic hyperplasia: correlation with prostate specific antigen value. *The Journal of urology* **157**, 1301-1303 (1997).
214. Gumus, B.H.*, et al.* Does asymptomatic inflammation increase PSA? A histopathological study comparing benign and malignant tissue biopsy specimens. *Int Urol Nephrol* **36**, 549-553 (2004).
215. Grossklaus, D.J., Shappell, S.B., Gautam, S., Smith, J.A., Jr. & Cookson, M.S. Ratio of free-to-total prostate specific antigen correlates with tumor volume in patients with increased prostate specific antigen. *The Journal of urology* **165**, 455-458 (2001).
216. el-Shirbiny, A.M., Nilson, T. & Pawar, H.N. Serum prostate-specific antigen: hourly change/24 hours compared with prostatic acid phosphatase. *Urology* **35**, 88-92 (1990).
217. Bogdanowicz, J.F., Bentvelsen, F.M., Oosterom, R. & Schroeder, F.H. Evaluation of prostate-specific antigen and prostatic acid phosphatase in untreated prostatic carcinoma and benign prostatic hyperplasia. *Scand J Urol Nephrol Suppl* **138**, 97-103 (1991).
218. Seamonds, B.*, et al.* Evaluation of prostate-specific antigen and prostatic acid phosphatase as prostate cancer markers. *Urology* **28**, 472-479 (1986).
219. Small, E.J. Prostate cancer: who to screen, and what the results mean. *Geriatrics* **48**, 28-30, 35-28 (1993).
220. Tawfik, A. Prostate-Specific Antigen (PSA)-Based Population Screening for Prostate Cancer: An Economic Analysis. *Ontario health technology assessment series* **15**, 1-37 (2015).
221. Pron, G. Prostate-Specific Antigen (PSA)-Based Population Screening for Prostate Cancer: An Evidence-Based Analysis. *Ontario health technology assessment series* **15**, 1-64 (2015).
222. Welch, H.G. & Black, W.C. Overdiagnosis in cancer. *Journal of the National Cancer Institute* **102**, 605-613 (2010).
223. Sandhu, G.S. & Andriole, G.L. Overdiagnosis of prostate cancer. *Journal of the National Cancer Institute. Monographs* **2012**, 146-151 (2012).
224. Bokhorst, L.P.*, et al.* Prostate-specific antigen-based prostate cancer screening: reduction of prostate cancer mortality after correction for nonattendance and contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer. *European urology* **65**, 329-336 (2014).
225. Bansal, A.*, et al.* Heritability of prostate-specific antigen and relationship with zonal prostate volumes in aging twins. *The Journal of clinical endocrinology and metabolism* **85**, 1272-1276 (2000).
226. Kim, J.S.*, et al.* Prostate-Specific Antigen fluctuation: what does it mean in diagnosis of prostate cancer? *International braz j urol : official journal of the Brazilian Society of Urology* **41**, 258-264 (2015).

227. Park, Y.H., *et al.* Prostate cancer detection rate in patients with fluctuating prostate-specific antigen levels on the repeat prostate biopsy. *Prostate international* **2**, 26-30 (2014).
228. Ankerst, D.P. & Thompson, I.M. Sensitivity and specificity of prostate-specific antigen for prostate cancer detection with high rates of biopsy verification. *Archivio italiano di urologia, andrologia : organo ufficiale [di] Societa italiana di ecografia urologica e nefrologica* **78**, 125-129 (2006).
229. Luboldt, H.-J., Schindler, J.F. & Rübben, H. Age-Specific Reference Ranges for Prostate-Specific Antigen as a Marker for Prostate Cancer. *EAU-EBU Update Series* **5**, 38-48 (2007).
230. Gronberg, H., *et al.* Prostate cancer screening in men aged 50-69 years (STHLM3): a prospective population-based diagnostic study. *Lancet Oncol* **16**, 1667-1676 (2015).
231. Coetzee, G. & Irvine, R. Size of the androgen receptor CAG repeat and prostate cancer: Does it matter? *Journal of Clinical Oncology* **20**, 3572-3573 (2002).
232. Krausz, C. An encore for the repeats: new insights into an old genetic variant. *The Journal of clinical endocrinology and metabolism* **97**, 764-767 (2012).
233. Lee, D.M., *et al.* Cohort Profile: The European Male Ageing Study. *International journal of epidemiology* (2012).
234. Berglund, G., Elmstahl, S., Janzon, L. & Larsson, S.A. The Malmo Diet and Cancer Study. Design and feasibility. *Journal of internal medicine* **233**, 45-51 (1993).
235. Ruhayel, Y., *et al.* Male infertility and prostate cancer risk: a nested case-control study. *Cancer causes & control : CCC* **21**, 1635-1643 (2010).
236. Orwoll, E., *et al.* Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. *Contemporary clinical trials* **26**, 569-585 (2005).
237. Brokken, L.J., *et al.* Association between polymorphisms in the aryl hydrocarbon receptor repressor gene and disseminated testicular germ cell cancer. *Front Endocrinol (Lausanne)* **4**, 4 (2013).
238. Richthoff, J., Rylander, L., Hagmar, L., Malm, J. & Giwercman, A. Higher sperm counts in Southern Sweden compared with Denmark. *Hum Reprod* **17**, 2468-2473 (2002).
239. Limer, K.L., *et al.* Genetic variation in sex hormone genes influences heel ultrasound parameters in middle-aged and elderly men: results from the European Male Aging Study (EMAS). *J Bone Miner Res* **24**, 314-323 (2009).
240. Eriksson, J., *et al.* Causal relationship between obesity and serum testosterone status in men: A bi-directional mendelian randomization analysis. *PloS one* **12**, e0176277 (2017).

241. Vastermark, A., *et al.* Polymorphic variation in the androgen receptor gene: association with risk of testicular germ cell cancer and metastatic disease. *Eur J Cancer* **47**, 413-419 (2011).
242. Ruhayel, Y. Lund University (2012).
243. Carlstrom, K., Eriksson, A., Stege, R. & Rannevik, G. Relationship between Serum Testosterone and Sex Hormone-Binding Globulin in Adult Men with Intact or Absent Gonadal-Function. *International journal of andrology* **13**, 67-73 (1990).
244. Machiela, M.J. & Chanock, S.J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555-3557 (2015).
245. Kent, W.J., *et al.* The human genome browser at UCSC. *Genome Res* **12**, 996-1006 (2002).
246. Giwercman, A., Richiardi, L., Kaijser, M., Ekbom, A. & Akre, O. Reduced risk of prostate cancer in men who are childless as compared to those who have fathered a child: a population based case-control study. *International journal of cancer. Journal international du cancer* **115**, 994-997 (2005).
247. Jorgensen, K.T. & Frisch, M. Linking fatherhood to prostate cancer risk. *Future oncology* **4**, 137-140 (2008).
248. Jorgensen, K.T., Pedersen, B.V., Johansen, C. & Frisch, M. Fatherhood status and prostate cancer risk. *Cancer* **112**, 919-923 (2008).
249. Andersson, A.M., Jorgensen, N., Frydelund-Larsen, L., Rajpert-De Meyts, E. & Skakkebaek, N.E. Impaired Leydig cell function in infertile men: a study of 357 idiopathic infertile men and 318 proven fertile controls. *The Journal of clinical endocrinology and metabolism* **89**, 3161-3167 (2004).
250. Muraleedharan, V. & Jones, T.H. Testosterone and mortality. *Clin Endocrinol (Oxf)* **81**, 477-487 (2014).
251. Hackett, G. An update on the role of testosterone replacement therapy in the management of hypogonadism. *Therapeutic advances in urology* **8**, 147-160 (2016).
252. Xue, W.M., *et al.* Genetic determinants of serum prostate-specific antigen levels in healthy men from a multiethnic cohort. *Cancer Epidem Biomar* **10**, 575-579 (2001).
253. Giwercman, Y.L., *et al.* Androgen receptor CAG repeat length correlates with semen PSA levels in adolescence. *Prostate* **59**, 227-233 (2004).
254. Schatzl, G., *et al.* Association of polymorphisms within androgen receptor, 5alpha-reductase, and PSA genes with prostate volume, clinical parameters, and endocrine status in elderly men. *Prostate* **52**, 130-138 (2002).
255. Henderson, R.J., *et al.* Prostate-specific antigen (PSA) and PSA density: racial differences in men without prostate cancer. *Journal of the National Cancer Institute* **89**, 134-138 (1997).

256. Bellaousov, S., Reuter, J.S., Seetin, M.G. & Mathews, D.H. RNAstructure: Web servers for RNA secondary structure prediction and analysis. *Nucleic acids research* **41**, W471-474 (2013).
257. Presnyak, V.*, et al.* Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111-1124 (2015).
258. Duan, J.*, et al.* Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Human molecular genetics* **12**, 205-216 (2003).
259. Capon, F.*, et al.* A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Human molecular genetics* **13**, 2361-2368 (2004).
260. Nackley, A.G.*, et al.* Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* **314**, 1930-1933 (2006).
261. Plotkin, J.B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature reviews. Genetics* **12**, 32-42 (2011).
262. Chen, H.*, et al.* Decreased hephaestin activity in the intestine of copper-deficient mice causes systemic iron deficiency. *J Nutr* **136**, 1236-1241 (2006).

Original publications



Magdalena Bentmar Holgersson holds a MSc in Molecular Biology from Lund University, Sweden, since 2012. The focus of her thesis is the relationship between polymorphisms in the androgen receptor, testosterone concentrations and risk of prostate cancer. The main results are the identification of a common androgen receptor haplotype with a reduced risk of prostate cancer but with higher PSA concentrations and the increased risk of all-cause mortality in younger men with low testosterone concentrations.

Printed by Media-Tryck, Lund 2017 NORDIC ECOLABEL, 3041 0903

LUND UNIVERSITY
Faculty of Medicine

Department of Translational Medicine

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2017:149
ISBN 978-91-7619-532-1
ISSN 1652-8220