

https://dspace.org/

Genetic analysis of Mild Androgen Insensitivity Syndrome (MAIS) and breast cancer in a South African Indian family

Chauhan, Samantha

2015-09-18
http://hdl.handle.net/10539/18696

Downloaded from a DSpace repository. https://dspace.org/

GENETIC ANALYSIS OF MILD ANDROGEN INSENSITIVITY SYNDROME (MAIS) AND BREAST CANCER IN A SOUTH AFRICAN INDIAN FAMILY

Samatha Chauhan

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree

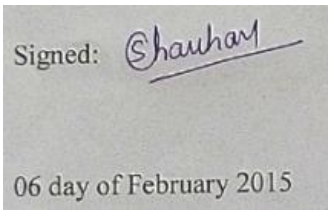
of

Doctor of Philosophy

Johannesburg, February 2015

Declaration

I, Samatha Chauhan declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.



For my Guruji

Yogiraj-Shri Vethathiri Maharishi

(Balge Samruddhiyondige)

Publications arising from this thesis

A. List of articles in preparation:

1. Variations of Androgen receptor Poly-glutamine and Poly-glycine tract in a family with Mild Androgen Insensitivity syndrome (MAIS) / Gynecomastia and breast cancer. (To be submitted to *American Journal of Medical Genetics*).

Authors: Samatha Chauhan, Caroline Dickens, Therese Dix-Peek, Raquel Duarte and Girish Modi

2. Novel and recurrent androgen receptor gene mutations in a family with MAIS/ gynecomastia and breast cancer patients. (To be submitted to *Human Mutation*).

Authors: Samatha Chauhan, Therese Dix-Peek, Caroline Dickens, Raquel Duarte and Girish Modi

B. Presentations

Abstract accepted at international Conference:

Identification of Novel and Recurrent Androgen Receptor Gene Mutations in a Family with MAIS/Gynecomastia and Breast Cancer Patients.

12th International Symposium on Mutation in the Genome: mutation detection & genome sequencing-detecting variants one-by-one or genome wide 22nd - 26th April 2013, Lake Louise, Canada.

ABSTRACT

Introduction: Androgen Insensitivity Syndrome (AIS) is an X-linked disorder caused by mutations in the androgen receptor (AR) gene. The phenotype is variable and ranges from a complete feminine syndrome to simple gynecomastia. The phenotypes are described in terms of complete, partial and mild forms (CAIS, PAIS and MAIS). We describe novel and previously reported (recurrent) mutations in the AR gene for a family in which segregation of breast cancer (BC) and gynecomastia/MAIS is present. **Methods:** We studied a family of 16 members spanning four generations. Based on the presentation of symptoms, the family was divided into affected, unaffected, and control groups. Seven patients (six males diagnosed with MAIS and one female diagnosed with BC) formed the affected group, four genetically related individuals (two males and two females) formed the unaffected group and five genetically unrelated family members (one male and four females) served as controls. In each of these individuals, PCR amplification, cloning and the sequencing of exon 1 were carried out. Exons 2-8 were sequenced directly after PCR amplification. Exon 1 (CAG)n and (GGN)n repeats were classified according to their length: short **(S)** (n<23), long **(L)** (n>23) and wild type **(WT)** (n=23). **Results: Part 1**-The (CAG)n repeats varied among individuals and generations. In the 2nd generation, the unaffected male was **S** and the control female was **WT**. In the 3rd generation, three affected males were **S**, 2 of the controls were **WT**, one control was **L** and the other **S**. In the 4th generation, the 4 affected individuals were **L**, 1 of the unaffected was **WT** and the other 2 unaffected were **L**. **Part 2**- The (GGN)n variations also differed among

individuals and generations. In the 2nd generation, the unaffected male and the control were **S**. In the 3rd generation, all three affected family members were **S** and among the controls, 1 was **WT**, 1 was **L** and 2 were **S**. In the 4th generation, 3 of the affected were **S** and one was **WT** and among the 3 unaffected, 2 were **S** and one was WT. **Part 3-** 30 unreported (novel) mutations as well as 13 recurrent (previously reported) mutations in exon 1 of the AR gene were identified. 17 novel and 5 reported mutations were identified in the affected group, 8 novel and 5 reported mutations, including one premature stop codon mutation, were identified in the related unaffected group and 7 novel and 4 reported mutations were found in the controls. Of the above-mentioned mutations, four mutations were identified in the activation function-1 (AF-1) domain of exon 1 in 4 members (3 affected: M-2, F-1 and 1 unaffected: F-1) of the family. All the point mutations identified were somatic in nature and were present in heterogeneous form i.e wild and mutant (mixture) as determined by cloning. The analysis of exons 2 through 8 revealed completely WT sequences. **Conclusions:** The (CAG)n and (GGN)n repeat analysis showed an indeterminate association with MAIS and BC in the family. Generation specific patterns of (CAG)n were detected and suggest generation specific modulation of the AR. Novel mutations including AF-1 region mutations were identified in exon 1. The disruption of the AF-1 domain may affect the transactivation activity of the AR.

Acknowledgements

This thesis would have been hard and impossible without the guidance and support of several people.

Foremost, I would sincerely like to thank my supervisor, Dr. Raquel Duarte. I am extremely grateful to her for letting me take this opportunity to work with her. I truly appreciate her brilliant ideas, all the support, encouragement and guidance that I obtained from her for my thesis and my career.

I would also like to thank my other supervisor, Prof. Girish Modi for his support and guidance during this work.

I would like to thank my parents for their infinite love, prayers and blessings throughout my education and for instilling in me the importance of hard work.

I'm also thankful to my husband, Dr. Ranjit Chauhan for his motivation to become involved in scientific research and providing support and encouragement in my work.

I would also like to acknowledge the CANSA (Cancer Association of South Africa) for funding this project; The Health Science Research Office (HSRO), for the travel grant and the Post graduate merit award committee of the University of Witwatersrand for their financial support.

Also, in addition, I would like to thank my friends and lab mates for all of their advice and encouragement. At last, I am extremely grateful to my very sweet girl, Samiran and my Son Shiven for always keeping me sane.

TABLE OF CONTENTS

DECLARATION..... ii

DEDICATION..... iii

PUBLICATIONS ARISING FROM THIS THESIS…..... iv

A. LIST OF ARTICLES IN PREPARATION…..... iv

B. PRESENTATIONS..... iv

ABSTRACT..... v

ACKNOWLEDGEMENTS..... vii

TABLE OF CONTENTS..... viii

LIST OF FIGURES..... xiv

LIST OF TABLES..... xix

LIST OF ABBREVIATIONS..... xx

1. INTRODUCTION..... 1

1.1 Significance of the Androgen Receptor..... 1

1.1.1 Androgen Receptor -Nuclear Receptor/Transcription Factor..... 2

1.2 Androgen regulation and production..... 3

1.2.1 Physiological role of androgens in males and females..... 5

1.3 Androgen Insensitivity Syndrome (AIS)……... 7

1.3.1 Prevalence of Androgen Insensitivity Syndrome..... 8

1.3.2 Complete Androgen Insensitivity Syndrome (CAIS)..... 9

1.3.2a Presentation of CAIS..... 9

1.3.3 Partial Androgen Insensitivity Syndrome (PAIS)..... 10

1.3.3a Presentation of PAIS..... 11

1.3.4 Mild Androgen Insensitivity Syndrome (MAIS)..... 12

1.3.4a Presentation of MAIS..... 12

1.4 Androgen receptor gene variation and AIS..... 14

1.4.1 Somatic mosacism…….. 15

1.5 Breast Cancer..... 17

1.5.1 Prevalence of Breast Cancer..... 18

1.6 Androgen receptor gene variation and breast cancer..... 20

1.7 REVIEW OF ANDROGEN RECEPTOR LITERATURE..... 24

1.7.1 Structure of the androgen receptor gene..... 24

1.7.2 Androgen receptor domains..... 27

1.7.2.1 Transactivation Domain..... 28

1.7.2.2 DNA Binding Domain..... 31

1.7.2.3 Hinge region..... 34

1.7.2.4 Ligand Binding Domain..... 35

1.7.2.4a 3D structure of LBD..... 37

1.8 Mechanism of AR action..... 39

1.9 AIMS……..... 43

2. MATERIAL AND METHODS……..... 44

2.1 Ethical consideration……..... 44

2.2 Family Pedigree included in the study, its 3 components and analysis. 44

2.3 Extraction of genomic DNA from patients……... 47

2.4 PCR amplification of coding region of AR gene...……... 48

2.4.1 The design of primers to amplify exons 1 to 8 of AR gene..... 48

2.4.2 PCR conditions for amplifying exon1 to exon 8 of the AR..... 49

2.4.3 Agarose Gel Electrophoresis..... 52

2.5 Analysis of sequencing data of exons 2-8 of AR gene..... 52

2.6 Analysis of exon 1..... 53

2.6.1 Purification of amplicons of exon 1 (part I and part II)..... 53

2.6.2 Preparation of chemically competent cells..... 54

2.6.3 Ligation reactions..... 55

2.6.4 Transformation of ligation products into

DH5 α cells using heat shock……... 55

2.7 Screening of transformants for inserts..... 56

2.7.1 Plasmid Extraction..... 56

2.8 Checking of positive clones by restriction digestion analysis..... 58

2.9 Sequencing of plasmid DNA of exon 1 (Part I and Part II) and

analysis..... 58

3. RESULTS..... 60

3.1 Amplification of exons 1-8 of androgen receptor gene, sequencing and analysis..... 61

3.2 Cloning of part I and II of exon 1of AR gene..... 65

3.3 Sequencing of plasmid DNA of exon 1(Part I and Part II) and analysis..... 66

3.4 Results of CAG repeat length (poly-glutamine tract variations)..... 67

3.4.1 Defining and dividing CAG repeat length of the AR gene..... 67

3.4.2 Analysis results of poly-glutamine stretch within the exon 1 region of the AR gene..... 70

3.4.3 Detection of variation in the poly-glutamine tract (CAG-repeat)..... 70

3.4.4 (a) poly-glutamine stretch among the 2nd generation..... 71

3.4.5 (b) poly-glutamine stretch among the 3rd generation..... 71

3.4.6 (c) poly-glutamine stretch among the 4th generation..... 72

3.5 Results of GGN repeat length (poly-glycine tract variations)..... 72

3.5.1 Defining and dividing GGN repeat length of the AR gene..... 72

3.5.2 Analysis results of poly-glycine stretch within the exon 1 region of the AR gene..... 76

3.5.3 Generation specific analysis of Poly-glycine tract (GGN-repeat)….. 76

3.5.3 (a) Poly-glycine stretch among the 2nd generation..... 76

3.5.4 (b) Poly-glycine stretch among the 3rd generation..... 76

3.5.5 (c) Poly-glycine stretch among the 4th generation..... 77

4. RESULTS OF AR GENE POINT MUTATIONS…..... 78

4.1 Analysis of sequencing data of all the exons of the AR gene..... 78

4.2 Analysis of exons 2-8 of androgen receptor gene..... 78

4.3 Identification of point mutations in exon 1..... 79

4.3.1 Point mutations identified in affected members..... 83

4.3.2 Point mutations identified in AF-1 domain of exon 1 in affected members..... 83

5. DISCUSSION..... 85

5.1 Poly-glutamine tract variations in androgen receptor gene..... 85

5.2 Poly-glycine tract variations in androgen receptor gene..... 88

5.3 Point mutations in androgen receptor gene..... 89

6. CONCLUSIONS..... 98

6.1 FUTURE DIRECTION..... 102

Appendix A..... 103

Appendix B..... 104

Appendix C..... 113

Appendix D..... 116

REFERENCES..... 139

LIST OF FIGURES

Page

Figure 1. Hypothalamic pituitary testicular axis showing the principal negative feedback mechanism..... 4

Figure 2. Chemical Structure of DHT and Testosterone..... 7

Figure 3. Geographical Prevalence and breast cancer statistics worldwide in 2012..... 19

Figure 4. Structure of the Androgen Receptor gene..... 25

Figure 5. The amino acid sequence of human AR gene..... 26

Figure 6. AR NH2-Transactivating domain amino acid sequence..... 30

Figure 7. AR DNA binding domain amino acid sequence..... 33

Figure 8. AR Hinge region and LBD amino acid sequence..... 36

Figure 9a. The 3D-crystallographic structure of AR LBD..... 37

Figure 9b. Structural organization of AR gene..... 39

Figure 10. Mechanism action of Androgen Receptor cycle..... 42

Figure 11. *Pedigree of family* kindred with *MAIS and BC*..... 46

Figure 11a. *Family Pedigree* figure superimposed with the results of CAG repeat length identified in present 69

study……

Figure 11b. *Family Pedigree* figure superimposed with the results of GGN repeat length identified in present study…… 75

Figure 12. Primer design of exon 1 of AR gene..... 49

Figure 13. The pGEM-T Easy Vector Map..... 113

Figure 14. Sequence and multiple cloning site of the pGEM-T easy vector..... 114

Figure 15. Gel picture of amplification of exon 1 of AR..... 61

Figure 16. Gel Picture of amplification of exon 2 of AR..... 61

Figure 17. Gel Picture of amplification of exon 3 of AR..... 62

Figure 18. Gel Picture of amplification of exon 4 of AR..... 62

Figure 19. Gel Picture of amplification of exon 5 of AR..... 63

Figure 20. Gel Picture of amplification of exon 6 of AR..... 63

Figure 21. Gel Picture of amplification of exon 7 of AR..... 64

Figure 22. Gel Picture of amplification of exon 8 of AR..... 64

Figure 23. Gel picture of restriction mapping of plasmid DNA of part I of exon 1..... 65

Figure 24. Gel picture of restriction mapping of plasmid DNA of part II of exon 1..... 66

Figure 25. Mutations identified in AF-1 domain of exon 1..... 84

Figure 26. Electropherogram of L108P and E113K novel mutations..... 116

Figure 27. Electropherogram of T231I novel mutation..... 116

Figure 28. Electropherogram of Q78R novel mutation..... 117

Figure 29. Electropherogram of E187G novel mutation..... 117

Figure 30. Electropherogram of L345P novel mutation..... 118

Figure 31. Electropherogram of T105I novel mutation..... 118

Figure 32. Electropherogram of C240R novel mutation..... 119

Figure 33. Electropherogram of E32V novel mutation..... 119

Figure 34. Electropherogram of Q360R novel mutation..... 120

Figure 35. Electropherogram of R386C novel mutation..... 120

Figure 36. Electropherogram of N25S novel mutation..... 121

Figure 37. Electropherogram of L347F novel mutation..... 121

Figure 38. Electropherogram of V132A novel mutation..... 122

Figure 39. Electropherogram of V245M novel mutation..... 122

Figure 40. Electropherogram of W527R novel mutation..... 123

Figure 41. Electropherogram of M528V novel mutation..... 123

Figure 42. Electropherogram of E155G reported mutation..... 124

Figure 43. Electropherogram of E289G reported mutation..... 124

Figure 44. Electropherogram of Q445R reported mutation..... 125

Figure 45. Electropherogram of Q456D reported mutation..... 125

Figure 46. Electropherogram of S221F reported mutation..... 126

Figure 47. Electropherogram of P153L novel mutation..... 126

Figure 48. Electropherogram of L389P novel mutation..... 127

Figure 49. Electropherogram of M191T novel mutation..... 127

Figure 50. Electropherogram of Q72R novel mutation..... 128

Figure 51. Electropherogram of Q78R novel mutation..... 128

Figure 52. Electropherogram of G326D novel mutation..... 129

Figure 53. Electropherogram of G173S novel mutation..... 135

Figure 54. Electropherogram of P422S novel mutation..... 130

Figure 55. Electropherogram of T440A reported mutation..... 130

Figure 56. Electropherogram of Q28 Stop codon reported mutation..... 131

Figure 57. Electropherogram of E442K reported mutation..... 131

Figure 58. Electropherogram of Y225C reported mutation..... 132

Figure 59. Electropherogram of G525D reported mutation..... 132

Figure 60. Electropherogram of T82A novel mutation..... 133

Figure 61. Electropherogram of Q79R novel mutation..... 133

Figure 62. Electropherogram of A372T novel mutation..... 134

Figure 63. Electropherogram of V33A novel mutation..... 134

Figure 64. Electropherogram of Q98R novel mutation..... 135

Figure 65. Electropherogram of P49L novel mutation..... 135

Figure 66. Electropherogram of Q79R novel mutation..... 136

Figure 67. Electropherogram of S361G novel mutation..... 136

Figure 68. Electropherogram of Y481C reported mutation..... 137

Figure 69. Electropherogram of Q488R reported mutation..... 137

Figure 70. Electropherogram of E155G reported mutation..... 138

Figure 71. A chart representing amino acid properties..... 115

LIST OF TABLES……. Page

Table 1. Types of androgen response elements (ARE)..... 32

Table 2. List of Primers for exons 1-8 of AR..... 51

Table 3. Ligation reaction setup and condition..... 111

Table 4. Restriction digestion and mapping of purified plasmid DNA by *Not I* enzyme..... 112

Table 5. Variation of Poly Q in 16 members of the family..... 68

Table 5a. Superimpose table on pedigree: Addition and deletion of Poly Q in 16 members of the family..... 69

Table 6. Variation of Poly G in 16 members of the family..... 75

Table 6a. Superimpose table on pedigree: Addition and deletion of Poly G in 16 members of the family..... 77

Table 7. Point mutations identified in affected members of the family..... 81

Table 7a. Point mutations identified in unaffected and control members of the family……..... 82

LIST OF ABBREVIATIONS

AR	Androgen Receptor
ARE	Androgen Response Element
AIS	Androgen Insensitivity Syndrome
AF-1	Activation Function-1
AF-2	Activation Function-2
AF-5	Activation Function-5
aa	Amino acid
BC	Breast Cancer
CAIS	Complete Androgen Insensitivity Syndrome
DBD	DNA Binding Domain
D-box	Distal Box
DHT	Di-hydro-testosterone
DHEA	Dehydroepiandrosterone
DHEA-S	Dehydroepiandrosterone-Sulfate
DSD	Disorder of Sex Development
DNA	Deoxyribonucleic acid
ER	Estrogen Receptor
EDTA	Ethylene diamine tetra acetic acid
EtBr	Ethidium Bromide
FSH	Follicle Stimulating Hormone
FlnA	Cytoskeletal protein filamin –A
GR	Glucocorticoid Receptor
GnRH	Gonadotropin releasing hormone
HSP	Heat Shock Protein
hCG	human Chorionic Gonadotropin
HDAC	Histone De-Acetylases
LBD	Ligand Binding Domain
LB	Luria Bertani

LH	Luteinizing Hormone
MAIS	Mild Androgen Insensitivity Syndrome
MIS	Mullerian inhibiting substance
MR	Minerelo Corticoid Receptor
NTD	Transactivation Domain
NHR	Nuclear Hormone Receptor
NR3C4	Nuclear Receptor subfamily 3, Group C, Member 4
NCoR	Nuclear Receptor Co-Repressor
NLS	Nuclear Localization Signal
NCI	National Cancer Institute
NIH	National Institute of Health
nt	Nucleotide
nts : A	Adenine
G	Guanine
C	Cytosine
T	Thiamine
ORD	Office of Rare Disease
PR	Progesterone Receptor
PAIS	Partial Androgen Insensitivity Syndrome
PCR	Polymerase Chain Reaction
SBMA	Spinal and Bulbar Muscular Atrophy
SMRT	Silencing Mediator for Retinoid and Thyroid
SDS	Sodium Dodecyl Sulphate
T	Testosterone
TAU1	Transcription Activation Unit-1
TAU5	Transcription Activation Unit-5
UTR	Untranslated region
VDR	Vitamin D Receptor

Amino acids code

A	Alanine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glysine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
Y	Tyrosine

Measurement units

˚C	Degree Celsius
kDa	Kilo Dalton
kb	Kilo base
ml	Milliliter
mg	Milligram
µg	Microgram
min	Minute
µM	Micro Molar
M	Molar
MW	Molecular Weight
µl	Microlitre
ng	Nanogram
nm	Nanometre
OD	Optical Density
rpm	revolution per minute
s	Second
U	Unit

CHAPTER 1

1. INTRODUCTION

1.1 Significance of the Androgen Receptor

In 1949 Arnold A. Berthold was the first to state that it was the testes that was responsible for the synthesis of a substance which acts through the bloodstream. The term 'Testosterone' was coined 86 years later by 'Ernst Laqueur' and his co-workers, and it was shown to be the key factor in controlling the development and maintenance of male sexual characteristics (Freeman *et al*., 2001; Nieschlag, 2010).

The action of testosterone occurs via a single intracellular receptor protein defined as the 'Androgen receptor' (AR). The AR fulfills prominent roles in males as well as females. In males the AR is primarily involved in normal virilization of the urogenital tract during embryogenesis, masculinization of the male phenotype at puberty, and maintenance of the male sex organ function during adulthood (Migeon *et al*., 1984; Cunha, 1994; Sheckter *et al*., 1989; Yong *et al*., 2003).

In females, it helps to maintain the growth and development of the mammary gland. The biological roles of the androgen hormone and the involvement of the AR accomplish a vital role in normal maintenance of both the sexes (Zegarra-Moro *et al*., 2002).

1.1.1 Androgen receptor-nuclear receptor/transcription factor

The AR is grouped under the family, class I of the nuclear steroid receptor, also termed as NR3C4 (Nuclear receptor subfamily 3, group C, member 4 (Brinkmann *et al*., 1992; Brown *et al*., 1989). It is highly expressed in genital and some non-genital tissues which are typically androgen dependent, such as muscles, prostate, seminal vesicles, epididymis and testis. Though present in a lesser quantity the AR is also found in several other organs e.g. hypothalamus, pituitary, kidney, spleen, heart and salivary glands (Gao *et al*., 2005; Quigley *et al*., 1995; Lindzey *et al*., 1994; Keller *et al*., 1996; Gelmann *et al*., 2002).

The AR binds to its ligands, testosterone (T) and 5α-testosterone (DHT), with high affinity and through the transcription of various genes regulates a broad spectrum of physiological processes such as metabolism, organ development, and homeostasis (Bocklandt and Vilain, 2007; Komori *et al*., 2007). They also exert major effects on the reproductive system, viz., the prostate and testes, muscles, liver, skin, nervous system and immune system (Gao *et al*., 2002; Culig *et al*., 2000b; Gronemeyer *et al*., 2004; Das *et al*., 2000). To regulate efficient transcriptional control, co-activators and co-repressors also participate to develop a transcriptional network which is further regulated by a series of upstream signaling systems (O'Malley and Kumar, 2009; Beck *et al*., 2009).

As the AR transcriptional complex influences target gene expression any dysregulation in the dynamic intracellular AR complex, or resistance of AR binding to its ligand, alters the modulation of AR functions. Mutations in the

receptor or an alteration in the circulating level of androgens may disturb normal homeostasis causing multiple diseases and inherited traits, consisting of endocrine and neurological disorders, steroid associated cancers, hyper androgenic disorders, prostatic hyperplasia (BPH), androgen insensitivity syndrome (AIS) and polycystic ovarian syndrome (PCOS) (Xu *et al*., 2009; Han *et al*., 2009; O'Malley *et al*., 2008; Brinkmann *et al*., 2001; Araujo *et al*., 2007; Brooke *et al*., 2008).

1.2 Androgen regulation and production

Androgens are the primary sex steroid hormones of the males, which are also known as 'C19-steroid hormones'. They are essential for the development of the human, as it fulfills vital roles for both sexes (Gupta *et al*., 1977; Debes and Tindall, 2002; Heinlein and Chang, 2002). The secretion and regulation of androgens are controlled by the hypothalamus and anterior pituitary (Andrew T Kicman, 2010). See figure 1. As the result of secretion of the gonadotropin-releasing hormone (GnRH) in the hypothalamus the anterior pituitary synthesizes two gonadotropins, luteinizing hormone (LH) and follicle-stimulating hormone (FSH) (Krsmanovic LZ *et al*., 2009).

The secretion of gonadotropin is under control of a negative feedback mechanism operating between the gonads, pituitary and hypothalamus. In the Leydig cells of the testes, LH stimulates steroidogenesis, which results in the production and release of testosterone, whilst FSH regulates spermatogenesis (Nieschlag, 1997). Testosterone acts on the hypothalamus by reducing the hypothalamic pulse and

inhibits the secretion of LH (Sheckter, 1989), whereas the major feedback control and secretion of FSH is from the Sertoli cells through the hormone inhibin (O'Connor, 2004).

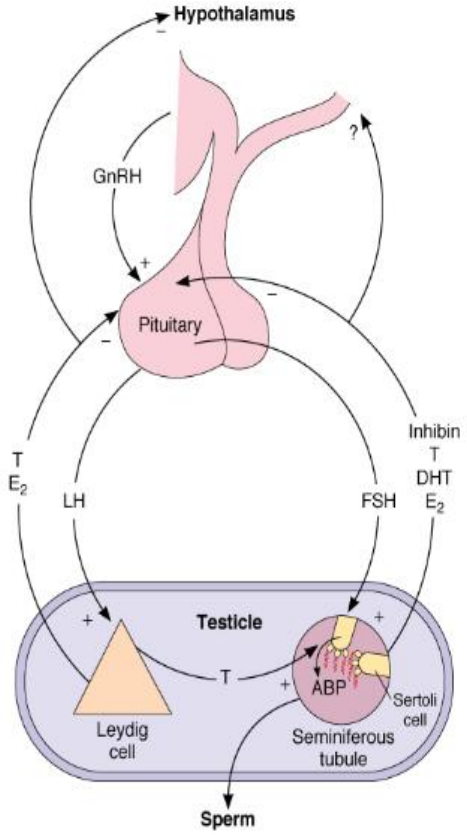


Figure 1: Hypothalamic-pitutary-testicular axis showing the principal negative feedback mechanism. The pitutary release of LH and FSH from the hypothalamus is mediated by GnRH, and binds to definite membrane receptors on Leydig cells and Sertoli cells correspondingly, testicular secretion (Leydig cells secrets the testosterone and estradiol and Sertoli cells secrets the hormone inhibin), act on the anterior pitutary via negative feedback mechanism where they excert control of gonadotropin release. (Adapted with permission from Gardner and Shoback, 2011)

1.2.1 Physiological role of androgens in males and females

The development of the male can only take place if androgens are present to act on target tissues and complete sexual separation (Hughes and Deeb, 2006). Testosterone and 5α-dihydrotestosterone (DHT) are the two most fundamental physiological androgens which act as ligands for the AR (Avila *et al*., 1998). These are synthesized by the testes and secreted into the bloodstream for circulation (Andrew T Kicman, 2010). See figure 2.

In the middle of first trimester, at around 7-8 weeks of gestation, the synthesis and secretion of testosterone, and its conversion to DHT occurs. The initial production of testosterone via the Leydig cells is mediated by placental human chorionic gonadotropin (hCG). Sexual separation and development prolong through weeks 14-16 of gestation (White and Speiser, 2000) and by the 16th week, the placental hCG levels decrease and fetal LH secretion begins to direct circulating androgen levels (Hughes and Deeb, 2006; Speroff, 2005) which helps to maintain normal virilization of the urogenital tract during embryogenesis (Bruchovsky and Wilson, 1968).

Testosterone is involved in the survival, proliferation and differentiation of the Wolffian duct into seminal vesicles, epididymis and ductus deferens, while the development of the male external genitalia, together with the penis and scrotum is due to the involvement of DHT. At puberty, androgen helps to initiate spermatogenesis and the development of accessory sex organs, including the

prostate and it is also responsible for the deepening of voice, phallus enlargement, and male pattern hair growth (Hughes and Deeb, 2006; Speroff, 2005; Wiener *et al*., 1997; Migeon *et al*., 1984).

In contrast, females produce androgens in their ovaries and adrenal glands and in smaller quantities in peripheral tissues. According to earlier studies, a minor role for the androgen/AR system was indicated in female physiology (Freeman *et al*., 2001; Nieschlag, 2010).

In 2006, the finding of dysfunction of the ovaries in female knockout mice missing the AR proved the vital role of androgen action in female reproduction (Shiina *et al*., 2006). Androgens, such as dehydroepiandrosterone (DHEA) and dehydroepiandrosterone-sulfate (DHEA) -S, secreted by the adrenal glands are converted into androstenedione; and also to testosterone by peripheral tissues and are the main androgens in normal postmenopausal women (Noman, 1987; Labrie *et al*., 2003). DHEA, DHEA -S and androstenedione, considered as pro-hormones; act as precursors for androgen and estrogen synthesis as they are not able to bind to and activate the AR.

During fetal reproductive development Mullerian inhibiting substance (MIS) play a significant role. It is secreted in the Sertoli cells of the fetal testes which helps to control the degeneration of the Mullerian ducts. In the lack of MIS, the Mullerian system develops into the female reproductive organ and the uterus. During the postnatal period, adrenal and ovarian androgen helps to develop pubic and axillary hairs (Hughes and Deeb, 2006; Speroff, 2005).

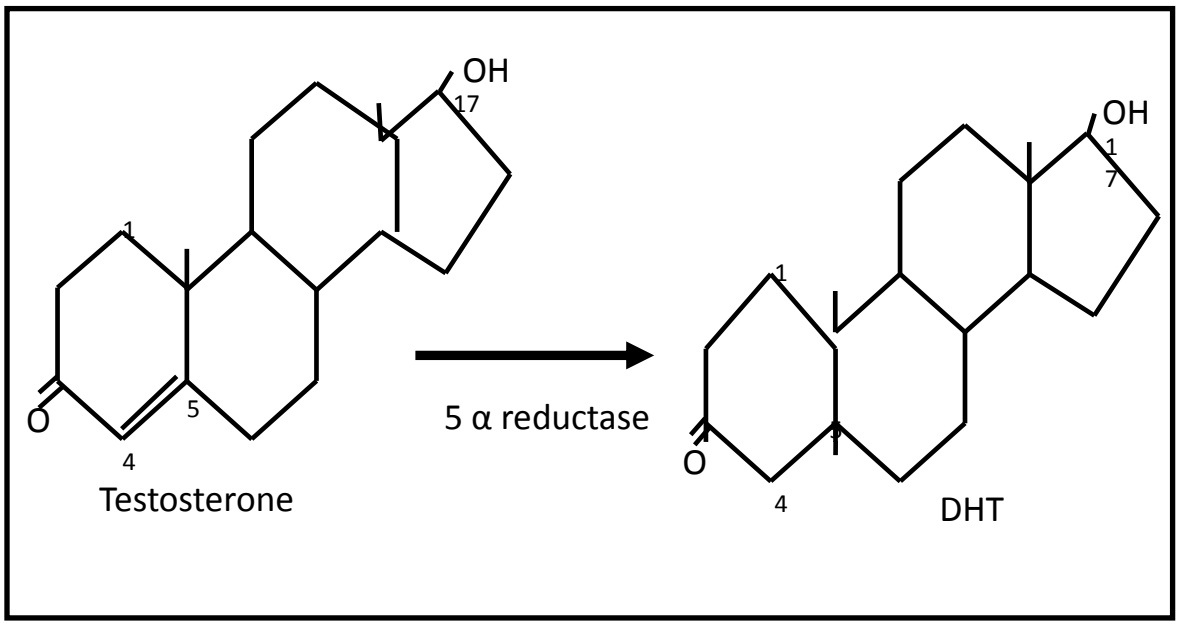


Figure 2: The chemical structure of dihydrotestosterone (DHT) and testosterone (T). The conversion of T to DHT by the enzyme 5α reductase. The double bond existing between position 4 and 5 of T is reduced to a single bond in DHT.

1.3 Androgen insensitivity syndrome (AIS)

Androgen action mediated by the AR determines the sexual phenotype of males and the promotion of spermatogenesis (Noman, 1987). Mutations in the AR gene cause functional defects in the AR protein that prevent normal androgen action, leading to impaired virilization and producing a wide range of clinical presentations, one of which is androgen insensitivity syndrome (AIS) (Grace *et al*., 1970; Hsiang *et al*., 1987). Androgen insensitivity syndrome (AIS) is a common cause for the virilization disorders.

Androgen insensitivity is an X- chromosome linked disorder in which individuals with a 46, XY karyotype develop secondary female sexual characteristics. In this

syndrome, androgens are synthesized but the receptor fails to respond resulting in failure of complete male genital development and is mainly due to mutations in the AR gene. Typically, there is elevated feminization or under masculinization of external genitalia during birth and at puberty growth of abnormal secondary sexual characters and infertility (Abdullah *et al*., 1998). Some mutations in the AR are also associated with gynecomastia and male breast cancer.

AIS includes gynecomastia, testicular atrophy, oligospermia, azoospermia (absence of sperm in semen) and elevated serum gonadotropins. Depending on the degree of virilization of the external genitalia the AIS phenotype has been classified into 7 types, clustering into one of the 3 broad subtypes i.e. Complete Androgen Insensitivity Syndrome (CAIS), Partial Androgen Insensitivity Syndrome (PAIS) and Mild Androgen Insensitivity Syndrome (MAIS) (Aiman *et al*., 1979).

1.3.1 Prevalence of Androgen Insensitivity Syndrome

The prevalence of AIS is estimated to be between "1 in 20,000 to 1 in 90,000 male births" (Hughes and Deeb, 2006) and CAIS is expected to occur in "1 out of every 20,400 46, XY births" (Mazen *et al*., 2010). PAIS is estimated to occur in "1 in 130,000 46, XY births" (Mazen *et al*., 2010) whereas the prevalence of MAIS is unknown (Galani *et al*., 2008). A literature search found no study documenting any difference in prevalence based on ethnicity, geography or exposure.

1.3.2 Complete Androgen Insensitivity Syndrome (CAIS)

CAIS was first described in the 19th century. Both Queen Anne and Joan of Arc were seemingly affected by CAIS (Hughes and Deeb, 2006; Bakan, 1985). This disease is also known as "testicular feminization" syndrome and was first described by John Morris in 1953 (Morris, 1953). It is a rare form of male pseudo-hermaphroditism characterized by a 46, XY karyotype (Speroff, 2005; Robby, 2003).

In 1989 the exact localization of the human AR gene was discovered, located on the Xq11.2-12 chromosome (Brown *et al*., 1989). The first indication that the AR itself is defective in androgen insensitivity came from androgen binding studies on genital skin fibroblast cell lines from AIS subjects (Keenan *et al*., 1974). Brown *et al*., in 1989, published the first evidence that mutations in the AR gene cause AIS (Brown *et al*., 1989). The absence of, or defective, androgen binding was observed in the majority of CAIS cases and suggests a severe deficit in androgen receptor function (Lubahn *et al*., 1988b). The final evidence for the molecular nature of AIS was reported on the basis of sequencing results of the intron - exon boundary of the human AR gene in 1989 (Lubahn *et al*., 1989).

1.3.2a Presentation of CAIS

It was in the early description of CAIS that the mechanism of androgen resistance, and not androgen deficiency, was found to be responsible for clinical presentation. The nomenclature of testicular feminization has been changed to androgen insensitivity syndrome (AIS), and was redefined by discovering the normal level of urinary 17-ketosteroid (Morris, 1953).

In complete AIS, the activity of AR is completely absent. This results in the birth of a phenotypically normal female with a male (46, XY) genotype, and high physiological levels of T and DHT can be seen in these patients due to the lack of the negative feedback mechanism (Morris, 1953; Lubahn *et al*., 1989). The female infant presents with an inguinal hernia. Recent statistics reveals a 1.1% incidence rate of CAIS in children with a premenarcheal inguinal hernia while 80-90% of girls with CAIS eventually develop an inguinal hernia (Gans and Rubin, 1962; Sarpel *et al*., 2005; Viner *et al*., 1997). However at puberty, the female adolescent presents with primary amenorrhea. Scant or absent axillary pubic hair is the most common symptom, but breast development is normal. Height is generally normal, and examination of sex organs shows normal external genitalia with a rudimentary blind end vagina (Viner *et al*., 1997).

According to Griffin *et al.*, CAIS individuals have completely female external genitalia, a paucity of axillary and pubic hair and the absence of Wolffian duct derivatives (Griffin *et al* 1980; Chaung, 2001). Whereas, according to Quigley *et al*, CAIS has complete female external genitalia without pubic hair, but remnants of Wolffian duct derivatives may be found (Quigley *et al*., 1992).

1.3.3 Partial Androgen Insensitivity Syndrome (PAIS)

PAIS is only one etiology of DSD (Disorder of sex development), also known as “Reifenstein” syndrome or "Infertile male syndrome" (Quigley *et al*., 1995). It is a heterogeneous form of CAIS presenting with varying degrees of female

virilization or male feminization due to differing degrees of AR activity. (Quigley *et al*., 1995; Giwercman *et al*., 2002).

In PAIS, the response of cells towards the androgens is incomplete (partial) which results in the impairment of the masculinization of male genitalia in the developing foetus, and it also affects male adolescent, secondary sexual characters during puberty. In females sexual development is normal (Quigley *et al*., 1995; Hughes and Deeb, 2006; Galani *et al*., 2008; Boas *et al*., 2006). This explains the clinical significance of genetic males with a 46, XY karyotype who show insensitivity to androgens (Hughes and Deeb, 2006).

1.3.3a Presentation of PAIS

In PAIS individuals, the symptom of feminization is less than that compared with CAIS patients. PAIS individuals present with a micropenis, perineal hypospadias, and cryptorchidism. Depending on the biochemical nature of the AR, the Wolffian duct derivatives are incomplete or completely developed. Hormonal analysis reveals increased level of LH, T and estradiol at puberty. Affected individuals may develop gynecomastia, small size of the phallus, reduced number of sperm or azoospermia in the testes can be seen. Individuals at puberty may develop *in situ* carcinoma (Melo *et al*., 2003). The PAIS presentation also shows the symptom of "Anosmia", in which the sense of smell is lost (Vincent *et al*., 1991). Assignment of sex is extremely complicated in these patients, as they show ambiguous genitalia during birth. The genitalia may resemble a male; but may

include abnormalities such as cryptorchidism, a micropenis and hypospadias. Alternatively, the appearance may be essentially feminine (Quigley *et al*., 1995).

1.3.4 Mild Androgen Insensitivity Syndrome (MAIS)

MAIS was described in 1979, by Aiman *et al.*, in 3 unrelated men with long histories of infertility and a normal male phenotype with idiopathic oligospermia (Aiman *et al*., 1979). Their group later confirmed it as MAIS based on qualitative and quantitative abnormalities in the cytoplasmic androgen receptor (Aiman *et al*., 1979; Aiman and Griffin, 1982), and it has also been described that the condition for MAIS was the result of male infertility evaluations (Yong *et al*., 2003; Gottlieb *et al*., 2005).

1.3.4a Presentation of MAIS

MAIS currently represents an under explored area. MAIS has been identified in small subset of men with male infertility characterized by oligospermia or azoospermia, normal serum testosterone levels and elevated LH concentrations. In MAIS males genetic studies have demonstrated defects and mutations of the AR gene. These men have normal male external genitalia and male gender orientation. An unambiguously male phenotype with some degree of undervirilization with or without some degree of infertility, is characteristic of MAIS. The incidence of MAIS is unknown but may account for a proportion of male infertility resulting from azoospermia or oligospermia. These individuals may present with mild hypospadias, a normally formed but small phallus, or a male phenotype accompanied by sperm deficiency. Some individuals may have

gynecomastia resulting from increased testosterone levels and its conversion to estrogen. Some subjects are undervirilized and have gynecomastia, but are fertile (Grino *et al.,* 1988). Gynecomastia may eventually turn into a cancerous tumour. Breast cancer in men, which is extremely rare, constitutes only 0.2% of all cancerous tumours in men (Boring *et al*., 1994; Bembo and Carlson, 2004; Derkacz *et al*., 2011).

In one subject with MAIS, administration of pharmacologic doses of androgen improved virilization, facial and pubic hair development, and lowered voice pitch (McPhaul *et al*., 1991). In another MAIS subject, administered androgen successfully corrected the infertility (Yong *et al.,* 1994). All the MAIS subjects classically present with high plasma LH levels, despite high plamsa concentrations of T (Boyar et al., 1978). Some MAIS subjects have androgen-binding defects including upregulation of receptor levels, and increased dissociation rates of AR (Grino *et al.,* 1988). In phenotypically normal men (Ferlin *et al*., 2007; Pinsky *et al.,* 1989) idiopathic azoospermia or oligosperima (Aiman and Griffin 1982) is common. Reduced potency of the AR in MAIS subjects has been shown to be caused by mutations in the HBD of the AR.

These males show increased amounts of serum LH, T and estradiol levels and low volumes of seminal fluid during ejaculation, which is a clear indication of androgen insensitivity. However symptoms in MAIS affected males varies according to mutations in the AR gene (Ferlin *et al*., 2007).

In some patients the external genitalia might not be developed properly, which is the major reason for the infertility. The coronal hypospadias and the central raphe of the scrotum are simple (Concato *et al*., 1992). At puberty, MAIS can have 2 phenotypic forms, both presenting with varying degrees of gynecomastia.

There could be two different phenotypic forms of MAIS, in one form of phenotype, infertility and impairment of spermatogenesis can be seen (Migeon *et al*., 1984; Cundy *et al*., 1986; Mazen *et al*., 2010) whereas, in the other form of MAIS complete spermatogenesis occurs, which maintains fertility (Pinsky *et al*., 1984; Gottlieb *et al*., 1999a).

Physical examination of both phenotypic forms reveal high-pitched voices, sparse pubic hairs, varying degrees and development of gynecomastia in adolescent males. Infants show extremely less hypospadias (Hughes and Deeb, 2006).

1.4 Androgen receptor gene variation and AIS

The AR is one of the most mutated of the steroid receptors. According to the AR database (http://androgendb.mcgill.ca), of the three functional domains the LBD is highly mutated with most of the mutations being missense mutations, and about 20% of the mutations are located on the DBD (Gottlieb *et al*., 1999b; Gottlieb *et al*., 2004) and the remaining mutations are located on NTD. AIS is the leading genetic cause of male pseudohermaphroditism. An understanding of AIS genetic transmission and the related outcome are necessary for clinical decision making, especially, because since it has a significant impact on gender identity and the patient's future sex life (Sultan *et al*., 2002).

1.4.1 Somatic mosaicism

Somatic mosaicism occurs when somatic cells of the body are comprised of more than one genotype. In other words, mosaicism refers to the presence of genetically distinct cell populations within an organism (Youssoufian and Pyeritz, 2002**).** Boveri *et al.*, in 1929 first formulated the concept of somatic mutation and associated these as the cause of cancer. Nonetheless, such changes were also linked to ageing and death (Szilard, 1959). One of the main characteristics of mosaicism is that when cells undergo changes during development one group of cells may differ from a neighboring group due to spontaneous DNA mutations, spontaneous reversion of an existing DNA mutation, epigenetic changes and chromosomal abnormalities. The phenotypes associated with mosaicism depend on the extent of mosaic cell gene population. Equally interesting is the observation that even in absence of clear phenotypic effects, individuals may accumulate somatic variations. Mosaicism has important consequences with respect to human disease and it also results in variations among all humans at molecular level, even among identical twins. If the event leading to mosaicism occurs during development, it is possible that both somatic and germ line cells will become mosaic. In this case, both somatic and germ line tissue populations would be affected, and an individual could transmit the mosaic genotype to his or her offspring. Conversely if the triggering event occurs later in life, it could affect either a germ line or somatic cell population. If the mosaicism occurs only in the somatic cell population, the phenotypic effect will depend on the extent of the mosaic cell population;

however there would be no risk of passing on the mosaic genotype to offspring. On the other hand if the mosaicism occurs only in germ line cell population, the individual would be unaffected, but his or her offspring could be affected. Somatic mutations in the androgen target tissue have been proposed to contribute to phenotypic variation, however data is lacking regarding somatic mutations and only few cases i.e. 25 cases of phenotypic variations have been reported to date (Gottlieb *et al.,* 2001a; Gottlieb *et al.,* 2001b).

There are a few studies which have highlighted the role of somatic mutations in the androgen receptor (Holterhus *et al.*, 1997; Segawa *et al.,* 2002 and Rajender *et al.,* 2007a). Holterhus *et al*., have reported an exceptional case in an individual presenting with PAIS, in whom the outcome of the SHBG-test and the phenotype did not match, due to somatic mosaicism (Holterhus *et al*., 1997).

There is a global database, the androgen receptor database: (http://androgendb.mcgill.ca), which tracks known mutations and their associated phenotypes. The total number of mutations has risen from 605 since 2004 to 1029 in 2012 (Gottlieb *et al*., 2012). There are approximately 800 known AIS-specific AR mutations causing CAIS, PAIS and MAIS (Gottlieb *et al*., 2012). In the AR gene, 4 different types of mutations have been detected in DNA from individuals with AIS and include:

(a) Amino acid substitution or premature stop codons ensuing from a single point mutation.

(b) Nucleotide insertions or deletions, typically leading to either termination of the premature codon or a frame shift mutation.

(c) Complete or partial gene deletions (>10 nucleotides) and

(d) Splicing of the AR RNA due to mutations in the splice donor site or the acceptor site (Gottlieb *et al*., 1999a).

1.5 Breast Cancer

Although the focus of much research into the normal and irregular functioning of the AR has been conducted in the context of male physiology, the AR also plays a pivotal role in female physiology. The normal development and maintenance of breast tissue is under the control of the balanced presence of androgens and their receptors. When androgen levels rise above normal levels, pre and post-menopausal females are at a greater risk for the development of breast carcinoma (Kaaks *et al*., 2005a; Kaaks *et al*., 2005b; Tiefenbacher and Daxenbichler, 2008).

In 1836, Cooper *et al.,* already suggested an association between ovarian secretions and the development of human breast cancer. The expression of the AR has been shown in 70% to 90% of primary breast tumors (Kuenen-Boumeester *et al*., 1996; Buchanan *et al*., 2005).

Breast cancer is the most common malignancy, and the second leading cause of cancer deaths in women throughout the world (Umar *et al*., 2012; Benson and

Jatoi, 2012). By far, the most significant risk factors for breast cancer are gender, age, race/ethnicity, and family history (Thomson, 2012).

The disease is rare before the age of 20, but the rate of growth of the disease associated with aging is higher, as compared to the different forms of the cancer (Stegeman and Bossuyt, 2012). Several studies have demonstrated an excess of familial clustering of breast cancer, suggesting an inherited genetic predisposition to the disease (Manjili *et al*., 2012). As androgens control gene expression, it is essential for the mammary gland to maintain the AR function to avoid the unwanted effects of androgens (Secreto and Zumoff, 2012).

The involvement of estrogen and the progesterone mediated activities in both the normal growth of breast and in the cancerous breast has been documented (Zava and McGuire, 1977). Importantly, the androgen hormone is the predominant sex steroid hormone in post menopausal women, a stage of life during which many cases of breast cancer are diagnosed. The expression of AR has been found in normal breast epithelial cells in approximately 70-90% of invasive breast carcinomas (Hu *et al*., 2011).

1.5.1 Prevalence of Breast Cancer

The most recent breast cancer statistics were reported by the National Cancer Institute (NCI), USA in 2012. Newly diagnosed breast cancer cases reached 226,870 with expected related deaths of 39,510. This represents 621 new breast cancer cases in females and 108 deaths due to breast cancer occurring everyday worldwide per 100,000 of women population.

The highest incidence rate is in western countries. American and British women accommodate the increased rates of (~92%) breast cancer, along with the rest of North Americans, Australians, Eastern and North Europe, and New Zealand women (Perkin *et al* 2005). The lowest incidence rate of (~22%) breast cancer is observed among the Japanese, Chinese and possibly other Asian populations (Parkin, 2007).

Demographically, the difference in the incidence rate of the disease between these native Asians and African-Asians is approximately 50-fold (Benson and Jatoi, 2012). Recent statistics indicate that every 1 out of 29 women in South Africa are diagnosed with breast cancer (Wadler *et al*., 2011). See figure 3.



Figure 3: Geographical Prevalence and statistics of breast cancer worldwide in 2012. The world map highlights the percentage of breast cancer with different colors: occurring everyday worldwide per 100,000 of women population. The

highest rate is 92%, in pink, 74% in light brown, 57% in yellow, 40% in light green and lowest rate is 22%, in light blue. **Source:** (http://www.faqs.org/sec-filings/120716/Immunovative-Inc_10-K).

1.6 Androgen receptor gene variation and breast cancer

Since 1970 it has been known that many cases of human breast cancer have defective androgen receptors (AR). The first evidence of the involvement of the AR gene in breast cancer came from the observation of a mutation (R607Q) in two brothers with breast cancer and partial androgen insensitivity syndrome (Wooster *et al*., 1992). This was followed by another report of a mutation (R608L) in another case of male breast cancer (Lobaccaro *et al*., 1993).

Several case control studies, in different populations revealed that longer CAG repeats in the exon-1 region increase breast cancer risk (Giguere *et al*., 2001; Kadouri *et al*., 2001; Liede *et al*., 2003). In contrast many other studies have limited the impact of AR CAG repeats on breast cancer (Dunning *et al*., 1999; Haiman *et al*., 2002; Spurdle *et al*., 2005).

The length of this polymorphic tract varies among different population groups. Ethnic differences were proposed as a factor affecting the possible variation of CAG repeats length (Gonzalez-Perez *et al*., 2006). It has been shown that Asian population has the longest CAG repeats, whereas the African population has the shortest, and Caucasians falling in the middle (Nicolas Diaz-Chico *et al*., 2007; Sasaki *et al*., 2003).

Studies conducted in Quebec by Giguere *et al.,* suggest, that the mean CAG repeat alleles are 21 (Giguere *et al*., 2001). However, studies from other population groups show differing results. CAG repeats varied from 14 to 31 in Japanese patients (Sasaki *et al*., 2003); 19 to 27 in Filipinos (Liede *et al*., 2003); and 14 to 30 were seen in the Spanish Tenerife population (Gonzalez *et al*., 2007).

Several studies, examining CAG repeat polymorphisms in relation to breast cancer risk conducted in Caucasian populations have been inconclusive (Gonzalez *et al*., 2007; Elhaji *et al*., 2001; Rebbeck *et al*., 1999; Anghel *et al*., 2006). Some, but not all the studies, including those limited to BRCA1 or BRCA1/2 mutation carriers, have found an increased risk of breast cancer associated with longer CAG repeats, particularly at the younger age (Spurdle *et al*., 2005; Slattery *et al*., 2007; Rebbeck *et al*., 1999; Jakubowska *et al*., 2010).

It is well recognized that cancer is the result of several aberrations in the same or different pathways. Therefore, some studies have analyzed the CAG repeats length in individuals who were predisposed to cancer due to some other aberration; for example, BRCA mutation(s).

Out of the five studies (Rajender *et al*., 2007) which analyzed CAG repeat length variation in individuals who were BRCA mutation carriers only one study (Rebbeck *et al*., 1999) reported an association of longer CAG repeats with disease risk. Whereas, the other four (Kadouri *et al*., 2001; Spurdle *et al*., 2005; Dagan *et al*., 2002) reported no association. In these studies, a positive family history of breast cancer appeared to be more significant. Only two studies (Haiman *et al*.,

2002; Wang *et al*., 2005b), analyzed the CAG repeats length in individuals with a positive family history of breast cancer, and both reported a direct association of longer CAG repeats with disease risk.

Interestingly in other studies, no significant association has been made between the AR-CAG repeat length and breast cancer risk in high risk breast cancer families (Kadouri *et al*., 2001; Menin *et al*., 2001; Lillie *et al*., 2003).

A few *in vitro* studies have detected an association of the longer polymorphic allele with reduced AR transcriptional activity, which could further be attributed to the lower affinity of androgens to AR and lower androgenic effects (Kazemi-Esfarjani *et al*., 1995; Nicolas Diaz-Chico *et al*., 2007). Similar findings were outlined for male breast cancer. Some researchers highlighted a significant increase in breast cancer risk only in individuals with a first degree family history (FH) of breast cancer who bore alleles longer than 22 (Haiman *et al*., 2002; Wang *et al*., 2005b).

In addition, androgens might serve as a modifying factor in breast cancer risk among women who are prone to breast cancer (Haiman *et al*., 2002). AR is also known to interact with many tumor suppressor genes (Heinlein and Chang, 2004). There are prospective epidemiologic reports that higher circulating androgen levels in women are associated with a higher carcinogenic risk of developing breast cancer.

In a review of eight prospective cohort studies, an association of endogenous testosterone levels with breast cancer risk was observed (Lillie *et al*., 2003).

Moinfar *et al*, studied AR expression in formalin-fixed, paraffin embedded archival specimens in 200 cases of breast cancer (Moinfar *et al*., 2003); 60% of 145 cases of invasive carcinoma and 82% of 55 cases of ductal carcinoma, *in situ* were AR positive. The great majority of well differentiated carcinomas were both AR and ER positive. Notably, in poorly differentiated carcinomas, 39% were ER negative, but AR positive. Also, in inadequately differentiated carcinomas, 30.5% were both AR and HER2/neu positive, while 42.5% were both AR and HER2/neu negative.

Furthermore, carcinomas with apocrine differentiation were usually AR positive and ER negative, while most invasive lobular carcinomas were AR positive and ER positive. Doane *et al.*, point out the heterogeneity of AR expression in breast cancer, as well as, the complexity of AR signaling in breast cancer (Doane *et al*., 2006).

Agoff *et al.*, studied AR expression in a series of ER negative breast cancers. This subset is of particular interest because ER negative cancers would not be expected to respond to ER blockade or aromatization inhibition (Agoff *et al*., 2003). ER negative cancers which are AR positive might shed light on the biological significance of AR expression. Overall, AR expression correlated with significantly longer survival rates, in a cohort of women with ER negative tumors.

1.7 REVIEW OF ANDROGEN RECEPTOR LITERATURE

1.7.1 Structure of the androgen receptor gene

Along with other nuclear receptors such as, Progesterone receptor (PR), Glucocorticoid steroid receptor (GR), Mineralocorticoid receptor (MR), Vitamin D receptor (VDR) and Estrogen Receptor (ER), the Androgen Receptor (AR) also falls under the "NR3C4" super family (Brinkmann *et al*., 1992; Brown *et al*., 1989).

The human AR gene is localized to the proximal, long arm of the X-chromosome at position Xq11.2-12 (Lubahn *et al*., 1988b; Kuiper *et al*., 1989). It occupies over 90 Kb of DNA, consisting of 920 amino acids according to the updated version of AR database (http://androgendb.mcgill.ca) (Gottlieb *et al*., 2012), with a molecular mass of 110 kDa (Lubahn *et al*., 1988a).

The AR gene is encoded by 8 exons and has two untranslated regions: a large untranslated region towards the 5' end (5'-UTR, 1.1kb in size) and a very large 3'-untranslated region (3'-UTR, 6.8-kb in size) (Faber *et al*., 1991; Kuiper *et al*., 1989). Figure 4 illustrates the AR gene organization and domain structure of the protein. The AR has four functional domains:

(i) a NH2-transactivation domain (NTD) (Amino acids 1 to 538),

(ii) a DNA binding domain (DBD) (Amino acids 539 to 628),

(iii) a Hinge region and

(iv) a COOH-terminal ligand binding domain (LBD) (Amino acids 629 to 920).

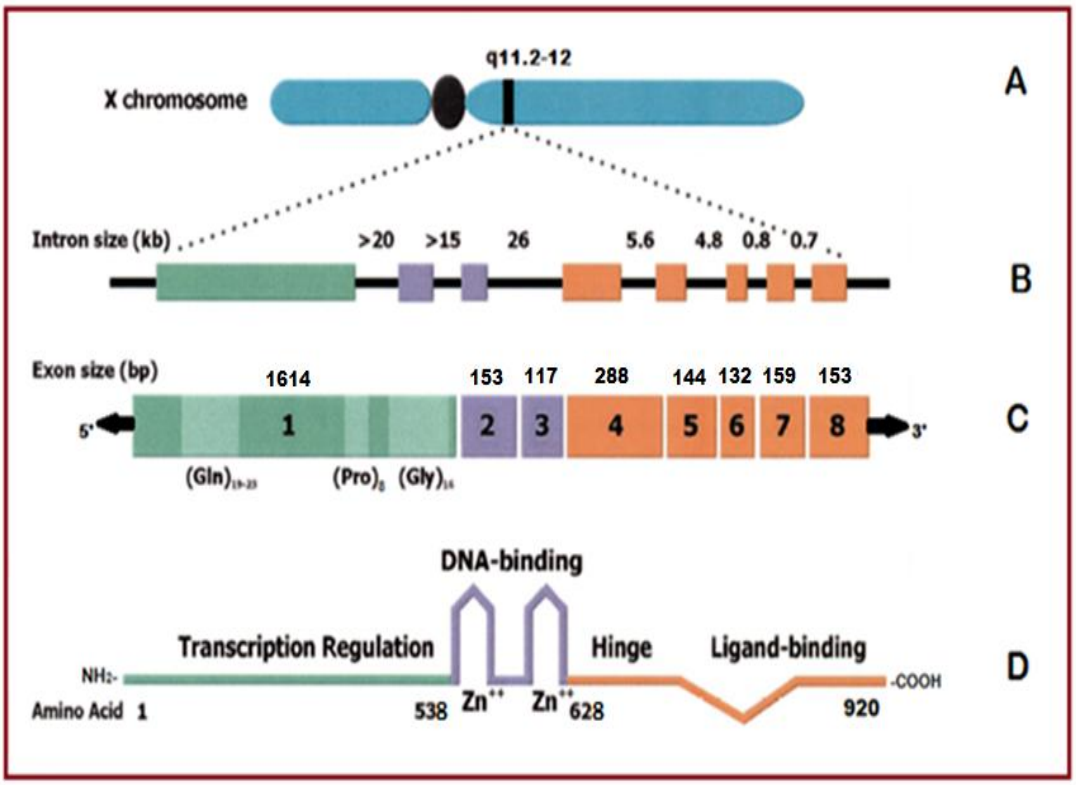


Figure 4: Structure of the AR gene: Location of the AR gene on the chromosome - X is shown in A. The intronic and exonic arrangement of the genome, spanning more than 90Kb is shown in B. The location of three trinucleotide repeat regions, in the exon 1, which codes for transactivation domain, is shown in C. The representation of the protein structure demonstrates how the exon organization translates into the distinct functional regions of the receptor. Reprinted with permission from, American Society of Clinical Oncology, 2002. (Adapted from Edward P. Gelmann, 2002).

The NCBI accession number for the nucleotide and amino acid sequence of the human AR is as referred to in a recent updated version of the androgen receptor database (http://androgendb.mcgill.ca) NM_000044.2 (Gottlieb *et al*., 2012). Figure 5 illustrates the amino acid sequences of each domain.



Figure 5: The amino acid sequence of human androgen receptor gene, NCBI accession number NM_000044.2. The diagram highlights the domains by different colors; NH2-Transactivation domain (Blue), DNA binding domain (Red) and Ligand binding domain (Green).

1.7.2 Androgen Receptor Domains

1.7.2.1 Transactivation Domain

The transactivation domain is the largest domain, representing nearly half of the receptor coding sequence (Quigley *et al*., 1995) consisting of amino acids 1 to 538, encoded by exon 1. The NTD is involved in significant transactivation function and also assists in homeostatic control and AR signaling complexity (Li and Al-Azzawi, 2009).

Within the NTD, Activation function-one (AF-1) has been identified (aa 142-485), which is modular in nature. It serves as an important transactivation function, due to the presence of hydrophobic amino acids (Simental *et al*., 1991; Jenster *et al*., 1995). A few residues (224-258) in the AF-1 domain are highly conserved in exon 1 of the AR, from fish through to primates, and these highly conserved hydrophobic amino acids have been shown to be essential for AR activity and protein-protein interaction (Betney and McEwan, 2003).

AF-1 is composed of sequences which are constitutively active in the absence of a ligand. It is generally responsible for the recruitment of co-activator/co-repressor molecules, to create active pre-initiation complexes; through a TATA box and pol- II, for the regulation of the target gene transcription (Lamont and Tindall, 2010). The NTD of AR AF-1 is known to robustly interact with the LBD, which may further stabilize the AR bound androgen network, and play a role in receptor dimerization (He and Wilson, 2002; Farla *et al*., 2004).

AF-1 consists of two sub-regions, named AF1-a and AF1-b corresponding to residues '154 SSC to SEA 167' and '295 SAG to DYY 359' respectively (Chamberlain *et al*., 1996). AF1-a is predicted to contain a β-turn followed by an acidic amphipathic α-helix. AF1-b is defined by a glutamate (E) and aspartate (D) residue organization, which is similar to an acidic activation domain (Triezenberg, 1995; Chamberlain *et al*., 1996). The definite function of these two sub-regions, in mediating androgen regulated transcription of genes has not been determined (Ikonen *et al*., 1997).

Two transcription activation units (TAUs) have been recognized in the N-terminal domain; TAU1 at amino acid residues 100-370 and TAU5 at the amino acid residues 360-528. They are accountable for receptor dependent trans-activation (Jenster *et al*., 1995; Callewaert *et al*., 2006).

The first 20 amino acids of the NTD contains a "FxxLF" motif (where x is any amino acid) which is beneficial for interacting with the LBD and results in an amino/carboxyl terminal interaction (N/C interaction). It has been implicated that this kind of interaction could be crucial for the transcriptional activation of some AR target genes (He and Wilson, 2002; Steketee K *et al*., 2002).

Inside the exon 1 transactivation domain, three microsatellite regions are described; each made up of the amino acid repeats glutamine (Q), glycine (G) and proline (P) which characterize the AR from other receptors of the NR3C4 super family (Mooradian *et al*., 1987; Janne and Shan, 1991; Janne *et al*., 1993). Of the three microsatellite regions, two are highly polymorphic in length; these are the

upstream glutamine, a classic repeat (referred to as the Poly-Q region) and the downstream glycine complex repeat (referred to as the Poly-G region) (Hughes and Deeb, 2006). The reason for the high polymorphic nature is due to the 'slippage' of DNA polymerase on the multiple glutamine in the template DNA, which results in inconsistencies "in the final set of glutamine repeats copied during DNA replication" (Gelmann, 2002). While the proline repeats are a non-polymorphic stretch of 8 residues from 374-381, the glutamine repeats begin at amino acid residue 58 and the glycine repeats begin at residue 451 (See figure 6). The glutamine repeat sequence, ranges from approximately 10 to 30 repeats, with an average length of 22 repeats (Lundin *et al.,* 2006). Ethnic differences in the glutamine repeat range are well known (Kittles *et al.,* 2001) as discussed previously in section 1.6.

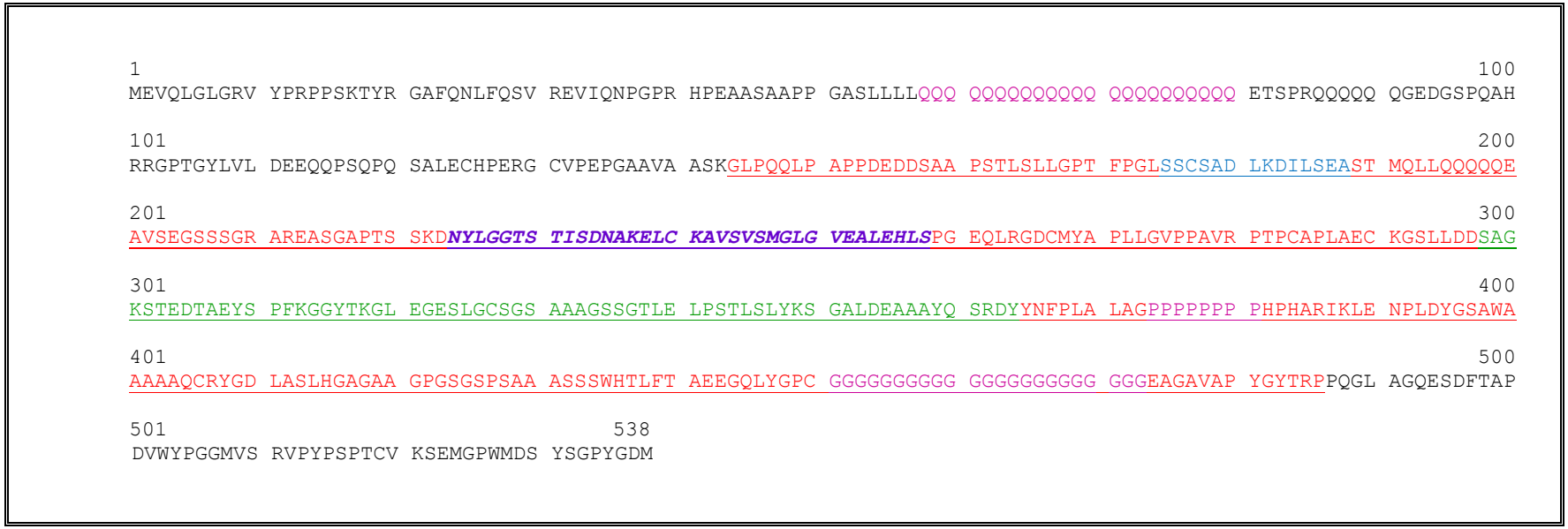


Figure 6: Androgen receptor NH2-transactivating domain amino acid residue 1 to 538, NCBI accession number NM_000044.2. The activating function-1 (AF-1) region (residues 144 to 486) are underlined; AF1-a sub region is shown in blue and AF2-b is shown in green; poly Q, poly P and poly G repeat sequences are shown in pink. The most conserved region is shown in bold italic purple.

1.7.2.2 DNA Binding Domain

The DNA binding domain encoded by exon 2 and 3 is highly conserved. It is rich in amino acid cysteine, which is involved in the creation of 2 zinc clusters; arranged as a pair of loop structures folded to create a single structural unit, made up of two zinc binding motifs (Gelmann, 2002; McEwan, 2004; Verrijdt *et al*., 2003; Hughes and Deeb, 2006; Alvarez-Nava *et al*., 1997; Simental *et al*., 1991; Wong *et al*., 1993; Freedman, 1992) followed by a C-terminal extension.

The first zinc finger contains the so-called proximal (P)-Box, which interacts with the DNA major groove of the Androgen Response Elements (ARE) and determines the specificity of DNA recognition. The different types and categorization of androgen response elements are given in the table 1, (Claessens, 2001). The second zinc finger contains a distal (D)-box, consisting of 5 amino acids (ASRND) and is involved in AR homo-dimerization and DNA/receptor complex stabilization (Gelmann, 2002).

Table 1: Different types and classification of androgen response elements. Reprinted with permission from, the Journal of Steroid Biochemistry and Molecular Biology, 2001. (Adapted from Claessens 2001).

AREs	Sequence	Reference
High affinity, non-specific		
GRE177	GTTACA AAC TGTTCT	Beato *et. al.* 1989
C3(1) ARE	AGTACT TGA TGTTCT	Claessens *et. al.* 1989
GRE2 TAT	TGTACA GGA TGTTCT	Beato *et. al.* 1989
PSA ARE-1	AGCACT TGC TGTTCT	Reigman *et. al.* 1991
SLP-HRE-3	GAAACA GCC TGTTCT	Loreni *et. al.* 1988
High affinity, AR specific		
PB-ARE-2	GGTTCG TGG AGTACT	Rennie *et. al.* 1993
SLP-HRE-2	TGGTCA GCC AGTTCT	Loreni *et. al.* 1988
SC ARE 1, 2	GGCTCT TTC AGTTCT	Verrjdt *et.al.* 1999
Low affinity, non-specific		
PB-ARE-1	ATAGCA TCT TGTTCT	Rennie *et. al.* 1993
MVDP pARE	TGAAGT TCC TGTTCT	Darne *et. al.* 1997
GPX5	ATCCTA TGT TGTTCT	Lareyre *et al.* 1997
CRP2	AGAAGA AAA TGTACA	Devos *et. al.* 1997
Low affinity AR-specific		
SC ARE	AGCAGG CTG TGTCCC	Haelans *et al.* 1999

In general, these zinc fingers are responsible for specific binding of transcription factors to cognate DNA elements that are organized as inverted repeats of hexameric binding sites (5'-TGTTCT-3'NNN5'-TGTTCT-3'), separated by three nucleotide spacers (See figure 7). Nucleotide sequence-specific DNA binding of receptors to their respective response elements leads to DNA binding suggesting that interactions between different *cis*-acting factors bound to DNA can occur, by drawing them into proximity within the transcriptional activation system. In addition, the DNA binding of AR may be influenced by its interaction with other nuclear proteins (Gelmann, 2002; McEwan, 2004; Verrijdt *et al*., 2003; Hughes and Deeb, 2006; Alvarez-Nava *et al*., 1997).

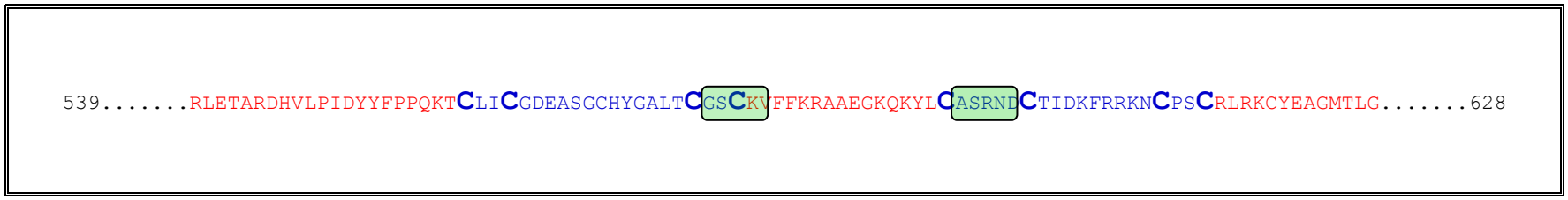


Figure 7: Amino acid sequence (539 to 628) of the AR DNA binding domain. NCBI accession number NM_000044.2. Zinc finger sub domains are shown in blue; Cysteine residue coordinating with zinc atoms are in large bold, P and D boxes are highlighted in the rectangle boxes.

1.7.2.3 Hinge region

A hinge region is located between the 3' end of exon 3 and 5' region of exon 4, consisting of flexible amino acid sequences (Janne and Shan, 1991). It contains the main nuclear localization signal (NLS) of the receptor, which is encoded by the 5’ portion of the exon 4, and spans approximately 51 amino acids (625-676) (MacLean *et al*., 1997; Jenster *et al*., 1991).

The NLS consists of two essential amino acid clusters found in the DBD and hinge, from amino acid 617-633 (**RK**CYEAGMTLGA**RKLKK**) (Zhou *et al*., 1994). The binding of androgen to the AR causes a receptor conformational change that regulates the NLS activity and translocation of the androgen in the AR complex into the cell nucleus (MacLean *et al*., 1997; Zilliacus *et al*., 1992).

The 280 kDa of cytoskeletal protein Filamin-A (FlnA) is derived from an actin binding domain (ABD), and 24 sequential stick like domain repeats also involve in nuclear translocation (Ozanne *et al.,* 2000; Wang *et al.,* 2005). The interaction of the cytoskeletal protein FlnA with the DBD, LBD and hinge, promotes the nuclear translocation of AR.

The hinge region is also required for intra-protein interaction between AF-1 and AF-2 domains. Deletion of NLS in the hinge region causes a partial reduction of AR's translocation into the cell nucleus, suggesting that a second NLS is also involved in translocation, which is available in the LBD (Zhou *et al*., 1994).

1.7.2.4 Ligand Binding Domain

The carboxyl terminal region of the AR contains the ligand binding domain (aa 677-920) encoded by 3' region of exon 4 and exons 5-8 (http://androgendb.mcgill.ca/). The structure of the androgen receptor-LBD (AR-LBD) is similar to the LBD of other steroid hormones such as PR, GR, ER, MR and others (Matias *et al*., 2000; Sack *et al*., 2001).

The high affinity interaction between the AR and its steroidal and non-steroidal ligands occurs through LBD, and is also involved in the recruitment of co-regulators during transcriptional activation (Heemers and Tindall, 2007). The LBD is also involved in the interaction of AR with chaperons. Please refer to figure 8.

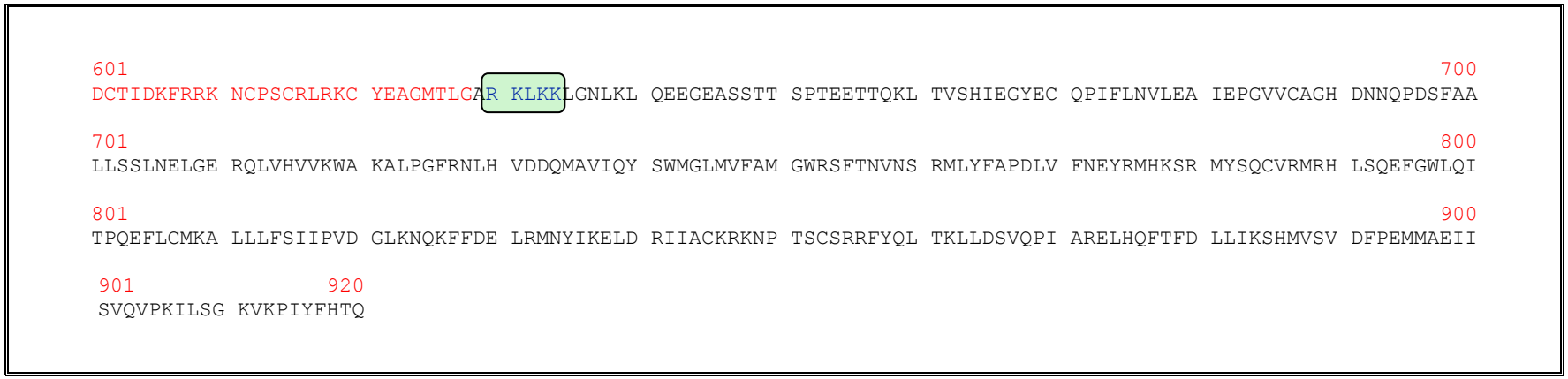


Figure 8: Amino acid residues 629 to 920 of the human androgen receptor (AR), Hinge region and Ligand binding domain. NCBI accession number NM_000044.2. The residues involved in electrostatic interaction with co-regulators are highlighted in the rectangular box.

1.7.2.4a 3D- Structure of the LBD

The 3D-crystallographic structure of the LBD was described early in the previous decade (Sack *et al*., 2001). Similar to other steroid hormone receptors the 3D-structure of AR- LBD established 12α helices and 4 short β–strands arranged in 2-β sheets, folded into a three-layer sandwich-like structure (Wadler *et al*., 2011; Matias *et al*., 2000; John *et al*., 2000) Figure 9 a.

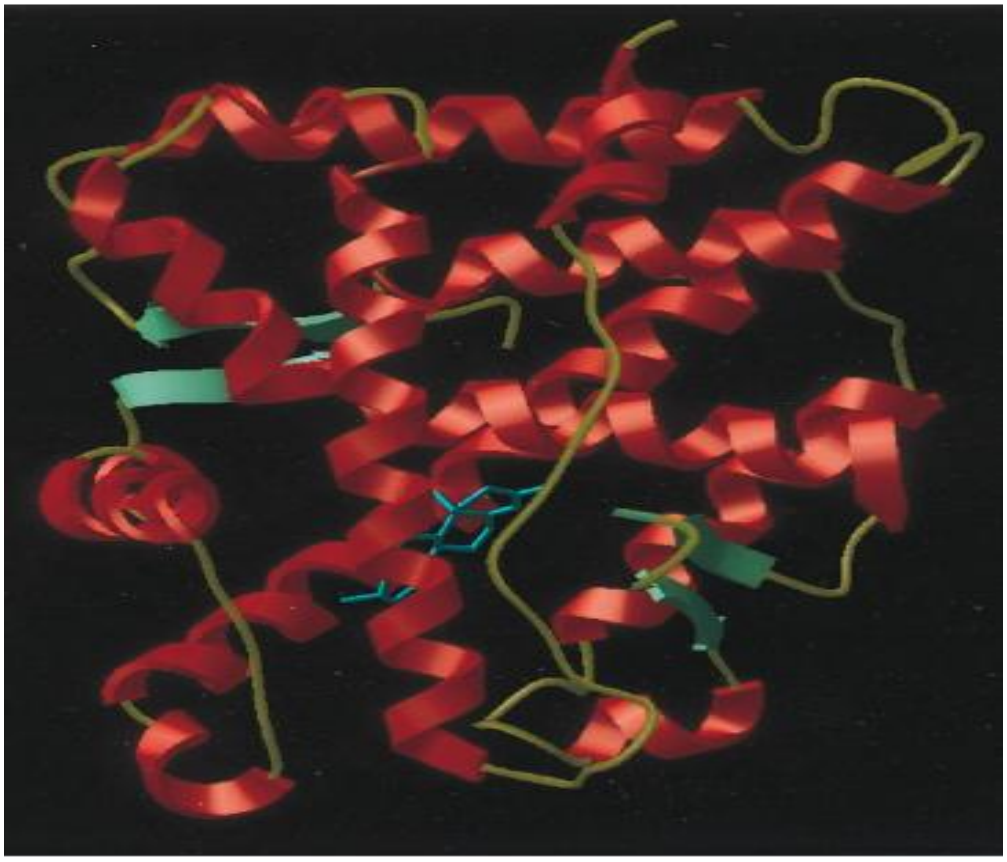


Figure 9a: The 3D-crystallographic structure of AR LBD. Reprinted with permission from Proceedings of the National Academy of Sciences, 2001, USA. (Adapted from John.S.Sack *et al.*, 2001).

The 3D-crystallographic structure revealed the ligand binding pocket allowing interaction with surface binding proteins, and with co-activators and co-repressors. The ligand binding pocket of the AR-LBD consists of about 20 amino acid residues which communicate with the bound ligand (Matias *et al*., 2000) and is flexible in nature capable of recruiting ligands with different structures.

Following hormone binding helix12 folds back into the core of the LBD allowing closure of the ligand binding pocket (Giammanco *et al*., 2005; Singh *et al*., 2006; Lubahn *et al*., 1988b; Brinkmann *et al*., 1989). Thus, conformational change occurs in the LBD resulting in the stabilization of helix 12 and the creation of functional AF2 on the surface. The structural organization of AR-LBD is shown in the figure 9b.

The role of AF2 in the LBD is mainly dependent on availability of nuclear receptor co-activators. A ligand dependent functional interaction of AF2 region in the LBD with the NTD is supported by *in vivo* experiments (Langley E *et al.*, 1995; Berrevoets CA *et al*., 1998).

Deletions in the LBD entirely disrupt hormone binding (Jenster G *et al*., 1991), while the deletions in the NTD and DBD do not affect the hormone binding. Five codons within exons 6 and 7 of the human AR encoding the LBD act as “hot spots” for mutations based upon their relative frequency of mutation and the alternative amino acid substitutions that occur within each codon (Murono *et al*., 1995).

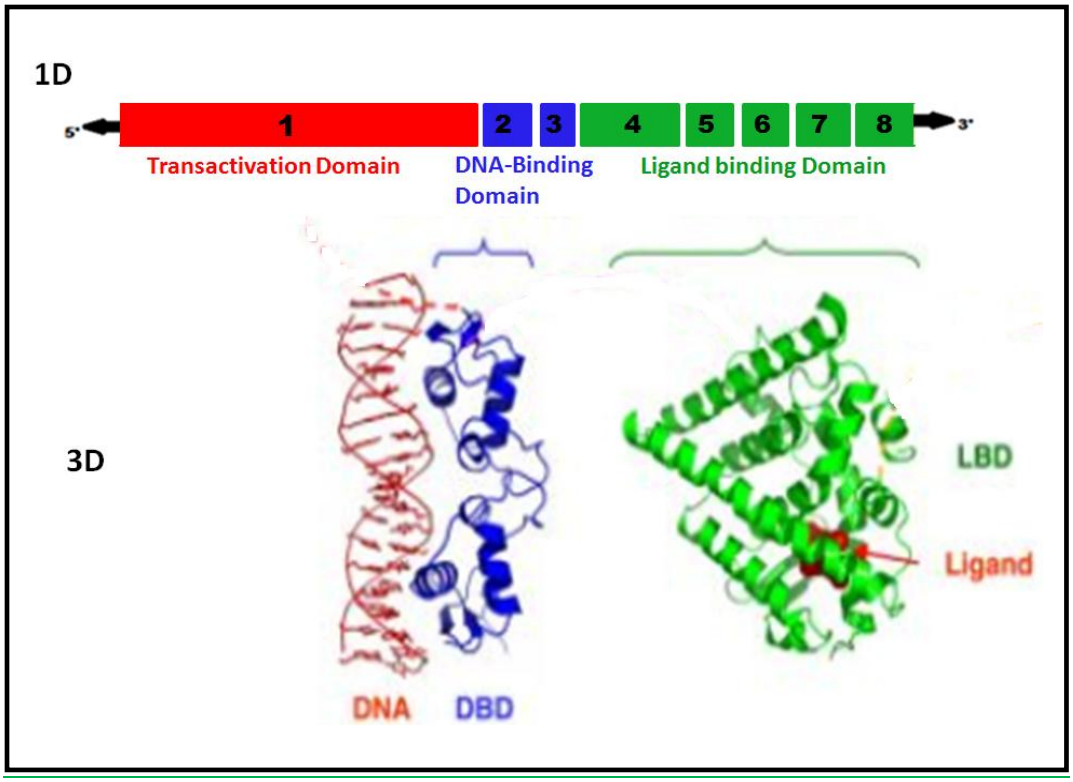


Figure 9b: Structural organization of androgen receptor**. Top:** Schematic 1D amino acid sequence of androgen receptor. **Bottom:** 3D structures of DBD bound to DNA and LBD bound to hormone regions of androgen receptor. This figure is adopted from the website (http://en.wikipedia.org/wiki/Nuclear_receptor).

1.8 Mechanism of AR action

An outline of how AR mediates the biological effects of androgens is known (Mooradian *et al*., 1987; Lamont and Tindall, 2010; Janne and Shan, 1991; Janne *et al*., 1993). The regulation of cellular proliferation and differentiation of androgens occurs through the AR. A schematic diagram showing the mechanism of action of the AR is shown in figure 10. This process requires at least 30-40 minutes altering the expression of genes, whereas it requires hours to create significant levels of newly expressed proteins.

Testosterone enters the target cell via the plasma membrane and is converted into DHT by the help of enzyme 5-α-reductase type II, and interacts with the AR present in the cytoplasm. When the ligand is not present, the AR is maintained in an inactive form by molecular chaperon heat shock proteins (HSP) 90, 70, 56 and p23-a. A co-chaperon stabilizes the binding of HSP- 90 to the receptor together with various co-chaperons including cyclophilin 40 (Cyp40), and FK506-binding protein 52(FKbp52), to prevent the constitutive activation of the receptor and is dispersed throughout the cytoplasm (Culig *et al*., 2000a).

When the ligand binds the C-terminal helix-12 in the LBD shifts position to block the ligand binding pocket. This results in the separation of heat-shock proteins that results in facilitating the AR to network with co-regulators, such as ARA-70 (binds to AR-DBD and AR-LBD), cytoskeleton protein filamin-A and importin-α, which binds to the AR nuclear localization signal (NLS).

This network encourages the nuclear targeting of AR and the formation of homo-dimers. It also facilitates a series of conformational changes in the AR-LBD with the help of helices 3 and 5 and the subsequent release of co-repressor complexes (Bennett *et al*., 2010). The AR-DHT complex then travels to the nucleus where it binds to specific chromosomal DNA sequences (ARE) in the regulatory regions (promoter/enhancers) of the AR-regulated gene. Subsequently, the recruitment of co-regulators and general transcription factors take place through the activated AR, which leads to the development or control of the target genes (Culig *et al*., 2000a; Bennett *et al*., 2010). The nature of AR-bound ligands determines the strength of AR-DNA complexes and eventually transcription is initiated. Under

normal physiological conditions, only half of the total cellular AR molecules are thought to be occupied by the ligand for mediating androgen action. The remaining half of unliganded AR is proposed to regulate other cellular events such as the cell cycle (Ma *et al*., 2001; Yamamoto *et al*., 2000).

Recent chromatin immunoprecipitation (ChIP) studies suggest that the ER and AR agonists recruit co-regulators cyclically in a time-dependent manner to suitable response elements (Shang *et al*., 2002; Wang *et al*., 2005a). Once transcription is initiated, the AR separates from the DNA and moves to the cytoplasm where it undergoes ubiquitination and proteolytic cleavage. In the next round, the transcriptional activation take place with a fresh receptor and the cycle continues in the presence of a fresh ligand (Cardozo *et al*., 1998; Pajonk *et al*., 2005). The unbound AR either shuttles back to the cytoplasm for a recycling process or gets degraded by the proteosomal degradation process.

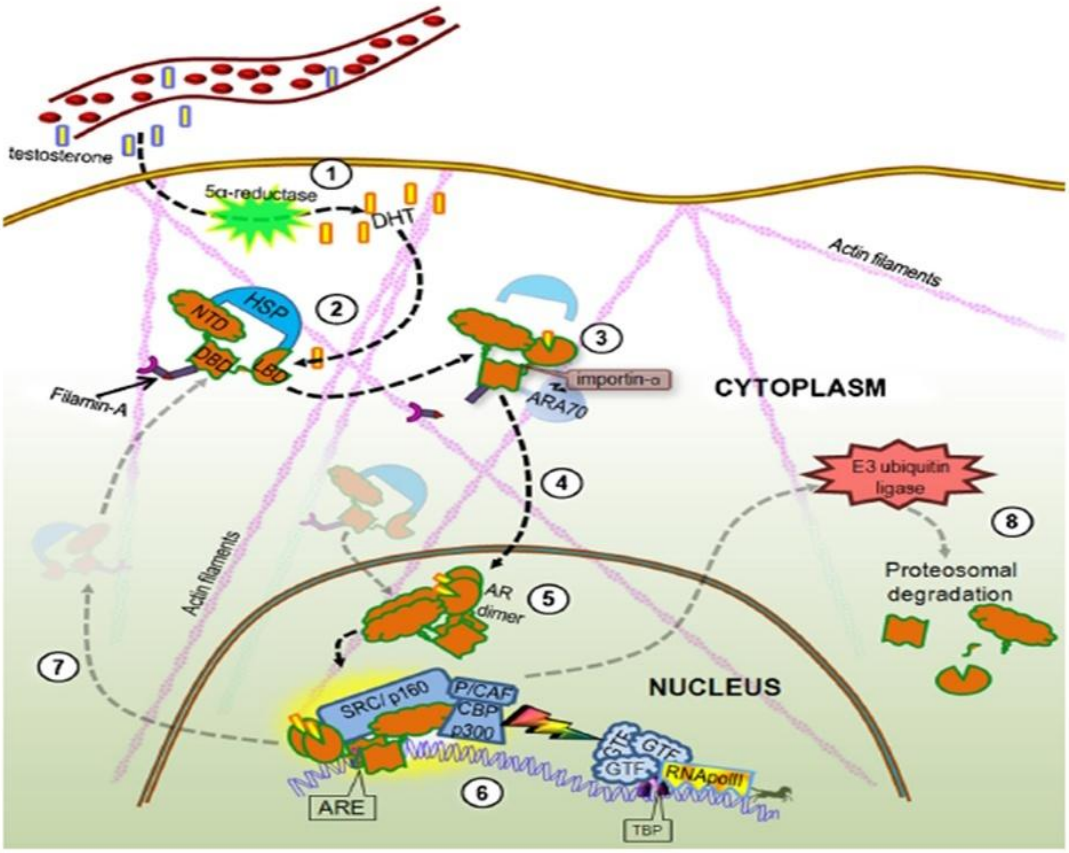


Figure 10: AR cycle. The androgen receptor (AR) (coloured orange) is illustrated as a modular protein with each of its domains represented; NTD (N-terminal transactivation domain), DBD (DNA-binding domain), the hinge and the LBD (ligand-binding domain). **(1)** Conversion of testosterone to DHT via the enzyme 5-reductase in basal epithelial cells within the prostate; **(2)** DHT moves into the cytoplasm of epithelial cells where it is bound by AR; **(3)** Ligand binding to AR induces an AR conformational change, whereby proteins such as heat shock proteins (HSPs) dissociate and others such as importin- and the androgen receptor-associatedprotein-70 (ARA70) are recruited to help stabilise AR and **(4)** promote AR nuclear translocation; **(5)** AR dimerises in the nucleus; **(6)** Numerous other co-activators bind to the AR in the nucleus, then the AR-DBD facilitates nucleic acid binding at androgen response elements (ARE) which promote the recruitment of co-activators with histone acetyltransferase (HAT) activity (CBP/p300, P/CAF), via SRC/p160 co-activator family members, resulting in chromatin remodelling. This allows the binding of the TATA binding protein (TBP), followed by general transcription factors (GTF) and RNApolII to begin transcription; **(7)** Non-ligand-bound AR is shuttled back to the cytoplasm and recycled in preparation for further ligand binding**. (8)** Alternatively, non-ligand-bound AR can be targeted for proteosomal degradation following ubiquitination by E3 ubiquitin ligase. Reprinted with permission from, the International Journal of Biochemistry and Cell Biology, 2010. (Adapted from Bennett *et al.*, 2009).

1.9 AIMS

The overall aim of this work is to examine the relationship between AR-CAG repeat length polymorphisms, AR-GGN repeat length polymorphisms of exon 1 of AR gene and point mutations of the AR gene in affected males with MAIS and gynecomastia and breast cancer affected females of the enrolled family. The specific aims are:

- To examine whether the CAG repeat length of exon 1, have any impact on the MAIS and breast cancer subjects participating in this familial study.
- To examine whether the GGN repeat length of NTD, have any impact on the MAIS and breast cancer subjects of the present study.
- To investigate, if any point mutations of the androgen receptor gene are linked to the MAIS/gynecomastia and breast cancer in the individuals in the enrolled family .
- To identify the somatic mosacism in exon I region using cloning.

CHAPTER 2

2. MATERIAL AND METHODS

2.1 Ethical considerations

This study was approved by the Human Research Ethics Committee (Medical) of the University of the Witwatersrand. The ethics clearance number is M121194, Appendix A.

2.2 A Family Pedigree included in the study, its three components and analysis

A South African family of Indian descent was enrolled in this study. In this family several males were affected with MAIS and gynecomastia, and the females were affected with an aggressive early onset of breast cancer. A pedigree chart was drawn up and is shown in the figure 11.

In this family pedigree, there were a total of 34 members, of which only 16 were involved in the present study and the other 12 members refused to participate. Six members of the family were deceased of which the 2 females died as a result of breast cancer and two males with MAIS and gynecomastia. The other deceased male and female died of natural causes.

The index case was a married phenotypic male of second generation (G2) who was unaffected but the father of the index case (G1) was suspected to be affected with MAIS and gynecomastia. A detailed family history was obtained by all the members of the family who were willing to participate in the study. All patients

were being treated by Prof. Girish Modi, Neuroscience department, Charlotte Maxeke hospital, University of Witwatersrand and clinically diagnosed with MAIS and gynecomastia in males and breast cancer in females. There were 7 affected members (6 males with MAIS and one female with breast cancer) in the present genetic study. All the males of the family had a 46, XY karyotype with a male phenotype and all the females in the family had a 46, XX karyotype with a female phenotype.

The special observation of this family was that all the affected males with MAIS and gynecomastia were in the left arm of the pedigree whereas the breast cancer affected females were in the right arm of the family. More precisely, both arms were descendents of the same father of (G1), but one arm (left) was comprised of male offspring's with MAIS and their genetic lineage males (i.e. next generation) were also affected with MAIS, whereas none of the females in the left arm of the pedigree had breast cancer although they are under observation for such a development. In the right arm, the daughter of the suspected father (G1) was affected with breast cancer and her genetic lineage females also had breast cancer, except one, who was under observation for such a development. There was a single male in the right arm and he was also under observation for androgen related disorders. The family was divided into following three groups:

(1) Affected members of same lineage [(n = 7) M: 6, F: 1].

(2) Unaffected members of the same lineage [(n = 4) M: 2, F: 2].

(3) Genetically un-related normal family members (by marriage), who served as controls [(n =5) M: 1, F: 4].

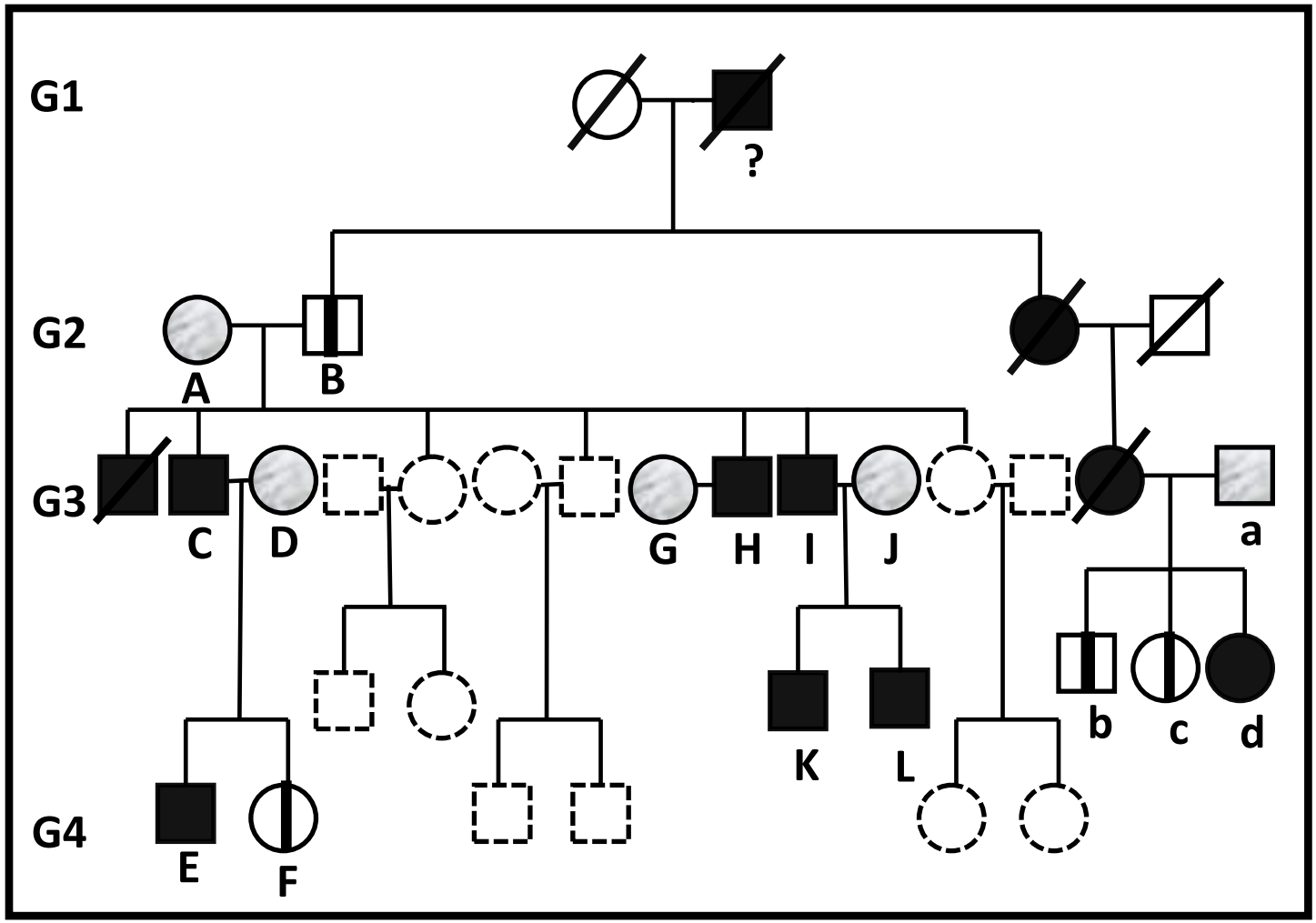


Figure 11: ***Pedigree of family kindred with MAIS and Breast Cancer:*** Circles and squares indicate Females and Males, respectively. G1, G2, G3 and G4 indicate the generations. Solid black circles and squares indicate the affected subjects. A solid vertical black line inside the circles and squares indicate the unaffected related subjects. Light-shaded circles and squares indicate the controls. Circles and squares with a cross represent the deceased subjects and the subjects who refused to participate in the study are indicated by dotted circles and squares.

2.3 Extraction of genomic DNA from patient whole blood samples

Genomic DNA was extracted from whole blood using a standard salting-out method as described by Miller *et. al*., 1988. This method involves salting out of the cellular proteins by dehydration and precipitation of DNA with a saturated NaCl solution (Miller *et.al.,* 1988).

Ten ml of the fresh whole blood was collected in EDTA vacutainer tubes (BD Biosciences, South Africa), mixed with 45 ml cold lysis buffer (Appendix B) and centrifuged at 2300 rpm for 2 mins. The supernatant was decanted and the pellets were resuspended in 20 ml ice cold lysis buffer, incubated on ice for 5 mins, and thereafter centrifuged briefly at 2300 rpm for 2 mins at 4°C in a centrifuge. The pellet was resuspended in 4 ml of nuclear lysis buffer and 500 µl proteinase K which contained 10% SDS (Appendix B) was added. The samples was mixed well and incubated at 37 °C in a water bath overnight.

The next day, 1 ml of a saturated NaCl solution (Appendix B) was added to each tube, mixed vigorously and incubated on ice for 5 mins, followed by centrifugation at 2300 rpm for 30 mins in a centrifuge. The supernatant was transferred into a clean 50 ml tube and the DNA precipitated with 20 ml absolute ethanol.

The samples were incubated at -20°C overnight and centrifuged at 8000 rpm for 30 mins at 4^0C. To remove the salt used to precipitate the DNA, the pellets were washed with ice cold 70% ethanol, centrifuged at 8000 rpm for 20 mins in a

centrifuge and, the DNA pellet air dried at room temperature and the DNA finally dissolved in 500 µl TE buffer (Appendix B).

The purified genomic DNA was quantified using a Nanodrop Spectrophotometer ND-1000 (NanoDrop Technologies Inc., Wilmington, Del.).

2.4 PCR amplification of the coding region of the AR gene

2.4.1 The design of primers to amplify exons 1 to 8 of AR gene

Exon 1 is the largest exon of the AR gene. At 1614 bp it constitutes over 58 % of the complete AR (2760 bp) gene. Exon 1 is also a guanine and cytosine (GC) rich region, since it contains stretches of CAG and GGC triplet repeats. These two factors presented a problem when it came to amplifying exon 1 as we were unable to amplify one PCR product that could be sequenced with accuracy. The strategy adopted was to therefore to amplify exon 1 in two parts designated Part I and Part II. Since there was no problem with exons 2-8, they were amplified as a single product.

Based on the most recent reference sequence of the human AR gene (Gene Bank ID NM_000044.2, http://androgendb.mcgill.ca/), two sets of primers were designed for exon 1, in such a way to cover the whole 1614 bp sequence. The resultant products overlap by 220 bps as shown in the figure 12. We designed the exon primers for the amplification of each exon using the primer3 software (http://primer3.wi.mit.edu/) and these were synthesised by Inqaba Biotech Ltd (Pretoria, South Africa). The sequence of the primer pairs for all the exons are listed in the table 2.

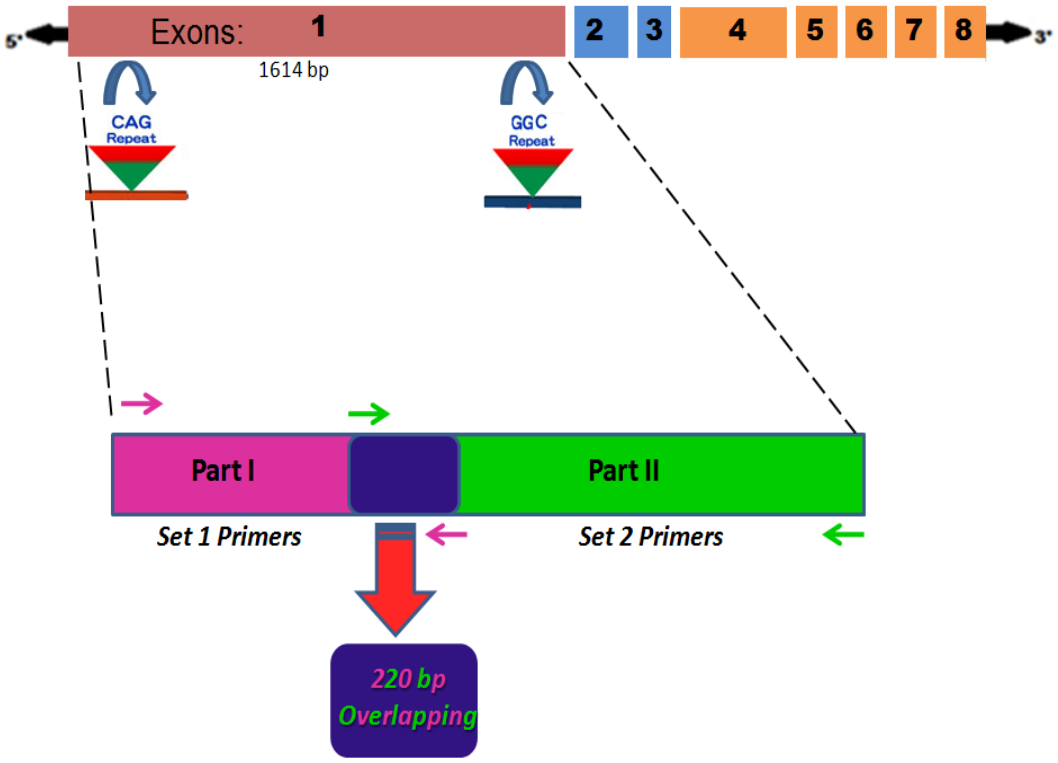


Figure 12: This diagram represents the primer design for exon 1 of the AR gene. Exon 1 was amplified in two parts. Part I is shown in pink and amplified by primer set 1. Part II is shown in green and amplified by primer set 2. The overlapping region of 220 bp is shown in purple.

2.4.2 PCR conditions for amplifying exon1 to exon 8 of the AR

PCR reactions for amplification of all exons were optimized in a total volume of 50 µl, containing 2.5 mM of each deoxynucleoside triphosphate (dNTP), 20 µM of each set of primers, 1.5 mM MgCl2, 10x PCR buffer, and 50 ng/µl genomic DNA. 5U BIOTAQ ™ DNA polymerase was used for all PCR reactions (Bioline Ltd, Luckenwalde, Germany), except part II of exon 1, in which, Hotstar Taq DNA (QIAGEN) was used. Additionally 5 µl of 5x Q-solution (Qiagen) was used

for amplification of exon 1 (part I and part II), according to the manufacturer's protocols. Q solution is a proprietary solution recommended for amplification of templates containing GC rich sequences which is capable of resolving the formation of secondary DNA structures. By influencing the melting behavior of DNA, Q solution is able to diminish polymerase failure as a result of the hairpin structures and loops caused by high GC content.

Reactions were performed in a thermal cycler (MJ Mini, Bio-Rad) with the following reaction conditions: 94 ˚C initial denaturation for 5 min, 94 ˚C denaturation for 1 min, annealing for 1 min, 72 ˚C extension for 2 min, for a total of 35 cycles, followed by 72 ˚C final extension for 10 mins. The annealing temperatures of each PCR reaction are listed in the table 2. All PCRs were performed in duplicate and a negative control (no template DNA) was included in the PCR analysis to rule out DNA contamination.

Table 2: Primers for PCR amplification, annealing temperature and location of each exon on the X chromosome (Gene Bank accession # NM_000044.2)

Exons	Primer Sequence	Location on X chromosome (start-end)	Size of amplicons (bp)	Annealing Temperature
Exon 1:Set 1	Forward: 5'-AGAGGAGGCGACAGAGGGAAA-3'	5061-5081	1605	65 ˚C
	Reverse: 5'-ATGTCTTTAAGGTCAGCGGAGCAG-3'	6643-6666		
Exon 1:Set 2	Forward: 5'-TGGATGAGGAACAGCAACCT-3'	6443-6463	1533	59 ˚C
	Reverse: 5'-GGTTGGGGGAGTTGAAGAAT-3'	7957-7976		
Exon 2	Forward: 5'-TGAAGACCTGAGACTTCACTTGCC-3'	104149-104172	311	58 ˚C
	Reverse: 5'-AATCCTGGGCCCTGAAAGGTTAGT-3'	104437-104460		
Exon 3	Forward: 5'-GACTCTGGAAACTCATTATCAGGTCT-3'	146927-146952	205	60 ˚C
	Reverse: 5'-AGAGGAAGGAGGAGGAAGAGAAAG-3'	147109-147132		
Exon 4	Forward: 5'-TAGAGTCTGTGACCAGGGAG-3'	172248-172268	525	54 ˚C
	Reverse: 5'-ATGAAGACTGCCTGAGTTAATG-3'	172752-172773		
Exon 5	Forward: 5'-ACTGCCTCTGCCTCTTCTTCTC-3'	178422-178443	255	58 ˚C
	Reverse: 5'-TCACTGTCACCCCATCACCATC-3'	178656-178677		
Exon 6	Forward: 5'-ATCAGAGACATTCCCTCTGGGCTT-3'	182728-182752	230	54 ˚C
	Reverse: 5'-CATTCCCTGCACTTCTAGGCACTT-3'	182935-182958		
Exon 7	Forward: 5'-CCCATTCTGTCTTCATCCCACATC-3'	183770-183793	219	60 ˚C
	Reverse: 5'-TTCTCCCTGATAAAGCACCCTCCA-3'	183966-183989		
Exon 8	Forward: 5'-CACCTCCTTGTCAACCCTGTTT-3'	184608-184629	615	54 ˚C
	Reverse: 5'-TCCCCAGATAATCTTAGCTCTC-3'	185202-185223		

2.4.3 Agarose gel electrophoresis

Amplified products of each exon were resolved on agarose gels of the appropriate volume and concentration, prepared as described in Appendix B. Once completely solidified, the agarose gel was immersed fully in an electrophoresis tank containing 1x TBE buffer as given in Appendix B. 5 µl of amplified PCR product were mixed thoroughly by adding 1µl of 6x Blue/Orange loading Dye (Promega, Madison, WI, USA) and were loaded into an agarose gel and electrophoresed at ~90 to ~100 volts/cm until the bromo phenol blue tracking dye front had migrated ~75% of the length of the gel. An image of each gel was recorded using the Gel Doc XR Imaging System (Bio-Rad). A clear band of the predicted size (given in the table 2) of each exon was observed after the gel electrophoresis (see figures 15 to 22).

2.5 Analysis of sequencing data of exons 2-8 of the AR gene

Un-purified PCR products of exons 2-8 were sent to Inqaba Biotechnical Industries Ltd (Pretoria, South Africa) for direct DNA sequencing, where sequencing was carried out using an ABI 3130XL genetic analyzer, using same set of primers as was used for the PCR. Sequences were subjected to a standard nucleotide-nucleotide BLAST search (http://www.ncbi.nlm.nih.gov/Blast.cgi) to ensure that the amplicons were the human androgen receptor gene.

The analysis was divided into 3 categories; affected, un-affected and control. Sequences were analyzed at the nucleotide level as well as amino acid level, nucleotide sequences were converted into amino acid using the ExPASy translator

tool. Bioedit and Clustal Omega software's were used for aligning both nucleotide and amino acid sequences. The most recent human AR gene (Gene Bank ID NM_000044.2) was used for the alignment. 153 nucleotides of exon 2, 117 nucleotides of exon 3, 288 nucleotides of exon 4, 144 nucleotides of exon 5, 132 nucleotides of exon 6, 159 nucleotides of exon 7 and 156 nucleotides of exon 8 were used for the analysis.

2.6 Analysis of exon 1

The direct sequencing of exon 1 did not prove successful as it failed to resolve the GC rich regions of the two alleles and therefore a cloning strategy was adopted. In addition this method would be effective in separating out the parental alleles to assist in analyzing the inheritance pattern. Part I and part II of exon 1 were cloned into the pGEM®-T easy vector using the T/A cloning kit supplied by the (Promega, Madison, WI, USA) and individual clones were subsequently sent for sequencing. See the figure 13 and figure 14 in Appendix C for details on the pGEM®-T easy vector system, the vector map and multiple cloning site.

2.6.1 Purification of amplicons of exon 1 (part I and part II)

The PCR products of exon 1 (part I and II) were resolved on agarose gels of the appropriate volume and concentration as mentioned in the Appendix B. 45 µl of amplified PCR product were mixed thoroughly by adding 2-3 µl of 6x blue/orange loading dye (Promega, Madison, WI, USA) and were loaded into pre-cast agarose gel wells and electrophoresed at ~90 to ~100 volts/cm until the bromo-phenol blue tracking dye front migrated ~75% of the length of the gel.

Visualization of DNA bands was carried out by placing the gel on the UV Transilluminator (Kodak, Electrophoresis documentation and analysis system), with a very low exposure of UV to avoid the formation of pyrimidine dimers. The DNA fragments were excised from the agarose gel using a razor blade and transferred to previously weighed empty 1.5 ml eppendorfs and weighed again to determine the weight of the gel slice, a requirement of the gel purification method. The DNA was purified from the agarose gel using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Germany). All the purification steps were performed following the manufacturer's instructions.

The purified amplicons of exon 1 (part I and II) were quantified using spectrophotometric analysis on a Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc., Wilmington, Del.) and subsequently used in ligation reactions.

2.6.2 Preparation of Chemically Competent Cells

E. coli bacterial strain DH5 α was made chemically competent using the Rubidium Chloride method (Choudhury A *et al.*, 2009).

A 1000 µl overnight culture of DH5 α was inoculated into 100-500 ml of *Psi* broth (Appendix B). The inoculated *Psi* broth was incubated at 37 ºC and shaken at 235 rpm in a orbital shaker incubator (LM-510, Sci-Lab Technology Co. Ltd. China) until an optical density at 600 (OD 600) reached 0.6 - 0.7. Following incubation of this culture on ice for 15 mins, a 3 ml aliquot was centrifuged at 3-5000 rpm for 5 mins using Sorvall SS-34 rotor. The pellet was resuspended with

40-400 ml TfbI buffer (Appendix B) and incubated on ice for 15 mins and re-centrifuged. The pellets was then re-suspended in 0.04 volumes TfbII (Appendix B), incubated on ice for 15 mins and then snap frozen in liquid nitrogen for storage at -70 ºC.

2.6.3 Ligation reactions

Ligation reactions were set up in a 3:1, insert: vector molar ratio using the following formula

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert : vector molar ratio} = \text{ng of insert}$$

All the ligation reactions were performed according to manufacturer's instruction (Promega, Madison, WI, USA). See table 3 (Appendix B) for the ligation reaction conditions. Reactions included a positive control (supplied by manufacturer) and no insert control. The ligation reactions were incubated at 4 ºC overnight to produce maximum number of transformants.

2.6.4 Transformation of ligation products into DH5 α cells using heat shock

We used high efficiency competent cells, DH5 α cells with a transformation efficiency of $1X10^{8}$cfu/µg DNA for transformations in order to obtain a reasonable number of colonies. 2-3 µl of ligated products were transformed into 50 µl DH5 α chemically competent cells, as described (Kushner,1978). To ensure that the transformation results were unambiguous and to assess the efficiency of

ligation numerous controls were prepared and included in the transformation protocol. As a transformation positive control, a known amount of closed circular (uncut) plasmid was used in the transformation. All the steps were carried out on ice. Two µl of each ligated product was added into a 1.5 ml eppendorf tube containing 50 µl of previously thawed DH5α chemically competent cells. The tubes were kept on ice for 20 mins to allow the DNA to precipitate and thereafter the bacteria was heat shocked in a dry bath at 42 ºC for 45-50 sec and then immediately chilled on ice for 2-5 mins. The transformed cells were resuspended in 950 µl of room temperature SOC medium (Appendix B) and were incubated at 37 ºC with ~150 rpm agitation for 90 mins in an orbital shaker incubator (LM-510, Sci-Lab Technology Co. Ltd. China). The cells were pelleted by centrifugation at 1000 rpm for 5 mins. The upper 750 µl of SOC medium was discarded and the pellet was resuspended in the remaining 250 µl of SOC medium. For each ligation reaction two plates of LB ampicillin/IPTG/ X-gal were used. IPTG, in conjunction with X-gal, was used to detect β-galactosidase activity in order to differentiate recombinants from non-recombinants with vector containing the lac Z α-peptide gene) (Promega, Madison, WI, USA) and two plates of LB ampicillin/IPTG/X-gal were used for determining the transformation efficiency. All the plates were prepared at room temperature (Appendix B). Usually 100 µl - 150 µl of the transformation mix was plated onto each plate and they were incubated inverted at 37 ºC overnight in an incubator (Amersham Biosciences, Washington, USA).

2.7 Screening of transformants for inserts

Successful cloning of an insert into the pGEM®-T easy vector interrupted the coding sequence of β-galactosidase gene and the recombinant clones were identified by color screening on LB ampicillin/IPTG/X-gal plates. The PCR products cloned into the vectors significantly affected the ratio of blue: white colonies. The colonies containing the inserted PCR product produced white colonies. We screened 10 to 15 white colonies for each ligation reaction corresponding to a PCR product of exon 1 part I and II for each patients genomic DNA as described in the following steps.

2.7.1 Plasmid extraction

For screening of transformants for inserts, 10 to 15 white colonies of each sample for both part I and part II of exon 1, were inoculated in a 15 ml tubes containing 1-5 ml of Luria broth with 50 mg/ml ampicillin. The tubes were incubated at 37 ºC and shaken at 230-250 rpm in an orbital shaker incubator (LM-510, Sci-Lab Technology Co. Ltd. China) for about 12 hrs-16 hrs. The next day bacterial cultures were harvested by centrifugation at 6000 rpm for 2 mins at room temperature. The supernatant was decanted and all traces of remaining LB medium removed. High quality plasmid DNA was extracted using Gene JET plasmid-mini prep kit which is based on the SDS/alkaline lysis method (Promega, Madison, WI, USA).

2.8 Checking of positive clones by restriction digestion analysis

All isolated plasmid DNA was subjected to restriction digestion analysis to screen for and identify positive clones. The pGEM®-T easy vector contains numerous unique restriction sites within the multiple cloning sites. See figure 14 for the sequence and multiple cloning site of the vector (Appendix C). Plasmid restriction was carried out using the Fast digest enzyme *Not I* (Fermentas, USA). The reaction was set up for 10 µl and incubated at 37 ˚C in a dry heat block for 50-60 mins. Refer the table 4 for restriction digestion reaction conditions (Appendix B). The thermal inactivation of the restriction enzyme *Not I* was achieved by incubating the reaction at 80 ˚C for 5 mins. All the steps were performed in accordance to the manufacturer's instructions. An uncut control (DNA without enzyme) was included in the analysis to compare digest patterns and to identify bands corresponding to uncut vector. The entire restriction reaction was resolved on a 1% agarose gel (Appendix B) as described previously (Section 2.4.3). Images of the gels were recorded using the Gel Doc XR Imaging System (Bio-Rad) and used to identify the plasmids containing insert.

2.9 Sequencing of plasmid DNA of exon 1 (part I and part II) and Analysis

Six to eight positive clones of each sample of exon 1 (part I and part II) were selected and sent for sequencing. The plasmid DNA was sequenced using the ABI 3730xl DNA analyzer (Applied Biosystems, MACROGEN, Seoul, Korea) using

two universal primers, the T7 forward promoter and the M13 reverse primer of the pGEM®-T easy vector.

Sequences were subjected to a standard nucleotide-nucleotide BLAST search (http://www.ncbi.nlm.nih.gov/Blast.cgi) to ensure that the amplicons amplified were the human androgen receptor gene, and 1614 nucleotides of exon 1 was used for the analysis, carried out as described before in Section 2.5.

CHAPTER 3

3. RESULTS

3.1 Amplification of exons 1-8 of androgen receptor gene, sequencing and analysis

Genomic DNA was extracted from all sixteen members of the family who agreed to participate in the genetic study. Each exon of the AR gene was amplified successfully as a single PCR product, using single set of primer listed in the (table 2), except for the exon 1, which was amplified as two PCR products (part I and part II) using 2 sets of primer (listed in the table 2). Amplified products of each exon were gel electrophoresed and an image of each gel was recorded. A clear band of the predicted size (given in table 2) of each exon was observed after the gel electrophoresis. A gel picture of each exon is shown corresponding from figure 15 to figure 22.

Amplified products of exons 2-8 were sequenced directly, using the same set of primers, as were used for the PCR (table 2). The amplified products of exon 1 (part I and II) were cloned, as described in the section 3.2 below.

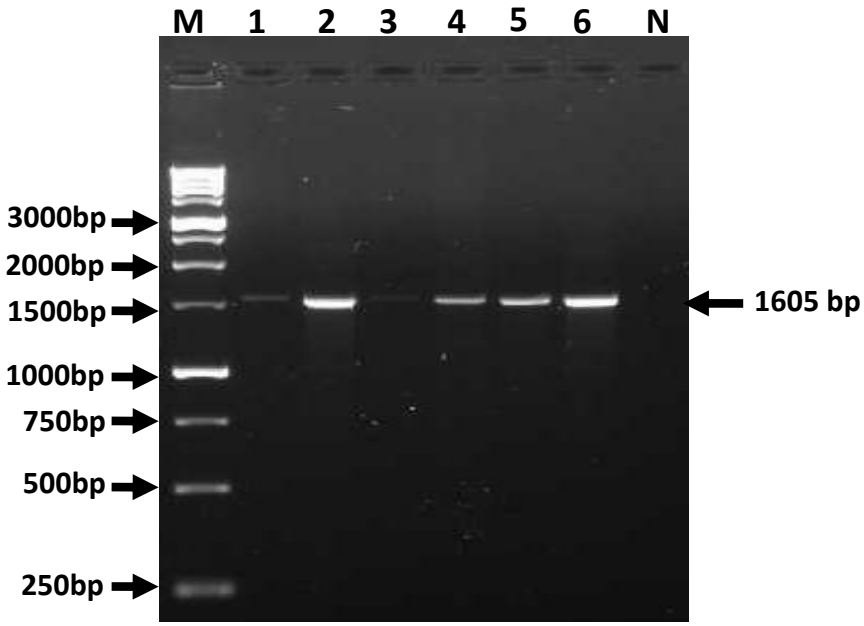


Figure 15: Amplification of exon 1 of AR gene, Lane 1-6 is positive amplification of a 1600 bp DNA band, M=1kb DNA ladder and N=negative control.

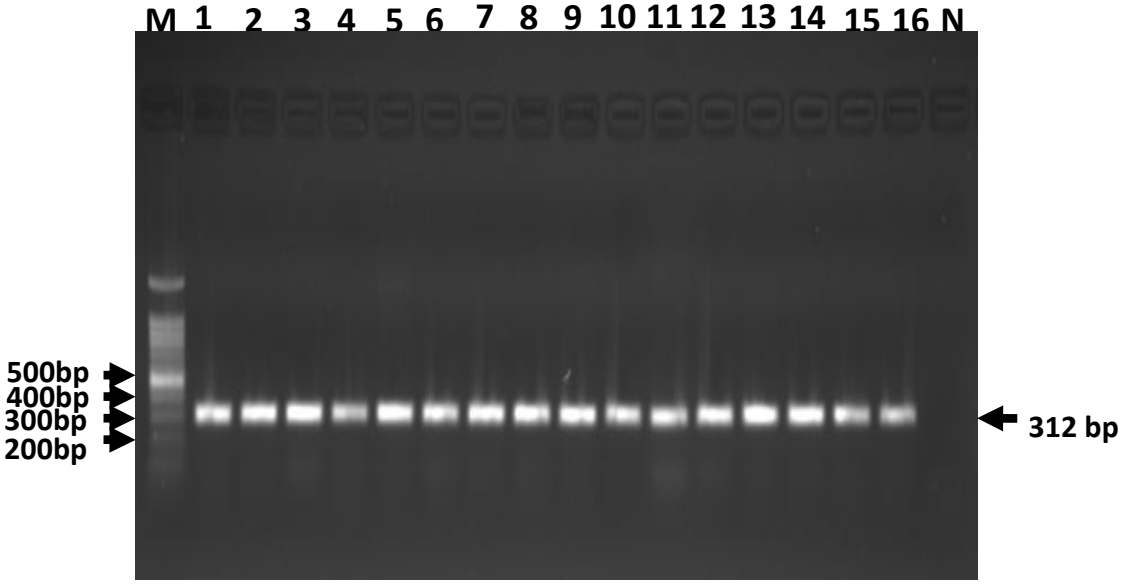


Figure 16: Amplification of exon 2 of AR gene, Lane 1-16 is positive amplification of a 312 bp DNA band, M = 100 bp DNA ladder and N = negative control.

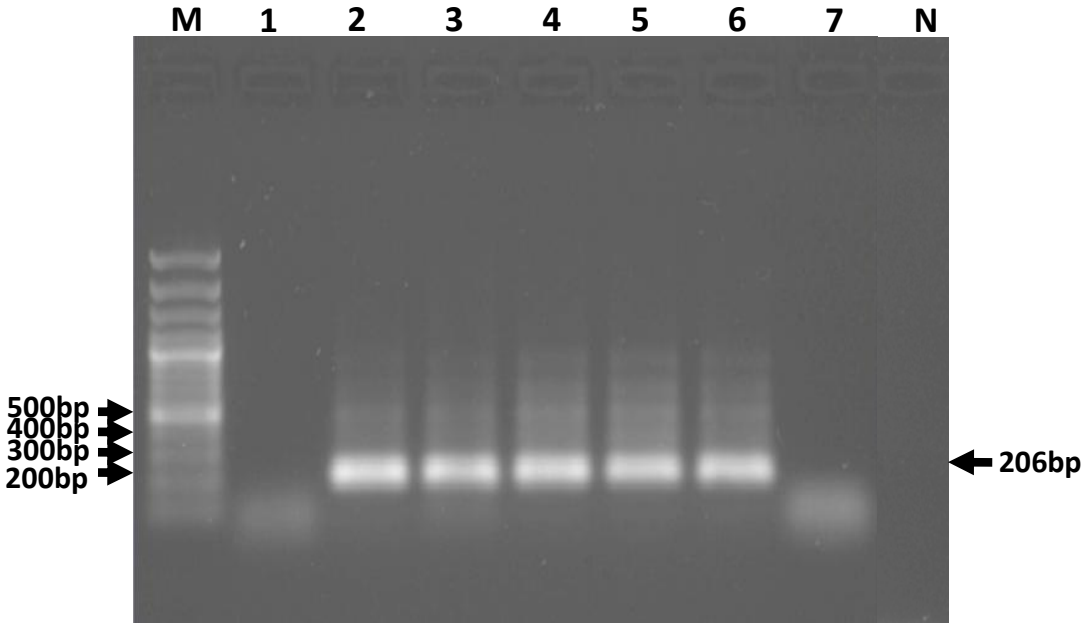


Figure 17: Amplification of exon 3 of AR gene, Lane 1-7 is positive amplification of a 206 bp DNA band, M = 100 bp DNA ladder and N = negative control.

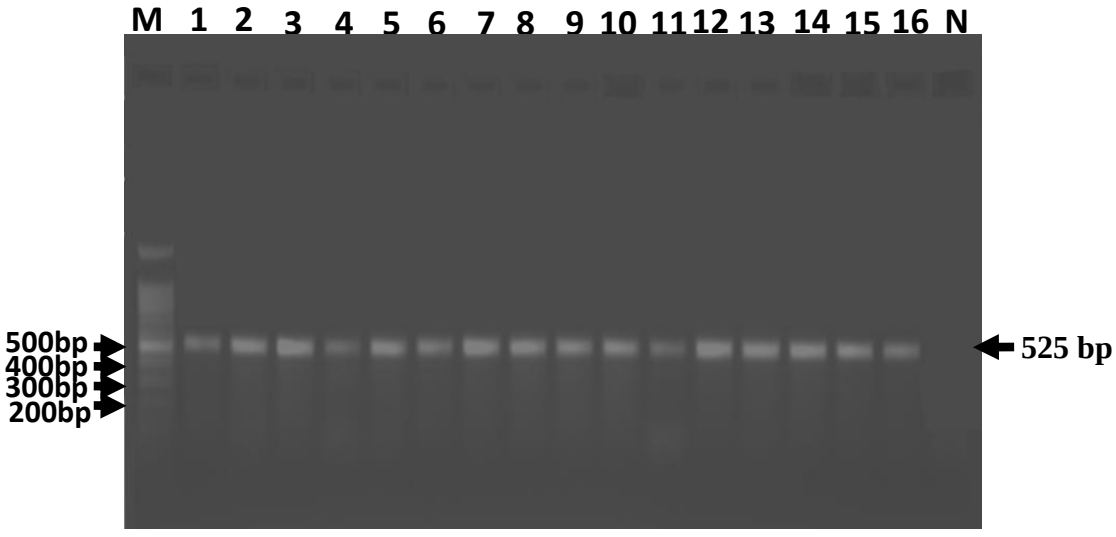


Figure 18: Amplification of exon 4 of AR gene, Lane 1-16 is positive amplification of 525 bp DNA band, M = 100 bp DNA ladder and N = negative control.

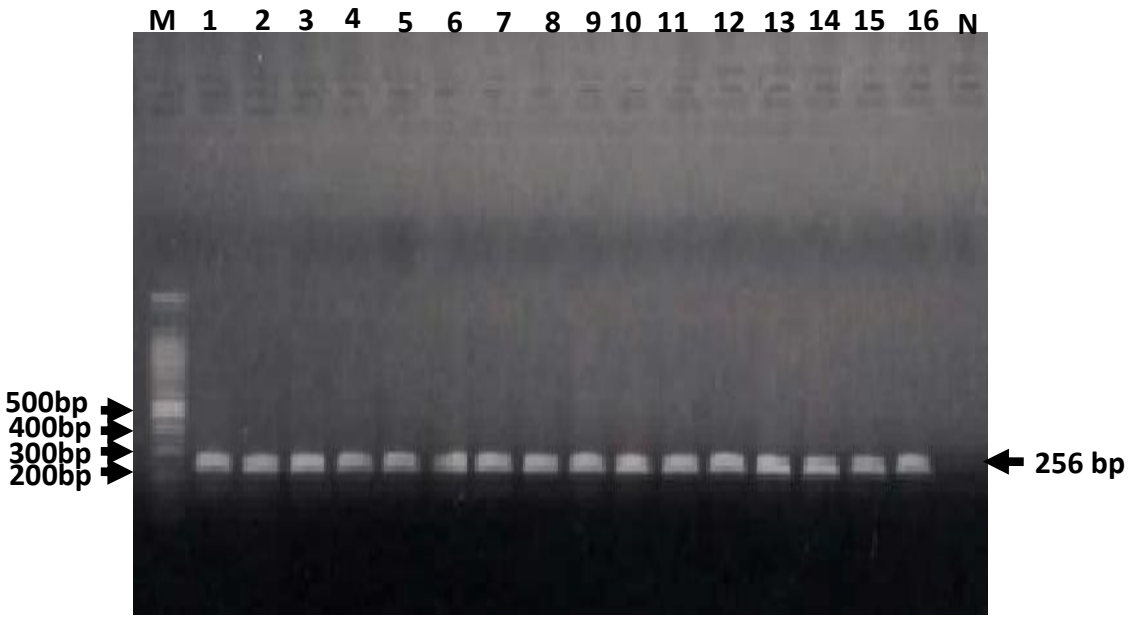


Figure 19: Amplification of exon 5 of AR gene, Lane 1-16 is positive amplification of a 256 bp DNA band, M = 100 bp DNA ladder and N = negative control.

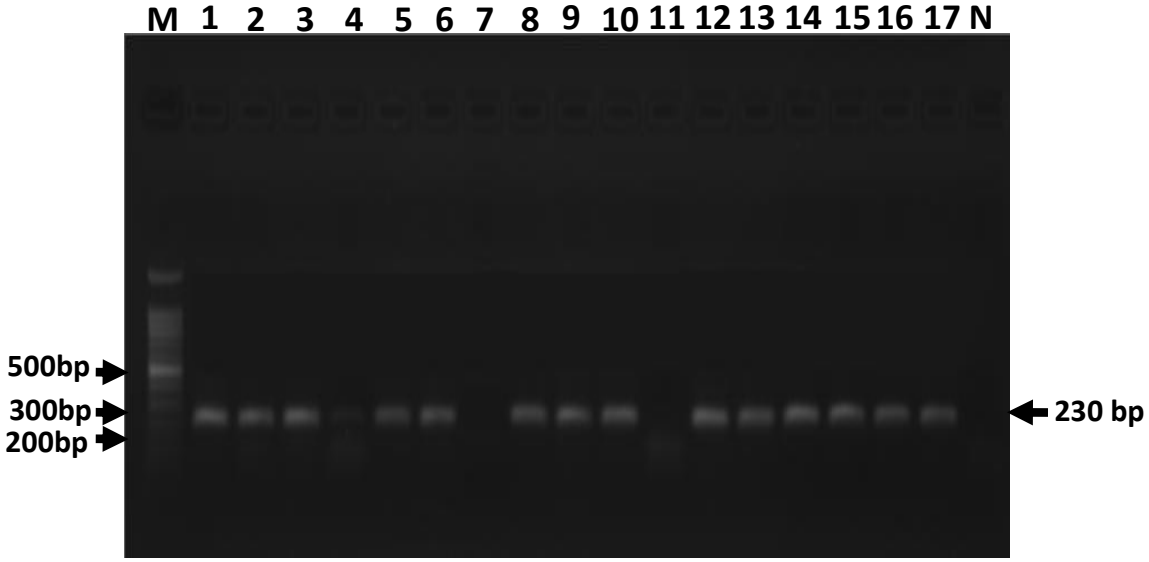


Figure 20: Amplification of exon 6 of AR gene, except lane 7 and 11 from lane 1-17 is positive amplification of a 230 bp DNA band, M = 100 bp DNA ladder and N = negative control.

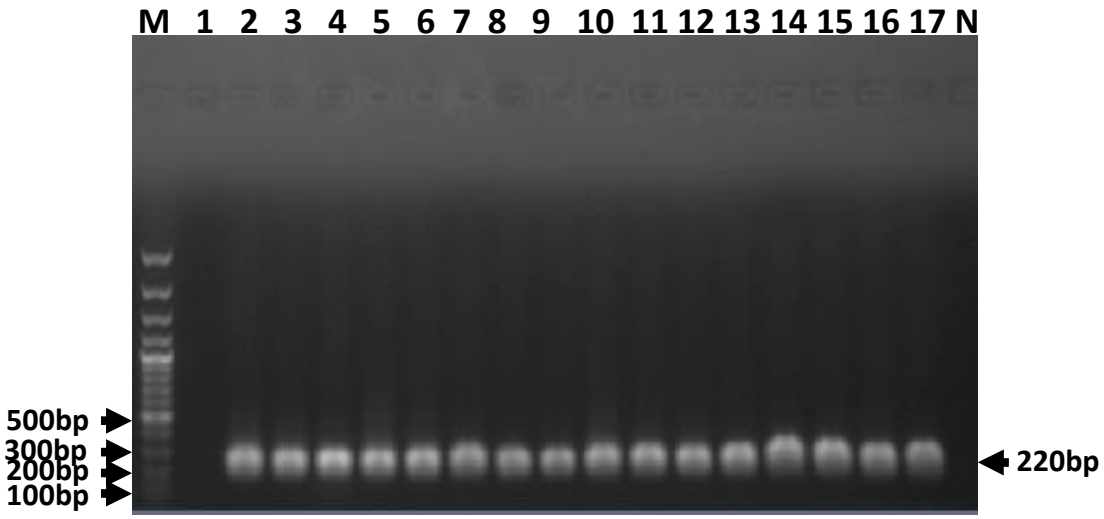


Figure 21: Amplification of exon 7 of AR gene, Lane 1-16 is positive amplification of a 220 bp DNA band, M = 100 bp DNA ladder and N = negative control.

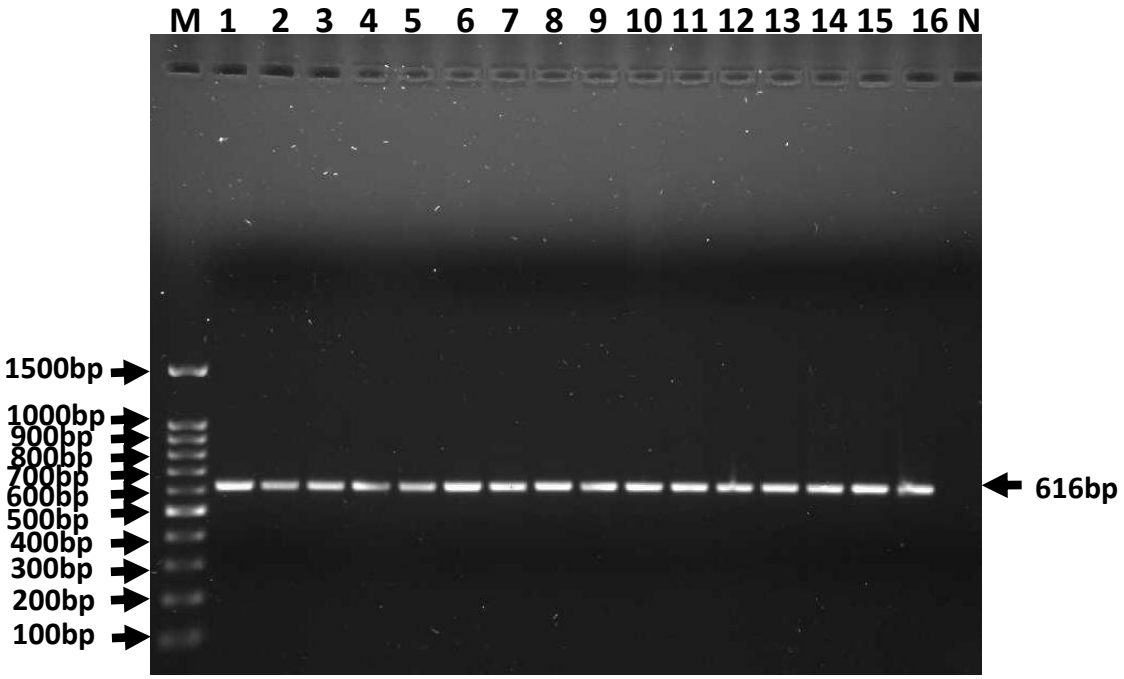


Figure 22: Amplification of exon 8 of AR gene, Lane 1-16 is positive amplification of a 616 bp DNA band, M = 100 bp DNA ladder and N = negative control.

3.2 Cloning of part I and II of exon 1 of AR gene

The PCR products of exon 1 (part I and II), of all the 16 members of the family were gel purified, quantified and ligated in 3:1, insert: vector ratio into a pGEM-T easy vector individually. Following transformation positive clones were checked by restriction enzyme digestion using Fast digest enzyme *Not I*. Digestion with *Not I* cuts the vector twice to produce to two clear bands: one band corresponding to vector sequence (3015 bp) and another the insert (of either 1605 bp for part I and 1533 bp for part II of exon 1). An image of each gel was recorded for both part I and II of exon 1 and is shown in the figures 23 and 24 respectively.

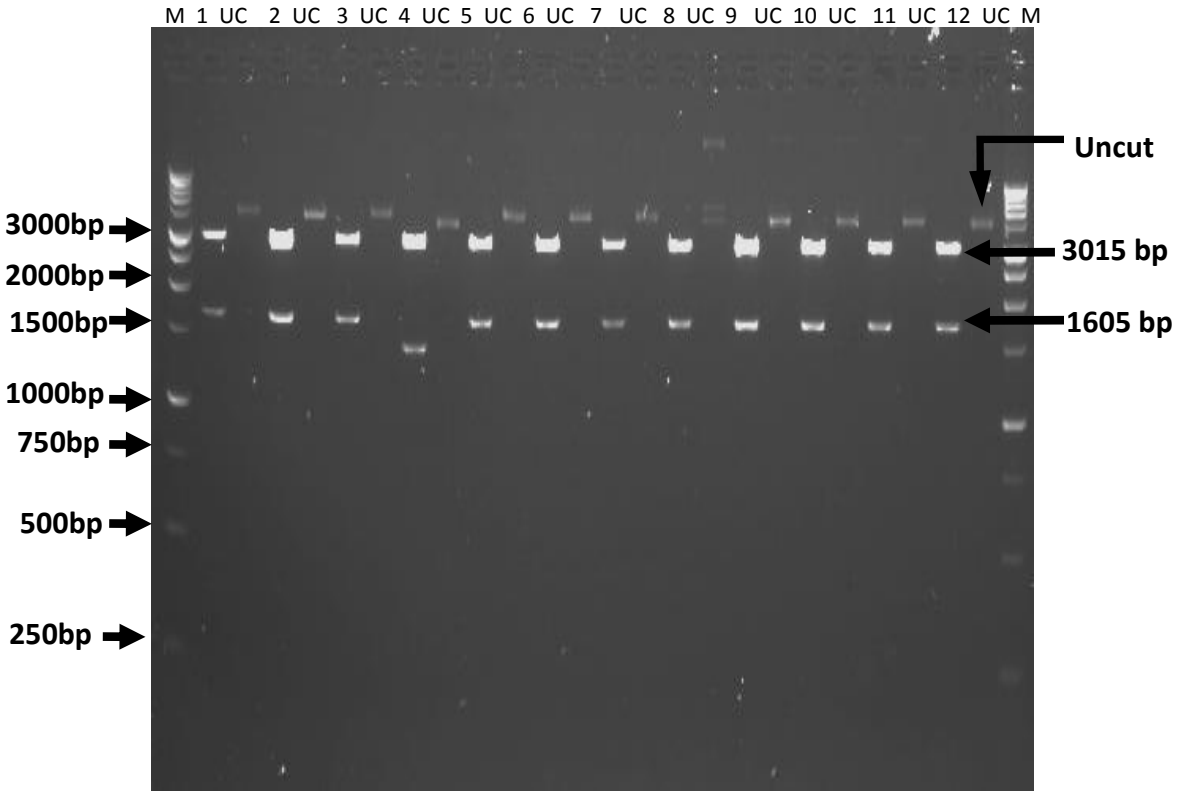


Figure 23: M = 1Kb gene ruler, Restriction mapping is done by *Not I*. Positive clone (of part I of exon 1) is shown by a 1605 bp DNA band and the pGEM-T easy vector DNA band is 3015 bp. UC= uncut.

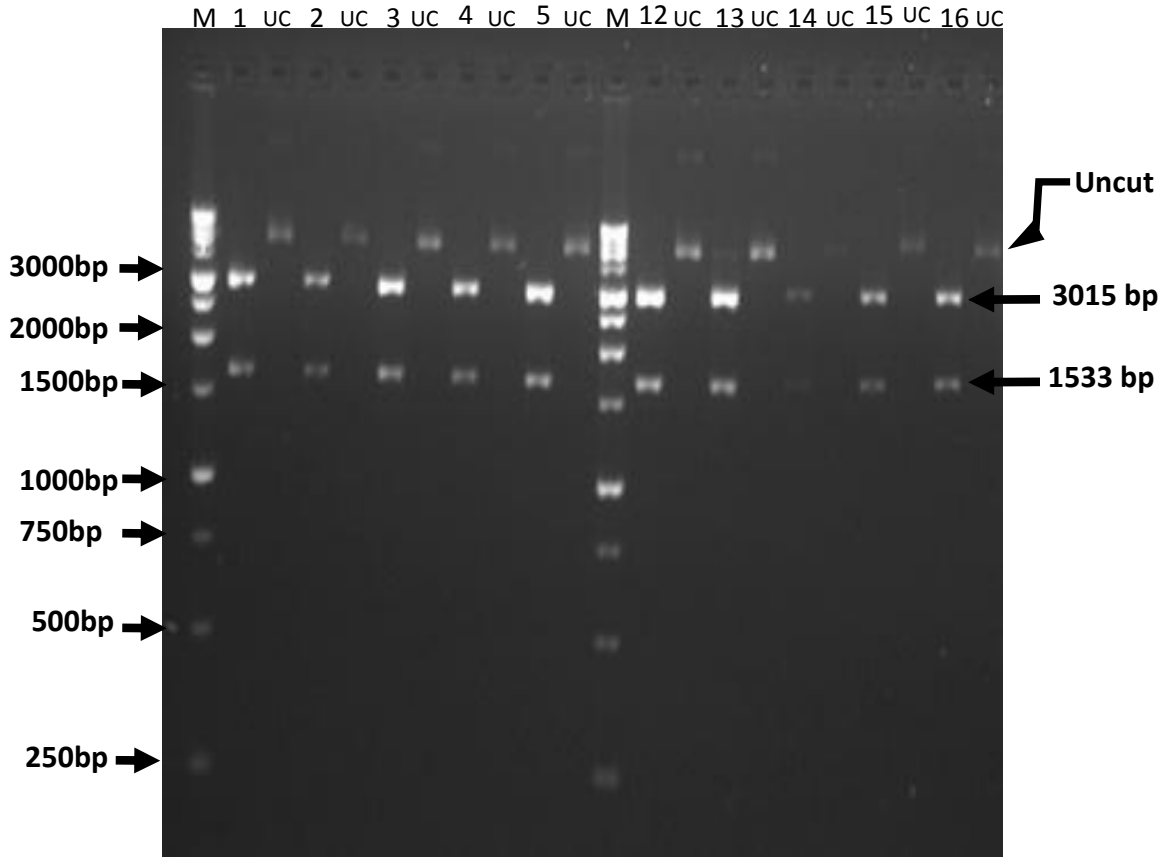


Figure 24: M=1Kb Gene ruler, *Not I* Restriction map of the pGEM vector. Positive clone (of part II of exon 1) is shown by 1533 bp DNA band and the pGEM vector DNA band is shown by 3015 bp and UC= uncut.

3.3 Sequencing of plasmid DNA of exon 1 (part I and part II) and Analysis

Six to eight positive clones of each sample of exon 1 (part I and part II) were sequenced using the two universal primers, T7 promoter forward and M13 reverse primer of the pGEM®-T easy vector and each part analyzed separately as described in section 2.5.

3.4 Results of CAG repeat length (Poly-glutamine tract variations)

3.4.1 Defining and dividing CAG repeat length of the androgen receptor gene

The length of poly Q varies from 14 to 35 repeats, with an average of 21 ± 2 repeats (Irvine *et al*., 1995). Analysis of CAG repeat length was divided into 3 categories based on their length as follows;

- (CAG)n repeat length < 23 were classified as short **(S)**.
- (CAG)n repeat length = 23 were classified as wild type **(WT)**.
- (CAG)n repeat length > 23 were classified as long **(L)**.

The CAG variation and superimposed poly-glutamine stretch among affected, unaffected and control individuals are shown in the table 5 and 5a.

Table 5: Representation of the addition and deletions of the Poly-glutamine tract identified in all 16 members of the family. M stand for Male and F stand for Female. Q stands for amino acid glutamine. Empty triangle represents deletion and solid black downward triangle represents addition of amino acid. ID- identifies each family member on the pedigree.

				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23			
	ID	Sex	Category	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q			
1	C	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	△	△			
2	E	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼		
3	H	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	△			
4	I	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	△			
5	K	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼		
6	L	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼		
7	d	F	BC Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼	▼	
8	B	M	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	△			
9	F	F	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·			
10	b	M	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼	▼	
11	c	F	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼	▼	
12	A	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·			
13	D	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	△	△	△	△			
14	G	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼	▼	▼
15	J	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·			
16	a	M	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·			

Figure 11 (a) Table 6 superimposed on pedigree as given in figure 11 Addition and deletions of the Poly-glutamine tract identified in all 16 members of the family.

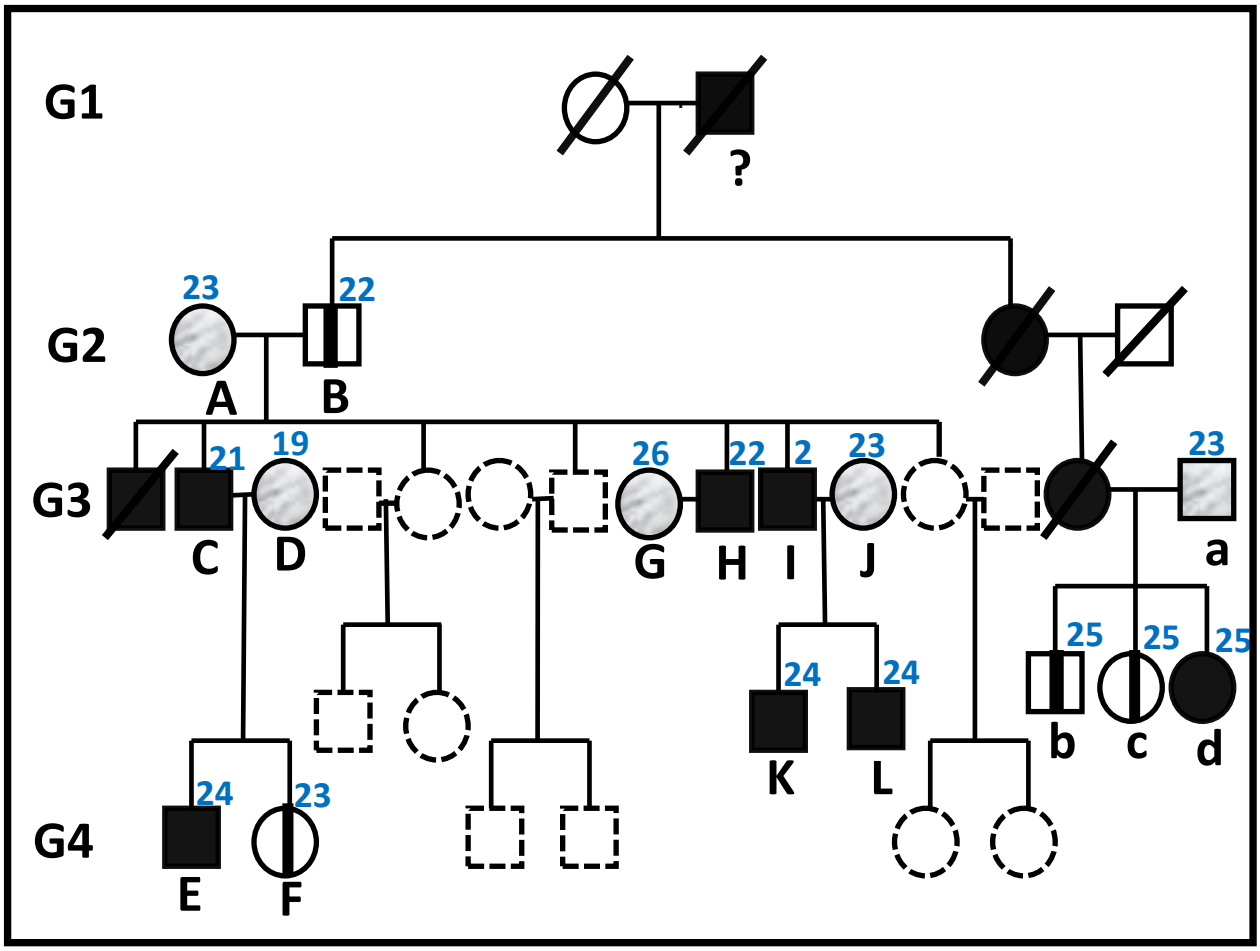


Figure 11a: ***Pedigree of family kindred with MAIS and Breast Cancer and correlation with poly-glutamine CAG repeat length identified in present study:*** Circles and squares indicate Females and Males, respectively. G1, G2, G3 and G4 indicate the generations. Solid black circles and squares indicate the affected subjects. A solid vertical black line inside the circles and squares indicate the unaffected family members. Light-shaded circles and squares indicate the controls. Circles and squares with a cross represent the deceased subjects and the subjects who refused to participate in the study are indicated by dotted circles and squares. There are two main arms of the study, left arm (capital alphabet) and right arm (small letters) and these two arms bifurcates in G2. Right arm (small letters) family members were affected with breast cancer and in left arm (capital alphabet) family members are affected with MAIS, gynecomytia. The number over each subject indicates the length of the CAG repeats identified in the present study

3.4.2 Analysis of results of poly-glutamine stretch within the exon 1 region of the AR gene

Four generations, with a total of 16 members of the family, which were segregated into two arms i.e. MAIS affected males in the MAIS arm and breast cancer affected females in the BC arm, were placed into three groups:

(1) Affected members of same lineage [(n = 7) (M: 6; F: 1)], six males ('C', 'E', 'H', 'I', 'K', 'L') had MAIS and gynecomastia, and one female ('d') had breast cancer.

(2) Unaffected members of the same lineage, who were under observation [(n = 4) (M: 2 F: 2)], two males ('B' and 'b') and two females ('F' and 'c').

(3) Unrelated normal members, which served as controls [(n =5) (M: 1 F: 4)] were a single male ('a') and four females ('A', 'D', 'G' and 'J').

3.4.3 Detection of variation in the poly-glutamine tract (CAG-repeat).

The variations of poly-glutamine stretch among affected members were (**S**: 3, **L**: 4). A short poly Q stretch was observed in three males affected with MAIS ('C', 'H', and 'I'), whereas the sequences of the remaining four affected members revealed long poly Q stretch, of which three were males ('E', 'K' and 'L') with MAIS and one female ('d') with breast cancer. Based on these results, it seems that poly Q stretch is variable and is not particularly associated with MAIS. Since there was a single female with breast cancer, we tried to correlate the results of poly Q stretch with controls and unaffected members of the family. On analysis,

the CAG repeat length of un-affected family members (n = 4) revealed wild type CAG repeat sequences in one, CAG repeat was longer in 2 and shorter in 1 member (**WT**: 1, **L**: 2, **S**:1). Among the four unaffected members, one male 'b' and one female 'c' were identified with a longer poly Q stretch. One male 'B' was identified with a shorter poly Q stretch and one female 'F' was identified as completely wild type. Of the five controls in the family, we identified one female ('D') with a short poly Q stretch, another female with a long poly Q stretch ('G') and three controls including two females ('A' and 'J') and one male ('a') were completely wild type. Interestingly, in the 3rd generation all affected members were identified with only **S** CAG repeats and in the 4th generation all the affected members and two unaffected members had **L** CAG repeat.

3.4.4 (a) Poly-glutamine stretch among the 2nd generation: There were only 2 members in the 2nd generation of the MAIS arm, a single control female and a single unaffected normal male. The control female ('A') was completely wild type, whereas an unaffected male 'B' showed a **S** CAG repeat,. In the BC arm, two members (a normal male and breast cancer affected female) died and were not included in the study.

3.4.5 (b) Poly-glutamine stretch among the 3rd generation: In the 3rd generation, there were 6 members (3 affected males and 3 control females) in the MAIS arm of the family. Poly-glutamine repeat analysis revealed CAG **S** repeats among all three MAIS affected males ('C', 'H' and 'I'), and among controls, one of each female ('J'), ('D') and ('G') control were identified as **WT**, CAG **S** and CAG **L**, whereas in the BC arm there was only a single control male ('a') member who was

completely **WT** and his partner, who died as a result of breast cancer and was not included in the study.

3.4.6 (c) Poly-glutamine stretch among the 4th generation: In the 4th generation, there were 4 members (3 affected males and a single unaffected female) in the MAIS arm of the family. Poly-glutamine analysis revealed CAG **L** repeat among all 3 MAIS affected males ('E', 'K' and 'L'), and a single unaffected female ('F') was completely **WT,** whereas in the BC arm, there were 3 members, 1 affected female (d'), 1 unaffected female ('c') and 1 unaffected male ('b'). All of them showed CAG **L** repeat and they were of the same genetic lineage. Refer to table 5.

3.5 Results of GGN repeat length (Poly-glycine tract variations)

3.5.1 Defining and dividing GGN repeat length of the androgen receptor gene

For the length of the poly-G repeats, the consensus appears to be that normal phenotypic males have 10 to 30 repeats in most of the population (Rajender *et al*., 2007). The functional consequences of variations in GGN repeat are less clear, the deletion of the poly-glycine tract reduces AR transcriptional activity in transient transfection assay (Gao *et al*., 1996). AR GGN repeat length polymorphism has been less studied in MAIS and BC. Variation in length can be associated with prostate cancer, but any relationships between the repeat length and disease characteristics are not clear (Montgomery *et al*., 2001; Edwards *et al*., 1999).

Analysis of GGN repeat length was also divided into 3 categories, based on their length as given below:

- (GGN)n repeat length < 23 were classified as short **(S)**.
- (GGN)n repeat length = 23 were classified as wild type **(WT)** and
- (GGN)n repeat length > 23 were classified as long **(L)**.

The GGN variations and superimposition of poly-glycine stretch among the affected, unaffected and controls are shown in the table 6 and 6a.

Table 6: Representation of addition and deletions of the poly-glycine tract in all 16 members of the family. M stand for Male and F stand for Female. G stands for amino acid glycine. Empty triangle represents deletion and solid black downward triangle represents addition of amino acid. ID- identifies each family member on the pedigree.

				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
	ID	Sex/Age	Category	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
1	C	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	Δ	Δ	Δ	Δ	
2	E	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	
3	H	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	Δ	Δ	Δ	Δ	
4	I	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	
5	K	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	Δ	
6	L	M	MAIS Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	
7	d	F	BC Affected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	
8	B	M	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	
9	F	F	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	Δ	Δ	Δ	Δ	
10	b	M	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	
11	c	F	Unaffected	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	
12	A	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	Δ	Δ	Δ	Δ	
13	D	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	
14	G	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	▼
15	J	F	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	Δ	Δ	
16	a	M	Control	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Δ	Δ	

Figure 11 (b) Table 6 superimposed on pedigree as given in figure 11: Addition and deletions of the Poly-glycine tract identified in all 16 members of the family.

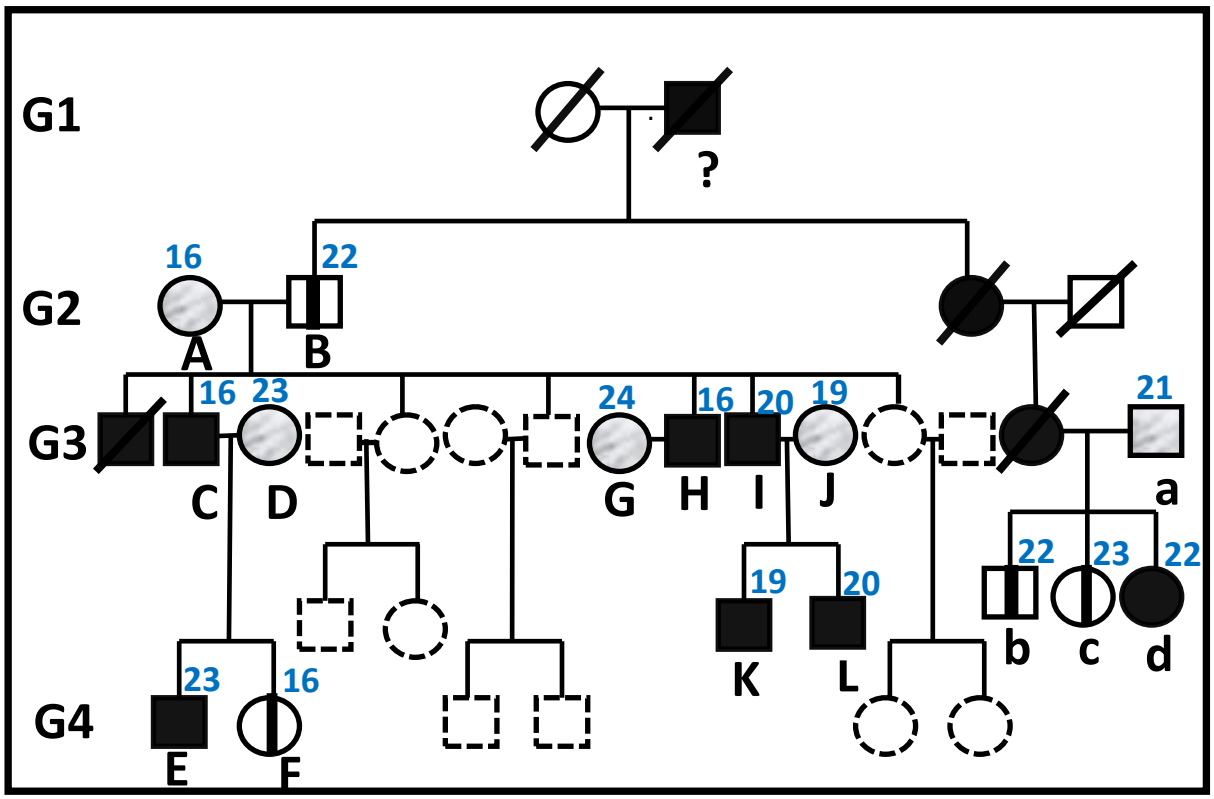


Figure 11b: ***Pedigree of family kindred with MAIS and Breast Cancer and correlation with poly-glycine GGN repeat length identified in present study:*** Circles and squares indicate Females and Males, respectively. G1, G2, G3 and G4 indicate the generations. Solid black circles and squares indicate the affected subjects. A solid vertical black line inside the circles and squares indicate the unaffected family members. Light-shaded circles and squares indicate the controls. Circles and squares with a cross represent the deceased subjects and the subjects who refused to participate in the study are indicated by dotted circles and squares. There are two main arms of the study, left arm (capital alphabet) and right arm (small letters) and these two arms bifurcates in G2. Right arm (small letters) family members were affected with breast cancer and in left arm (capital alphabet) family members are affected with MAIS, gynecomytia. The number over each subject indicates the length of the GGN repeats identified in the present study

3.5.2 Analysis of results of the poly-glycine stretch within the exon 1 region of the AR gene

We have identified, that the androgen receptor (GGN)n variations also differed among individuals and generations. Data analysis suggests that **S** GGN repeats [12/16 (75%), A:6, U:3, C:3] in a majority of the family members; however, only one [1/16 (6%) , A:0, U:0, C:1] of the family member showed **L** GGN repeats and three [3/16 (19%), A:1, U:1, C:1] of the family members showed **WT** GGN repeats.

3.5.3 Generation specific analysis of Poly-glycine tract (GGN-repeat)

3.5.3 (a) Poly-glycine stretch among the 2nd generation: Among two members of the MAIS arm in the 2nd generation, a single control female ('A') was identified as GGN **S** and a single unaffected normal male ('B') also showed GGN **S** repeat.

3.5.4 (b) Poly-glycine stretch among the 3rd generation: The three affected males ('C', 'H' and 'I') of the MAIS arm revealed GGN **S** repeat. Whereas among the 3 control females of the MAIS arm, one ('G') was GGN **L**, one ('J') was GGN **S** and the other female control ('D') was identified as **WT**. However, a single control male ('a') of the BC arm also showed a GGN **S** repeat and his partner, who died as a result of breast cancer, was not included in the study.

3.5.5 (c) Poly-glycine stretch among the 4th generation : In the 4th generation, among the 3 affected males of the MAIS arm, two affected males ('K' and 'L') showed GGN **S** repeat and the other affected male ('E') was **WT**. A single unaffected female ('F') of the MAIS arm was identified as GGN **S**, whereas, in the BC arm, 1 affected female ('d') and 1 unaffected male ('b') were identified as GGN **S**, and the unaffected female ('c') was completely **WT**.

CHAPTER 4

4. Results of androgen receptor gene point mutations

4.1 Analysis of sequencing data of all the exons of the AR gene

Analysis of point mutations was conducted on the complete androgen receptor gene covering all the exons (1-8). We analyzed 1614 nucleotides of exon 1, 153 nucleotides of exon 2, 117 nucleotides of exon 3, 288 nucleotides of exon 4, 144 nucleotides of exon 5, 132 nucleotides of exon 6, 159 nucleotides of exon 7 and 153 nucleotides of exon 8. The 16 members of the family with four generations were grouped into three as described in section 3.4.2.

All the nucleotide changes detected were translated into amino acids using the software ExPASy translator tool. The aligning of the nucleotide sequences and the amino acid sequences were carried out using Bioedit and Clustal Omega software's. The recent human AR gene sequence from the Gene Bank ID NM_000044.2 was used as a reference.

4.2 Analysis of exons 2-8 of androgen receptor gene

No mutations were identified in any of the family members in the exons 2-8 region of the androgen receptor gene. All sequences were completely wild type and well conserved.

4.3 Identification of point mutations in exon 1

We analyzed 5-6 clones for part I and II of exon 1 in all 16 members of the family. Each mutation reported were identified in at least 2 of the clones for each member. None of the mutations were identified in all 6 of the clones, however a few of the mutations were present ranging from 33.3%-66.6% of the clones. Mutational analysis revealed several important AR gene mutations in 16 members of the family. A total of 43 single nucleotide substitutions in exon 1 region were identified. Based on the analysis; we divided these mutations into two parts:

(1) Novel mutations and

(2) Previously reported mutations.

All the mutations were checked and matched with the updated version of the McGill University androgen receptor mutation database http://androgendb.mcgill.ca, (Gottlieb *et al*., 2012). The mutations obtained in the control members were matched with the reports of normal individuals with mutations in the AR gene of the ARDB data base. In addition, we also cross checked all the mutations identified in the present study with mutations reported in the PubMed literature database of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/pubmed).

Of these 43 point mutations, we classified 30 as novel mutations and 13 were previously reported. Furthermore, the distribution of each mutation among the members of the family is described below and is based on the clinical state and genetic lineage.

The distribution of all the novel and reported mutations of the affected members of the family were/are listed in detail in the table 7 and those of unaffected and control members of the family were listed in the table 7a.

Except for three mutations, none of the mutations were common among any of the family members; however, some of the mutations were coexisting. ‘Common’ was defined as a similar mutation at the same position in two different members and the ‘co-existing’ mutations could be double, triple and quadruple mutations identified in a single member.

Table: 7 Details of the complete novel and previously reported point mutations in affected members of the family. An ID code has been assigned for each affected member and is given in the extreme left column. M is Male and F is Female.

Member code (category)	Sex	Novel mutations at amino acid level (clones)	Novel mutations at nucleotide level	Reported mutations at amino acid level		Reported mutations at nucleotide level	References of reported mutations
C (MAIS Affected)	M	L 108 P(2) E 113 K (4)	CTG→CCG GAA→AAA	E 155 G (4) E 289 G (3) Q 445 R (3)		GAG →GGG GAA→GGA CAG→CGG	Copelli *et al.,* Holterhaus *et al.,* Gottlieb *et al.,*
E (MAIS Affected)	M	T 231 I (2)	ACC→ATC	G 456 D (3)		GGT→GAT	Steinkamp *et al.,*
H (MAIS Affected)	M	Q 78 R (4) E 187 G (3) L 345 P (3)	CAG→CGG GAG→GGG CTG→CCG				Novel Mutation
I (MAIS Affected)	M	T 105 I (5) C 240 R (2)	ACA→ATA TGT→CGT				Novel Mutation
K (MAIS Affected)	M	E 32 V (2) Q 360 R (3) R 386 C (3)	GAA→GTA CAG→CGG CGC→TGC	S 221 F (2)		TCC→TTC	Audi *et al.,*
L (MAIS Affected)	M	N 25 S (4) L 347 F (4)	AAT→AGT CTC→TTC				Novel Mutation
d (Breast cancer affected)	F	V 132 A (3) V 245 M (4) W 527 R (2) M 528 V (3)	GTC→GCC GTG→ATG TGG→CGG ATG→GTG				Novel Mutation

Table: 7a Details of the complete novel and previously reported point mutations in unaffected and control members of the family. An ID code has been assigned for each affected member and is given in the extreme left column. M is Male and F is Female.

Member code (category)	Sex	Novel mutations at amino acid level (clones)	Novel mutations at nucleotide level	Reported mutations at amino acid level	Reported mutations at nucleotide level	References of reported mutations
B (Unaffected)	M	P 153 L (2) L 389 P (3)	CCG→CTG CTG→CCG			Novel Mutation
F (Unaffected)	F	Q 72 R (3) Q 78 R (3) G 326 D (4)	CAG→ CGG CAG→ CGG GGC→GAC	T 440 A (4)	ACA→GCA	Steinkamp *et al.*,
b (Unaffected)	M	M 191 T (4)	ATG→ACG	Q 28 Stop (3) E 442 K (4)	CAG→TAG GAA→AAA	Katayama *et al*., Sun *et al., & Luo et al.,*
c (Unaffected)	F	G 173 S (3) P 422 S (2)	GGC→AGC CCC→TCC	Y 225 C (3) G 525 D (4)	TAC→TGC GGC→GAC	Hannema *et al.*, Hyytinen *et al.*,
A (Control)	F	Complete wild	Complete wild	Complete wild	Complete wild	Complete wild
D (Control)	F			Y 481 C (4) T 82 A (3)	TAC→TGC ACT→GCT	Gottlieb *et al ; Hannema et al.,& Yeh et al.*, Philibert et al.,
G (Control)	F	Q 79 R (3) A 372 T (4)	CAG→CGG GCC→ACC			Novel Mutation
J (Control)	F	V 33 A (3) Q 98 R (4)	GTG→GCG CAA→CGA			Novel Mutation
a (Control)	M	P 49 L (3) Q 79 R (4) S 361 G (3)	CCT→CTT CAG→CGG AGT→GGT	Q 488 R (3) E 155 G (3)	CAG→CGG GAG→GGG	Boehmer *et al.*, Copelli *et al.*,

4.3.1 Point mutations identified in affected members

In affected members of the pedigree (n=7) we identified a total of 22 point mutations, of which 17 mutations were novel and 5 were previously reported mutations. Of the 17 novel mutations, 13 were present in 6 of the MAIS affected males: L108P (CTG→CCG), E113K (GAA→AAA), T231I (ACC→ATC), Q78R (CAG→CGG), E187G (GAG→GGG), L345P (CTG→CCG), T105I (ACA→ATA), C240R (TGT→CGT), E32V (GAA→GTA), Q360R (CAG→CGG), R386C (CGC→TGC), N25S (AAT→AGT) and L347F (CTC→TTC).

There are five previously reported mutations which were found in these affected males at the position E155G (GAG→GGG), E289G (GAA→GGA), Q445R (CAG→CGG), G456D (GGT→GAT) and S221F (TCC→TTC).

Four novel mutations were identified in the breast cancer affected female at the positions V132A(GTC→GCC), V245M(GTG→ATG), W527R(TGG→CGG) and M528V(ATG→GTG); however none of the mutations were previously reported.

4.3.2 Point mutations identified in AF-1 domain of exon 1 in affected members

Interestingly, four mutations (3 novel and 1 reported mutations) were identified in tightly conserved residues (aa 224-258) of the activation function-1 (AF-1) domain of AR exon1. The three novel mutations were reported as one each in three patients. One unreported mutation was also present in a genetically related unaffected member in this highly conserved AF-1 domain. Of the three patients

two were affected with MAIS and the third was the female who developed breast cancer. Analysis revealed mutations at positions T231I and C240R in the two affected males ('E' and 'I') with MAIS and gynecomastia and a single mutation at position V245M was identified the female ('d'). The single mutation in the genetically related unaffected family member was at the position Y225C (unaffected female 'c'). The position of these mutations is shown in the figure 25. Electropherograms of all novel mutations identified in affected members are shown in the figures 26 to 41 (Appendix D) and those of reported mutations are shown in the figures 42 to 46, respectively in Appendix D.

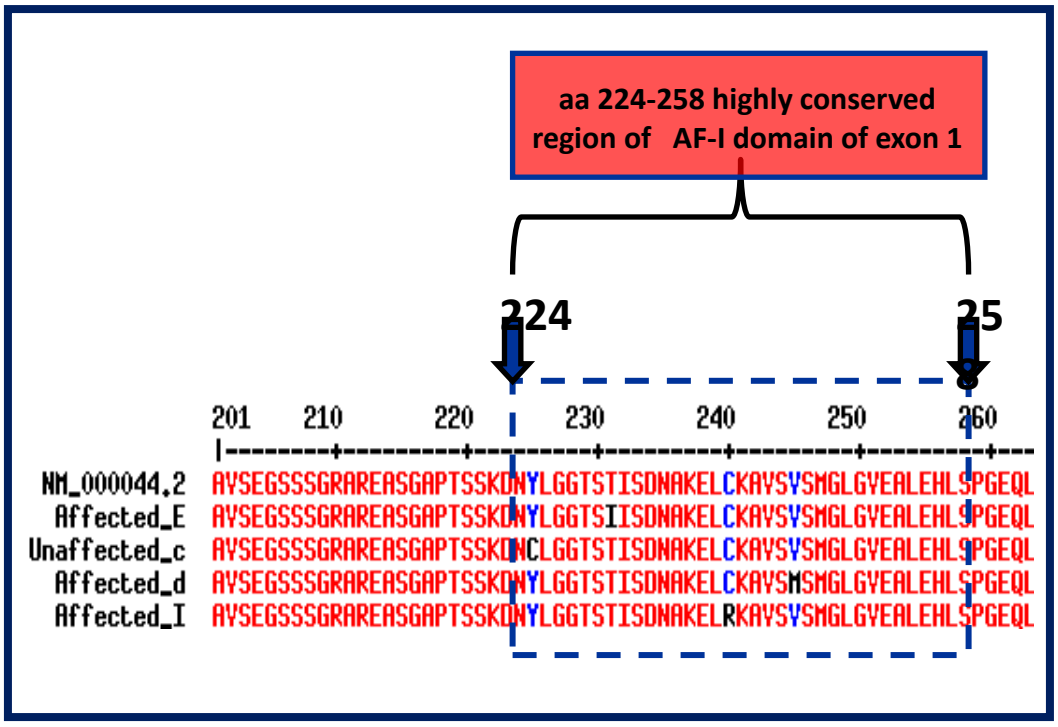


Figure: 25 Highly conserved region of AF-I domain of exon 1, lies between the aa 224-258. Three affected members (M:2, F:1) ‘d’, ’E’, ‘I’ and one unaffected female ‘c’ showed mutation with in this region at the position T231I, C240R, V245M and Y225C showed in black font.

CHAPTER 5

5. DISCUSSION

5.1 Poly-glutamine tract variations in the androgen receptor gene

The family described here is unique in that there is segregation of MAIS and breast cancer in different lines of the pedigree and to our knowledge this has not been reported previously. The clinical features and characteristics of the MAIS and breast cancer in the family together with mild infertility-related issues dictated that we explore the AR gene, and in particular to assess the CAG and GGN tracts and analyze the coding sequence of the AR gene for possible disease association. Data on Kennedy's disease was of particular interest to us in that Kennedy's disease was clinically characterized by androgen insensitivity together with spinal bulbar muscular atrophy (SBMA) and has been definitively linked to the expansion of CAG repeat of exon 1 of AR gene. The expansions of the CAG repeat regions in Kennedy's disease are significant and exceeded 40 (ranges 37 to 64).

Taken together our data analysis has shown longer than 23 CAG repeats in many of the family members [7/16 (44%), A: 4, U: 2, C: 1]. However, a few [5/16 (31%), A: 3, U: 1, C: 1] of the family members showed shorter CAG repeats as well. Interestingly, only the 2nd generation unaffected male and 3rd generation

affected males showed the shorter CAG repeats and in the 4th generation all affected members had longer CAG repeat.

In this study the androgen receptor gene of unaffected members of the family was also analyzed and we found that in two of the unaffected members had increased length of the CAG repeats Furthermore, the CAG repeat length was increased in the single female with breast cancer.

A few of the studies from the American population has demonstrated a higher association of disease risk with longer CAG repeats in individuals having a positive family history of breast cancer (Haiman *et al*., 2002; Wang *et al*., 2005b). Similarly, in a study from Québec a relationship between the long CAG repeat and breast cancer (Giguere *et al*., 2001) was also demonstrated. However two studies on Australian women (Spurdle *et al*., 1999; Spurdle *et al*., 2005) and two on Israeli women (Kadouri *et al*., 2001; Dagan *et al*., 2002) reported no association between the repeat length and disease risk. In contrast, a few studies have reported the inverse correlation (De Abreu *et al*., 2007; Tsezou *et al*., 2008). All the findings above strongly suggest that CAG repeat length exhibits association with breast cancer in a population/region-specific manner.

Unfortunately in this study in the affected male family members only minor CAG extensions were found, and we found no significant expansion (>40) of the CAG repeat length. The point to be observed in this family is that the majority of the family members (including affected, unaffected and controls) were found to have variation, either longer or shorter CAG repeats. Most importantly however was

that none of the affected members were found to have the normal CAG repeat length.

These results therefore show an indeterminate association of CAG repeats with MAIS and BC in this family. However since the entire 2nd and 3rd generation males were found to have reduced CAG repeat lengths we postulate that there could be generation specific modulation of AR. One study conducted on nurses (Haiman *et al*., 2002) found no association between the long AR alleles with increased breast cancer risk, but when the study was conducted specifically on those women with a first degree family history of breast cancer, an increased risk of disease was observed (Nicolas Diaz-Chico *et al*., 2007; Sasaki *et al*., 2003).

An inverse- correlation between the CAG repeat length and AR action has been demonstrated in *in vitro* studies (Chamberlain *et al*., 1994; Kazemi-Esfarjani *et al*., 1995). Furthermore, additional studies showed an inverse-correlation among the CAG repeat length and the transactivation activity of the AR (Morales *et al*., 2000; Zinn *et al*., 2005; Kamischke *et al*., 2003). There is also evidence that in patients with Kennedy's disease, an expansion of repeat length does not necessarily occur. Indeed the pedigree described by Ikezoe *et al.,* (Ikezoe *et al*., 1999) supports this observation.

Based on our observations we found an extended CAG repeat in one of the unrelated control family members and we speculate that this female may independently have an increased risk of breast cancer.

5.2 Poly-glycine tract variations in the androgen receptor gene

The AR GGN repeat length polymorphism has not been extensively studied in MAIS and BC patients. In this study shorter GGN repeats were identified in the majority of members. The distribution was [12/16 (75%), A:6, U:3, C:3] Only one [1/16 (6%), A:0, U:0, C:1] of the family members showed a L GGN repeat, and three members of the family members showed WT GGN repeats, [3/16 (19%), A:1, U:1, C:1]. Our data suggests an association of shorter GGN repeat with the disease, however, when the repeat length of all the family members were considered collectively; there were differences amongst the affected, unaffected and control members that made it very difficult to come to a definite conclusion on the role of GGN length in this family.

Several other studies have also failed to determine the relationship between GGN repeat length and disease characteristics (Edwards *et al*., 1999; Montgomery *et al*., 2001; Stanford *et al*., 1997; Platz *et al*., 1998; Kristiansen *et al*., 2002). The significant association of GGN repeat length and breast cancer in women before the age of 45 was supported by only a few studies (Dunning *et al*., 1999; Kadouri *et al*., 2001; Suter *et al*., 2003)

Moreover, earlier studies, on the distribution of GGN repeat length, among infertile men found no difference between infertile men and the general population (Lundin *et al*., 2006; Ferlin *et al*., 2007; Tut *et al*., 1997). The

functional consequences of variations in GGN repeat are therefore less clear. One of the studies reported that deletion of the poly-glycine tract could reduce AR transcriptional activity in transient transfections assay (Gao *et al*., 1996). In addition investigation of an association between GGN repeats and other diseases, specifically prostate cancer, have also produced inconsistent results. Although it seems that short GGN repeat length increases the risk of the disease, no clear conclusions were drawn from any of the previous studies either (Irvine *et al*., 1995; Hakimi *et al*., 1997; Chang *et al*., 2002; Cheng *et al*., 2002; Hsing *et al*., 2000; Edwards *et al*., 1999). One population based study from China had observed the combined effect of poly glutamine and poly glycine effects on prostate cancer and concluded that a shorter CAG together with a shorter GGN repeat length are associated with prostate cancer (Ann W Hsing *et al*., 2000). Shorter poly-G and poly Q repeats result in higher receptor activity of AR that inhibits the growth of the steroid related tissues; in contrast, longer poly Q and poly-G repeat lengths induce decreases in the trans-activation function in the AR receptor and are highly associated with cancers such as breast (Yu *et al*., 2000; Suter *et al*., 2003), endometrial (Young *et al*., 2000; Rodriguez *et al*., 2006), and ovarian cancers (Schildkraut *et al*., 2007; Spurdle *et al*., 2000). When we analyzed the results in generation specific manner, we still reached same conclusion of the indetermination association of GGN repeat with the disease.

5.3 Point Mutations in Androgen Receptor Gene

More than 800 mutations of AIS have been described in the androgen receptor database (http://androgendb.mcgill.ca) and the increasing number of new

mutations in different clinical categories of patients, and their importance, is an area of intense research (Gottlieb *et al*., 2012). In this study we have identified a total of 43 point mutations of which 30 were novel and 13 were previously reported. Four mutations were identified in most conserved AF-1 region of NTD in AR gene. The distribution of all these mutations were variable among the 16 members of the family and no specific pattern of inheritance of mutation, or familial clustering of mutations were observed; except for the existence of three common mutations (described in detail in the next paragraph).

On review of the literature and correlating all the 13 recurrent mutations of the present study with earlier reports we found that all of them were clinically relevant and had been shown to be associated with disease i.e. CAIS/PAIS/DMD/prostate cancer.

Among the previously reported mutations, the Q28X (X=Stop codon) was detected in two girls with Duchene muscular dystrophy (DMD) (Katayama *et al*., 2006). In this present study we identified the same mutation in one unaffected male 'b'.

A T82A deletion was previously reported in a CAIS female (Philibert *et al.*, 2010). In this study we identified the mutation at the same position 82 in control female ‘D’ where Threonine was substituted to Arginine.

Similarly the E155X mutation was found previously in two CAIS females (Copelli *et al*., 1999). In this study we found a mutation at the same position 155 but E was converted into G (E155G) in two members of the family; one was

MAIS affected male 'C' and the other was control male 'a'. A previous study identified S221fs (frame shift), a single nt deletion leading to a stop at codon 226 in one CAIS female (Audi *et al*. 2010). In this present study, we found a mutation in the MAIS affected male 'K' in which S (serine) is substituted to F (phenylalanine) at the same position 221. Additionally a Y225X mutation was found in a CAIS female in an earlier study (Hannema *et al*., 2004). We similarly found a mutation at the same position in an unaffected female 'c' but with Y being substituted with C (Y225C).

Upon conducting a more thorough investigation we also found similar mutations to be associated with disease state. Almost all of the previously identified mutations were correlated with, and have been demonstrated to play a significant role in disease development. For example, E289X was found in one previous study in a CAIS female (Holterhus *et al*., 2003). In this study a mutation was found at the same position E289G in affected male ‘C.

Similarly the P380R mutation was detected in one PAIS male and one CAIS female, where a 11bp duplication was the result of a mutation at the position P380fs (Audi *et al*., 2010, Philibert *et al*., 2010). In the present study, a mutation was identified at the same position; however, the amino acid 'P' was substituted with 'L' at position P380L in the one control female ‘D’. Despite the described importance of the mutation at this position we did not identify an associated 11 bp duplication. Given that the amino acid change was not the same as described in the literature; it is conceivable that such a mutant could be a naturally occurring variant.

T440P was previously reported in one male prostate cancer patient (Steinkamp *et al*., 2009). In this present study a T440A was found in one unaffected female ‘F, once again indicating the same position for the mutation. Interestingly a E442X mutation was detected in two previous studies where the mutation was reported in CAIS females (Luo *et al*., 2011, Sun *et al*., 2010). In contrast, in the present study a E442K mutation was identified in one unaffected male ‘b’.

Along similar lines a Q445R was reported in a previous study of one CAIS female (Gottlieb *et al*., 1999b). We found the same mutation in the MAIS affected male, ‘C’. G456S mutation was identified in another earlier study in two males, both of which had prostate cancer (Steinkamp *et al*., 2009). In this present study we identified G456D in the one MAIS affected male ‘E’. Importantly, another mutation, Y481X, was found in three females of which two were CAIS and the third who developed liver cancer (Gottlieb *et al*., 1999b; Hannema *et al*., 2004; Yeh *et al*., 2007). Similarly, we identified a mutation at position Y481C in one of the unrelated female control ‘D’. A Q488 deletion was reported previously in three CAIS siblings (Boehmer *et al*., 2001). In the present study we detected Q488R in a single control male ‘a’. Another somatic mutation G525S/D was described in two prostate cancer males (Hyytinen *et al*., 2002), and we similarly detected the G525D in unaffected female ‘c’. Since all these mutations were described to be disease associated we cannot negate the importance of the mutations in this family.

Since 2004 the number of reported mutations AR gene have been steadily increasing (Gottlieb *et al*., 2012). Some studies have proposed the role of a

'genetic background' or 'modifiers', in varying manifestations of AIS (Yong *et al*., 2003, Gottlieb *et al*., 2005). The mutations/polymorphisms in the genes involved in androgen action might affect the overall response to androgens (Gottlieb *et al*., 2005; Holterhus *et al*., 1997; Gottlieb *et al*., 2001). The exposure of the foetus to androgens at the embryonic stage has also been suggested to account for the variations (Houlston and Peto, 2004). Mutations in the AR gene are known to show phenotypic variation in different individuals. Substantial variations have been noted in familial cases bearing the same mutation.

Somatic mutations in the androgen target tissues and androgen responsive organs have been proposed to contribute to phenotypic variation in some cases (Gottlieb *et al*., 2005; Holterhus *et al*., 1997; Gottlieb *et al*., 2001). However, somatic mutations have only been able to explain the phenomenon in a few of the more than 25 cases of phenotypic variation reported to date (Gottlieb *et al*., 2001). In one of the study by Holterhus *et al.*, they revealed the presence of a single mutation in the AR gene in 4 members of the same family, but with different phenotypes (Holterhus *et al*., 2000).

In the present study, we identified a total of 30 novel mutations, in which single amino acid substitutions were seen. Interestingly, four of the mutations were seen in highly conserved residues (aa 224-258) of AF-1 domain of AR lies in exon 1. The AF1 domain is of immense importance due to the highly conserved hydrophobic nature of the amino acids, and shows complete homology from fish to primates, and shows 60-70% reduced functional activity if mutated (Betney and McEwan, 2003).

Among the four mutations, one each of the three mutations was present in three affected members and one was found in a genetically related unaffected member. They were located in this highly conserved AF-1 domain. Importantly, the three mutations identified in affected members were novel and of the three affected members, two developed MAIS and the third, a female developed breast cancer. Analysis revealed mutations at the position T231I and C240R in two affected males ('E' and 'I') with MAIS and gynecomastia and a single mutation at position V245M were identified in the female ('d') with breast cancer, and a single unaffected female 'c' showed mutation at the position Y225C, which was previously reported in CAIS female (Hannema *et al*., 2004). Since, such mutations are of high relevance, they all warrant future functional/structural studies. All the mutations and their details are given in the Figure 25.

The number of mutations in MAIS has been increased since 2004 from 7 to 22, and these mutations were due to nucleotide substitution. The reason for a significant number of missense mutations between aa residue 214 to 511 in the MAIS affected males remains a mystery; however, it has been suggested by Gottlieb *et al.*, that the missense mutations obtained in this region of NTD have a minor effect on the function of AR (Gottlieb *et al*., 2012). In the present study we identified 21, (50%) missense mutations between the aa residues 214 to 511 of AR transactivation domain, 12/21 mutations were novel and 9/21 were previous described. The majority (11) of the mutations were detected in affected members of the family at position T231I, L345P, C240R, Q360R, R386C, L347F, V245M and are novel. The remaining four mutations are recurrent mutations occurring at

position E289G, Q445R, G456D and S221F. The mutations identified in this region provide a strong basis for the MAIS with gynecomastia development in males and development of breast cancer in females. Among unaffected genetically linked members we found six mutations spanning this region. The three detected at position L389P, G326D, P422S were novel and those at position T440A, E442K and Y225C were previously reported. , surprisingly, we found four mutations at the position A372T, S361G which are novel, and Y481C, Q488R are reported mutations in unrelated members, who were also the sexual partners of the affected members and were used as controls, however, no genetic link can be established, and investigations are required to know their familial history and background.

Only 2 novel mutations were identified at the position W527R and M528V in the BC affected female. The remaining 19 novel mutations were found between the aa 25 to 213, and among affected members, we found 8 at the position L108P, E113K, Q78R, E187G, T105I, E32V, N25S, V132A. Among unaffected members, we found 5 mutations at the position P153L, Q72R, Q78R, M191T, G173S and the controls also showed 5 mutations at position T82A, Q79R, V33A, Q98R and P49L. Mutations spanning from 214 to 511 amino acids in the transactivation domain of AR in MAIS patients remain a mystery and is an area of intense future research, but at present, it does suggest the importance of exon 1 alone in asserting the overall modulation effect of the AR (Gottlieb *et al*., 2012).

In summary, we have identified several novel and recurrent mutations in the NTD of androgen receptor gene in the affected members of the family. Detailed

functional studies are now necessary to verify the mechanism and the pathogenicity of the novel mutations and are the basis of intense future research endeavors.

An important feature to arise from the genetic study of this family is the apparent genetic heterogeneity observed in AR exon 1. A study in 1997 was one of the first to first to appreciate the role of somatic mutation in the AR in causing PAIS (Holterhus *et al.,* 1997). In this study the authors found one of the previously reported mutations L172X abrogated AR expression and was the main reason for the observed androgen insensitivity. However in another study by the same group they identified the same mutation but patients this time presented with the PAIS phenotype. Functional assay of androgen binding further revealed that cells of PAIS patients has measurable androgen binding activity, despite the presence of stop codon mutation L172X. While looking for the cause of such phenomenon it was later discovered that the mutation was inherited from the patient's mother, however the same mutation had been acquired in the form of a mixed/heterogenous population (Holterhus *et al.,* 1999). In a study utilizing PBMCs and GSF the authors found that phenotypic variability caused by somatic mosaicism could also occur as a result of *denovo* post-zygotic forward mutations in AR gene (Hirot *et al*, 1998). Currently only five cases of variable AIS phenotypes are confirmed to be caused by somatic mosaicism (Holterhus *et al.,* 1999).

The somatic mosaicism of mutant and wild-type AR allele is a prime molecular factor of *in vivo* androgen action and the associated somatic mosaicism identified

in the present study could be the main cause of MAIS in members of the present family.

Due to somatic mutations variable expression of the AR has been observed in prostate cancer and is a primary source of variable expressivity. Somatic changes are common in cancer, but recent data indicates that many other diseases could also be caused by somatic mutations (Frank and Nowak, 2004). Recent literature has highlighted the role of somatic mutations in primary immune deficiency, neurofibromatosis, secondary hypertension and prostate cancer (Beuschlein *et al.*, 2013; Tsujimoto *et al.*, 2004 and Ackerman, 1998). Results in many of these studies have been driven by next generation sequencing. Therefore to investigate AR mutations in this family further and to determine the role played by somatic mosaicism will require more in depth investigation of the phenotypic-genotypic relationship *in vivo*/ *in vitro*.

Androgen receptor and breast cancer

Among all the family members right arm including affected and presently unaffected members revealed longest length of CAG repeat. Long poly Q stretch was 25 in numbers were identified in breast cancer affected female (d) and in addition longer poly Q were identified in her sibling (b,c), indicating possible chances of these subject to develop breast cancer and MAIS with gynecomytia in future. Similarly longer GGN repeat length were identified in female affected with breast cancer and both of her sibling. Four novel mutations were identified in the breast cancer affected female at the positions V132A (GTC→GCC), V245M (GTG→ATG), W527R (TGG→CGG) and M528V (ATG→GTG); however none

of the mutations were previously reported. Among all the risk factors family history is considered as top most risk factors in breast cancer patients, those members showing increased length of CAG repeat have higher chances of developing breast cancer. A recent review by Rajender *et al.*, 2011 described over 10 studies which found association of longer CAG repeat with breast cancer, however there are some studies which could not found such and 2 studies which found inverse correlation of CAG repeat length with breast cancer. In addition mutations in the Androgen receptor could compliment increased CAG repeat affect and may exacerbate the breast cancer development, this could possibly be due to reduced androgen receptor transactivation activity lowering androgen: estrogen balance and promoting pathogenesis

6. CONCLUSIONS

The family included in the study is unique and interesting as there is segregation of two diseases in two separate arms of the family. MAIS/gynecomastia affected males were present in one arm of the family and females affected with breast cancer were present in another arm of the family. Clinical features and symptoms of affected males typically showed gynecomastia in all affected male members of the family.

We decided to analyze the whole androgen receptor gene because it is the leading source of AIS. On the basis of clinical symptoms, we grouped the family into three groups: affected, unaffected and control members. We analyzed all the exons from 1-8 of androgen receptor gene and found substitution mutations throughout the exon 1 of AR gene in all three groups of family members. We extended our study and analyzed the remaining exons 2 to 8, however, all the

exons from 2 to 8 were found totally wild type and well conserved in all the 16 members of the family.

In exon 1, we initially focused on the highly polymorphic CAG repeat length, as an extension of CAG repetition is involved in different kinds of diseases including AIS. We were expecting the expansion of CAG repeat length in affected members of the family especially males with MAIS and gynecomastia, as gynecomastia is one of the symptoms associated with the neurodegenerative disease 'Kennedys disease' in which trinucleotide repeat of CAG extend more than 40 repeats. Surprisingly, we did not find an increased expansion of the CAG repeats in these MAIS affected males of the family.

On further analysis, we found generation specific patterns of (CAG)n length. In the second generation, we found shorter CAG repeat length in a single unaffected male. Also in the third generation all affected males showed shorter CAG repeat length. This led us to assume that shorter CAG repeat length had been transferred from second generation. However when we analyzed the fourth generation members we were surprised to discover longer CAG repeat lengths in all affected members of the fourth generation. Interestingly, unaffected males and females of the fourth generation also showed longer CAG repeat. Based on these observations we conclude an indeterminate association of CAG repeat length with breast cancer and MAIS with gynecomastia within affected members of this family. However there does appear to be some generational modulation of the AR gene with respect to the CAG repeat sequence.

Secondly, we focused on another polymorphic region of the AR gene i.e., GGN repeat length which varies from 10 to 30 in many population groups. In the second, third and fourth generation members, we found shorter GGN repeats except one control female in which we found longer a GGN repeat. Since all three groups of the family (viz. the affected, unaffected and controls) had shorter GGN repeat lengths we also conclude an indeterminate association of GGN repeat length with breast cancer and MAIS in this family.

The final focus was that of the point mutations which were identified throughout exon 1 of AR gene in all the members of the family except one control female who is completely wild type. In total 43 missense mutation were identified in the trans-activation domain of AR gene: 30 unreported (novel) mutations as well as 13 recurrent (previously reported) mutations. Seventeen novel and 5 recurrent mutations were identified in the affected group; 8 novel and 5 reported mutations, including one premature stop codon mutation were identified in the unaffected normal group and 7 novel and 4 reported mutations were found in the control group. Of the above-mentioned mutations, three novel mutations and one recurrent mutation were identified in the activation function-1 (AF-1) domain of exon 1 in 3 affected members (2 males and one female) and in one genetically related unaffected female member of the family. Although these 3 mutations of the AF-1 domain have not been previously identified and reported in patients, previous *in-vitro* studies have carried out using mutations specific for this regions and it was concluded that mutations of this AF-1 region could reduce the AR transactivation activity by 60-70%.

The reason for a significant number of missense mutations between aa 214 to 511 causing MAIS remains a mystery it has been suggested that missense mutations in this part of exon 1 have a slight effect on overall AR function (Gottlieb *et al*., 2012). In this study family we identified 21, (50%) missense mutations between the aa residues 214 to 511 of AR transactivation domain. The majority (11) mutations were found in affected members of the family and we conclude that the mutations identified in this region could affect the AR function and might be a strong cause for the diseases observed in this family. Since mutations were also detected in unaffected genetically linked members in this region, it is reasonable to speculate that these unaffected family members may be afflicted with an AR related disease in the future.

Regarding the identification of mutations in the single breast cancer affected female in the breast cancer arm, we conclude that changes in this regions may lead to altered modulation of transactivation function of androgen receptor,

changing its overall activity and thereby playing a role in breast cancer development.

6.1 FUTURE DIRECTION

The 43 different mutations spread along the entire transactivation domain of the exon 1 of AR gene are the defining feature of this family. The point mutations presented needs to be explored for their functional relevance and further studies are required to confirm if such mutations have an effect on the overall transactivation function and the resultant signaling mechanism of androgen receptor thereby changing the pathogenicity and severity of the disease.

Mutations identified in AF-1 domain of exon 1 of androgen receptor seems to be of major concern, followed by those identified within the aa 214-511 region, mainly in affected males with MAIS/gynecomastia and the female with breast cancer. Investigations are required to understand the population as well as familial history and background of the unrelated control members of the family which also showed few mutations in the exon 1 of AR.

Appendix A

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Ms Samatha Chauhan

CLEARANCE CERTIFICATE	**M121194**
PROJECT	Genetic Analysis of Mild Androgen Insensitivity Syndrome (MAIS) and Breast Cancer in a South African Indian Family
INVESTIGATORS	Ms Samatha Chauhan.
DEPARTMENT	Internal Medicine
DATE CONSIDERED	30/11/2012
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 30/11/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr Raquel Duarte

DECLARATION OF INVESTIGATOR(S)
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

Appendix B

Sucrose Triton-X Lysis buffer

10 ml 1 M Tris HCl pH 8

5 ml 1 M MgCl2

10 ml Triton X-100

Makeup to 1000 ml with distilled water. Autoclave and store at 4 ˚C.

Add 109.5 g sucrose just before use. Make fresh buffer for each use.

T20 E5 (pH 8)

20 mM Tris HCl

5 mM EDTA

Autoclave and store at room temperature.

Saturated NaCl

40 g NaCl to 100 ml sterile water until completely saturated.

Store at room temperature.

Proteinase K mix

10% SDS

0.5 M EDTA

10 mg /ml proteinase K: add 50 ml sterile water to bottle containing 500 mg. Aliquot and freeze.

TBE (10X Stock solution)

108 g Tris base

55 g Boric acid

40 ml 0.5 M EDTA pH 8

Makeup to 1 litre with distilled water.

PBS (Phosphate buffered Saline)

8 g NaCl

0.2 g KCl

1.44 g Na2 HPO4

0.24 g KH2 PO4

Dissolve in 800 ml of distilled water. Adjust pH to 7.4 Make up to 1000 ml, sterilize by autoclaving and store at room temperature.

10X TBE buffer

108 g Tris base

55 g Boric acid

9.3 g EDTA

Adjust volume to 1000 ml with distilled water. Sterilize by autoclaving and store at room temperature.

1X TBE buffer

100 ml 10 X TBE buffer

900 ml distilled water

Mix thoroughly.

Agarose gel, 0.7% or 1%

0.7 g (0.7%) or 1 g (1%) agarose

100 ml 1X TBE buffer

Dissolve agarose in 1X TBE buffer by heating in a microwave oven for 3 minutes, shake occasionally, until all the agarose particles are completely dissolved. Care must be taken that the high percentage gels do not boil over. Cool to 50 ˚C. Add 6 µl mg/ml Ethidium bromide and mix it well. Pour into a prepared casting tray and allow to solidify completely at room temperature.

0.5M Ethylene diamine tetraacetic acid (EDTA) (pH 8)

18.61 g Ethylene diamine tetra acidic acid

80 ml Distilled water

Stir vigorously on a magnetic stirrer. Adjust pH to 8 with 10 M NaOH. Adjust volume to 100 ml distilled water. Sterilize by autoclaving and store at room temperature.

Luria-Bertani (LB) medium

10 g Bacto Tryptone

5 g Yeast extract

10 g NaCl

Adjust volume 1000 ml with distilled water. Sterilize by autoclaving and store at 4 ˚C.

LB agar plates with ampicillin

10 g Bacto Tryptone

5 g Yeast extract

10 g NaCl

15 g Bacteriological agar

Adjust volume to 1000 ml with distilled water. Sterilize by autoclaving. Allow the medium to cool to 50 ˚C before adding ampicillin to a final concentration of 50 mg/ml. Pour 30-35 ml of medium into 85 mm petridishes. Let the agar solidify completely at room temperature and store at 4 ˚C for up to 1 month.

LB plates with ampicillin/IPTG/X-gal

Make the LB plates with ampicillin as above, then supplement with 100 µl of 100 mM IPTG and 20 µl of 50 mg/ml X-gal pour separately, and spread over the LB ampicillin plate and allow to absorb for 30 mins at 37 ˚C prior to use.

SOC medium 100 ml (pH-7)

2 g Bacto®- Tryptone

0.5 g Bacto®-yeast extract

1 ml 1 M NaCl

0.25 ml 1 M KCl

1 ml 2 M Mg^{2+} stock, filter sterilized

1 ml 2M glucose, filter sterilized

Add; Bacto tryptone, Bacto-yeast extract, NaCl and KCl to 97 ml distilled water. Stir to dissolve autoclave and cool to room temperature. Add 2 M stock Mg^{2+} and 2 M glucose, each to a final concentration of 20 mM. Bring to 100 ml with sterile, distilled water. The final pH should be 7.

2 M Mg^{2+} Stock

20.33 g MgCl2. 6H2O

24.65 g MgSO4. 7H2O

Add distilled water to 100 ml. Filter sterilize.

Psi broth (pH 7.6)

5 g Bacto yeast extract

20 g Bacto tryptone

5 g Magnesium Sulfate

Adjust volume to 1000 ml with distilled water.

Sterilize by autoclaving.

Tfb I (pH 5.8)

0.588 g Potassium Acetate (30 mM)

2.42 g Rubidium Chloride (100 mM)

0.294 g Calcium Chloride (10 mM)

2 g Magnesium Chloride (50mM)

30 ml Glycerol (15% v/v)

Adjust the volume to 200 ml with distilled water. Filter sterilize.

Tfb II (pH 6.5)

0.21 g M OPS (10 mM)

1.1 g Calcium Chloride (75 mM)

0.121 g Rubidium Chloride (10 mM)

15 ml Glycerol (15% v/v)

Adjust the volume to 100 ml with distilled water. Filter and sterilize.

6X Loading dye (Promega)

15% Ficoll ® 400

0.03% Bromo-phenol blue

0.03% Xylene Cyanol FF

0.4% Orange G

10 mM Tris HCl (pH 7.5)

50 mM EDTA.

Table 3: Ligation reaction component

Reaction component	Standard reaction	Positive control	Background control
2X rapid ligation buffer	5 µl	5 µl	5 µl
pGEM®-T easy vectors (50 ng)	1 µl	1 µl	1 µl
PCR product	X µl*	-	-
Control insert DNA	-	2 µl	-
T4 DNA ligase (3 Weiss U/µl)	1 µl	1 µl	1 µl
Nucleus free water to a final volume of	**10 µl**	**10 µl**	**10 µl**

* Molar ratio of (3:1) PCR product (insert): vector may require optimization.

Table 4: Restriction digestion and mapping of purified plasmid DNA with Fast digest enzyme *Not I*

Reaction component		**Uncut**
10 X Fast digest buffer	1 µl	-
Fast digest enzyme *Not I*	0.5 µl	-
DNA	Up to 1 µg	2 µl
Nuclease free water to a final volume of	**10 µl**	**10 µl**

Incubate at 37 ˚C for 50 to 60 mins.

Thermal inactivation of Fast digest enzyme, *Not I* at 80 ˚C for 5 mins.

Appendix C

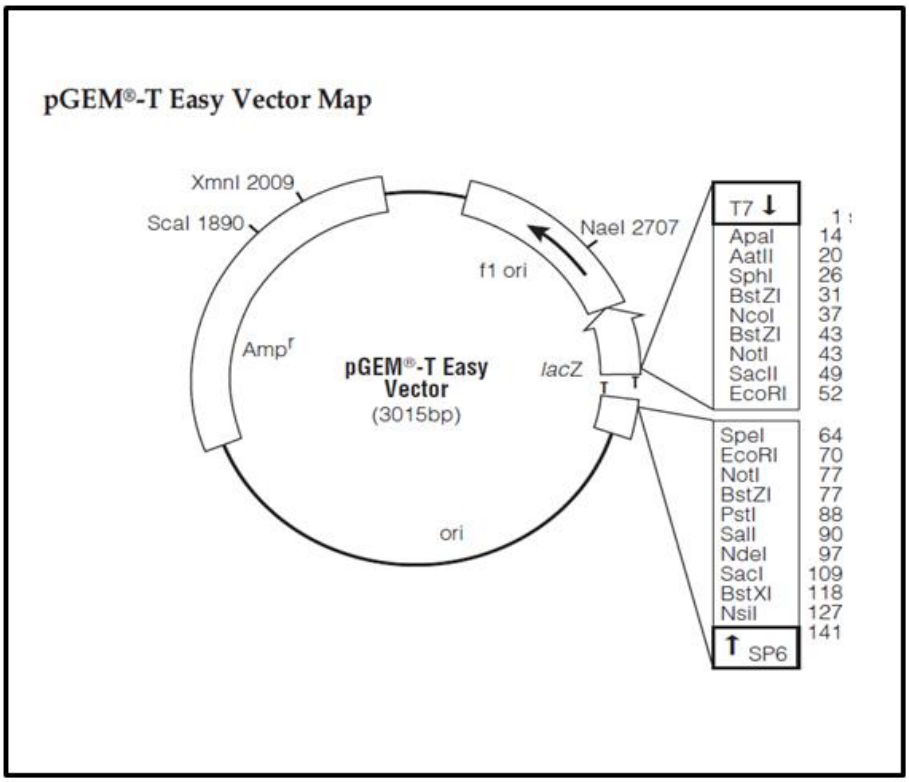


Figure 13: The p$^{GEM®}$- T Easy Vector Map: pGEM®-T easy vector 3015bp in size and it is a linearized vector with a single 3`-terminal thymidine (T) at both ends. The T- overhangs at the insertion sites greatly improve the efficiency of ligation of PCR products by preventing re-circularization of the vector and providing a compatible overhang for PCR products, generated by certain thermo stable polymerases. The pGEM®-T easy vectors are high copy number vectors containing T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α- peptide coding region of the enzyme β-galactosidase. Insertional inactivation of the α- peptide allows identification of recombinants by blue/white screening on LB agar plates with ampicillin. These vectors contained numerous restriction enzyme sites within the multiple cloning regions.

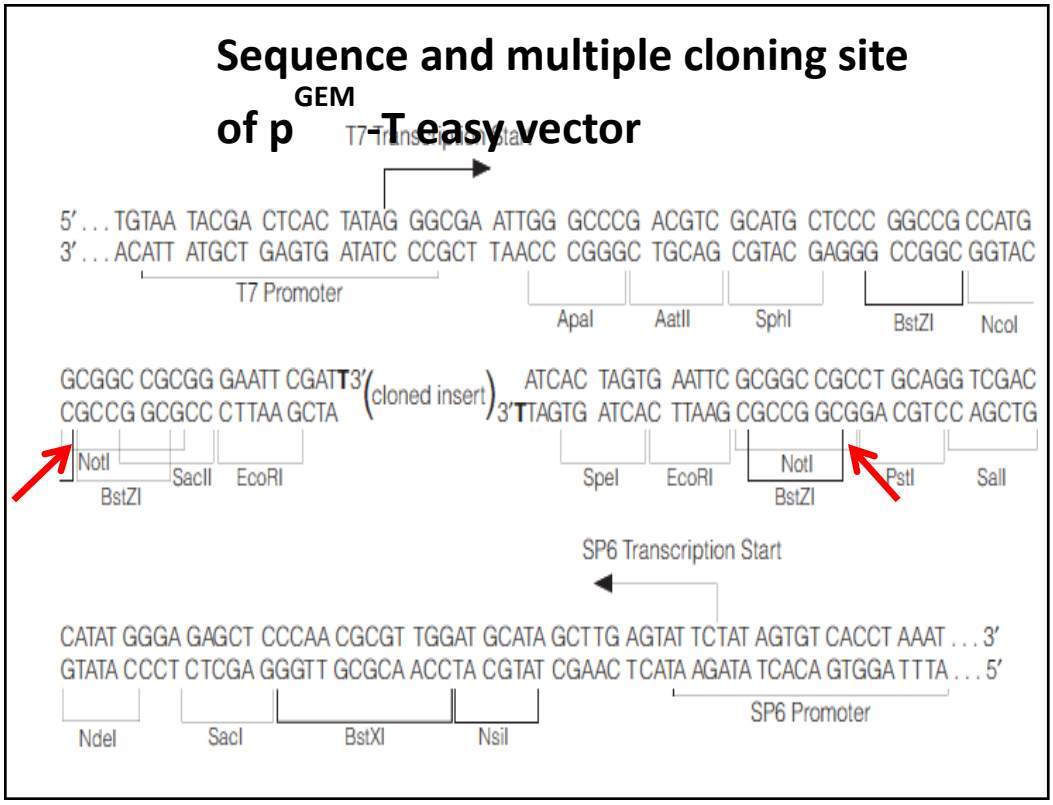


Figure 14: Sequence and multiple cloning site of the pGEM®-T easy vector. The top strand shown corresponds to the RNA synthesized by T7 RNA polymerase. The bottom strand corresponds to the RNA synthesized by SP6 polymerase.

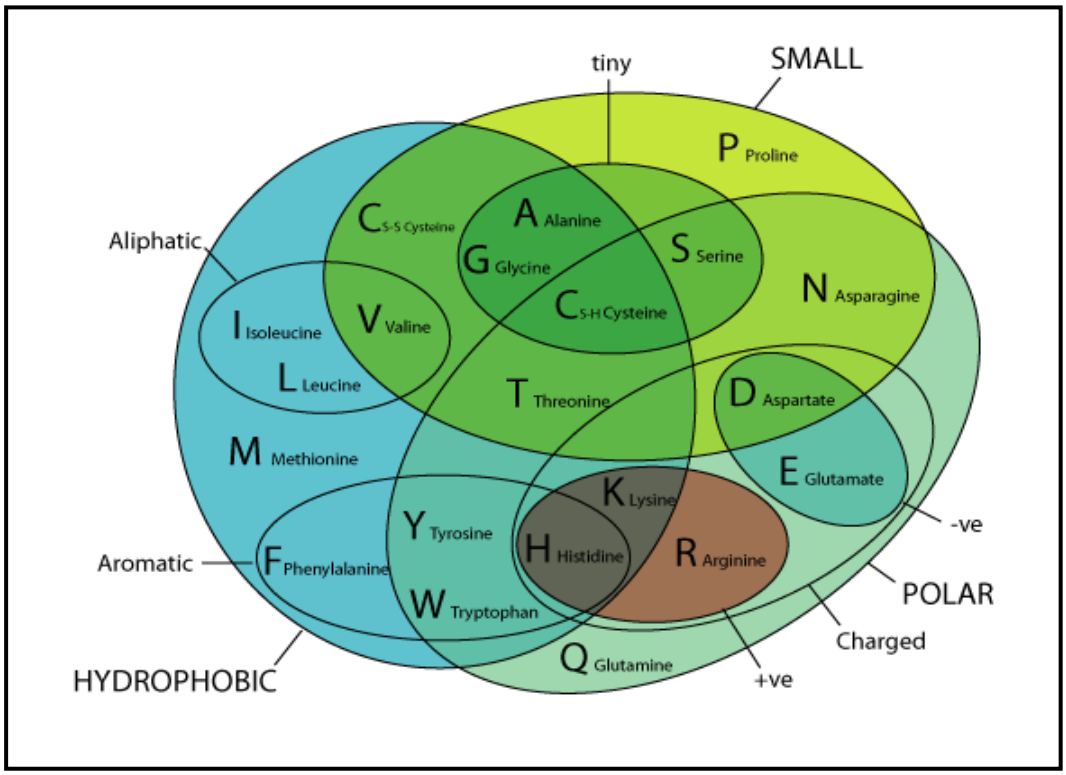


Figure 71: The chart representing the different properties of amino acid. Adopted from (http://www.geneinfinity.org/sp/sp_aaprops.html)

Appendix D

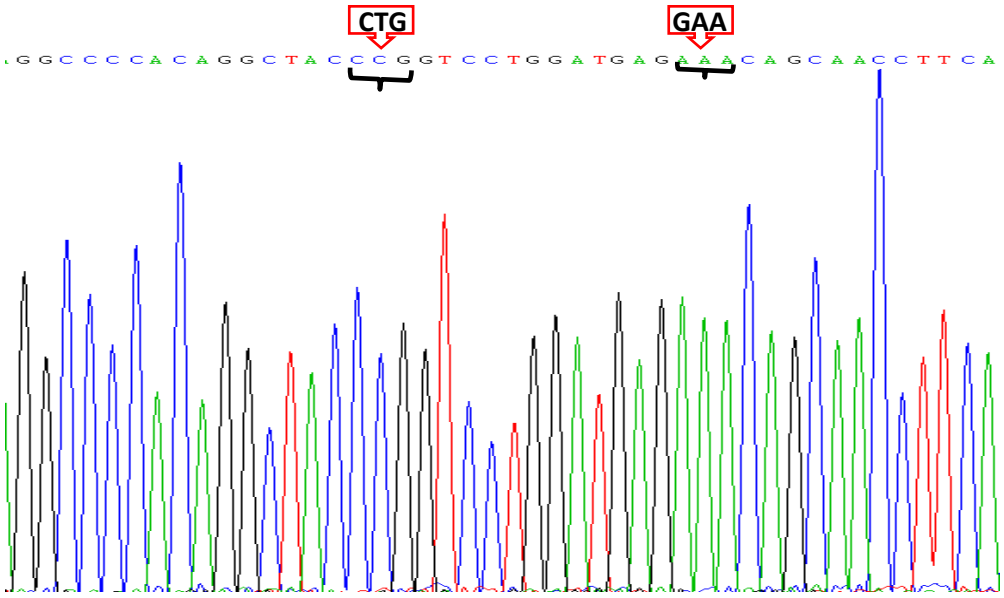


Figure 26: Electropherogram showing L 108 P and E 113 K novel mutations in exon 1 of AR gene in MAIS affected male 'C'. Where L = leucine, P = Proline and E = Glutamic acid, K = Lysine.

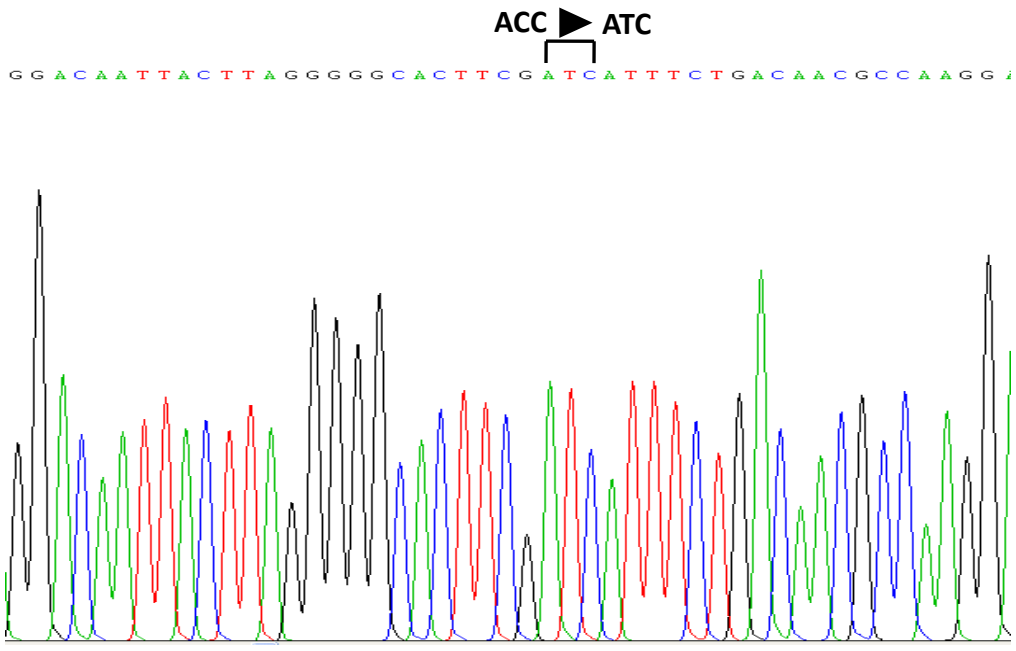


Figure 27: Electropherogram showing T 231 I novel mutation in exon 1 of AR gene in MAIS affected male 'E'. Where T = Threonine and I = Isoleucine.

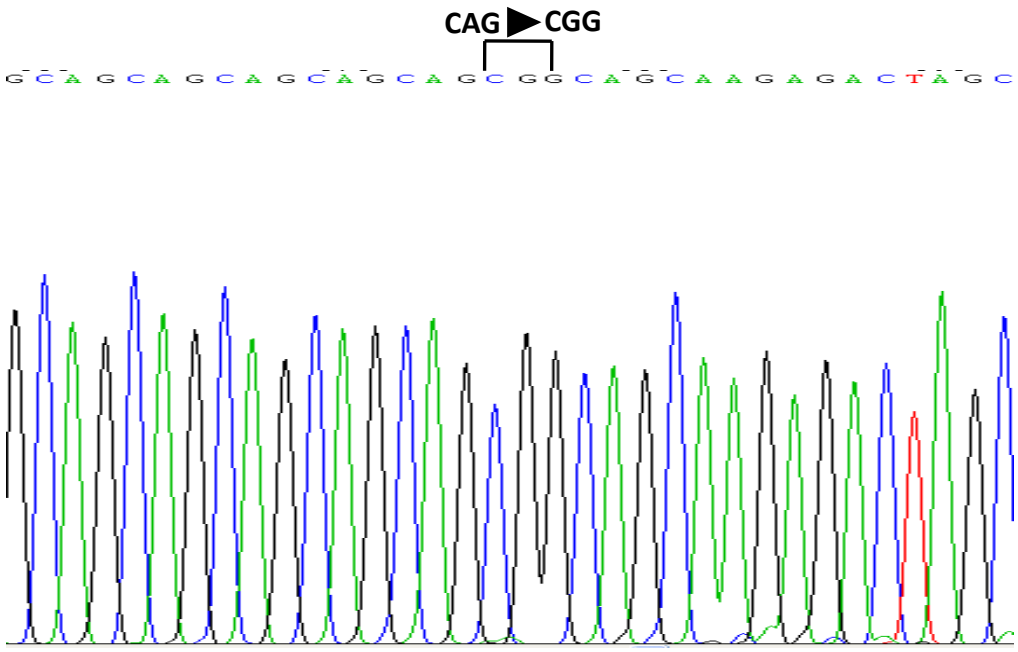


Figure 28: Electropherogram showing Q 78 R novel mutation in exon 1 of AR gene in MAIS affected male 'H'. Where Q = Glutamine and R = Arginine.

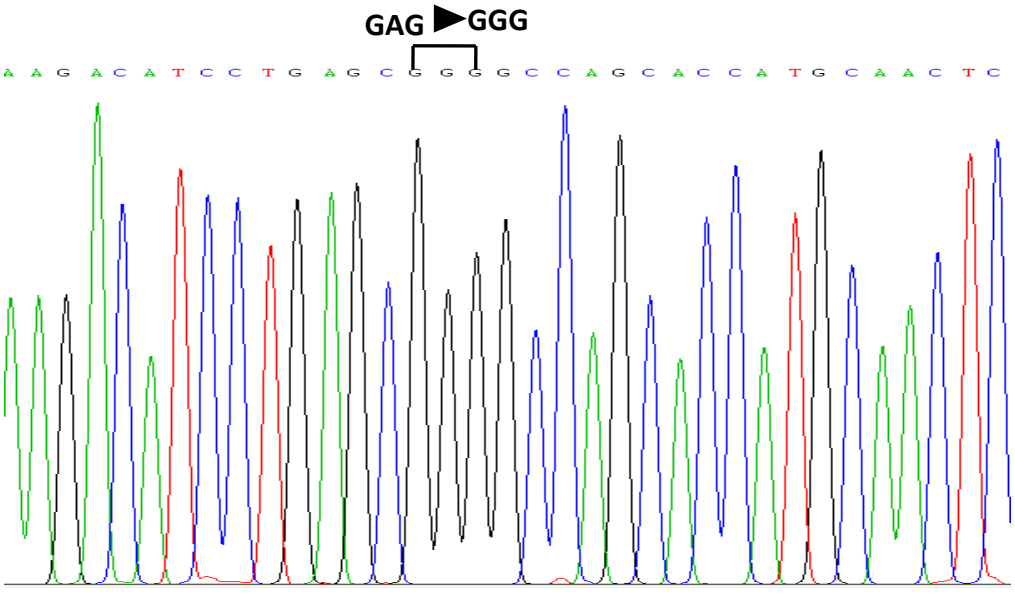


Figure 29: Electropherogram showing E 187 G novel mutation in exon 1 of AR gene in MAIS affected male 'H'. Where E = Glutamic acid and G = Glycine.

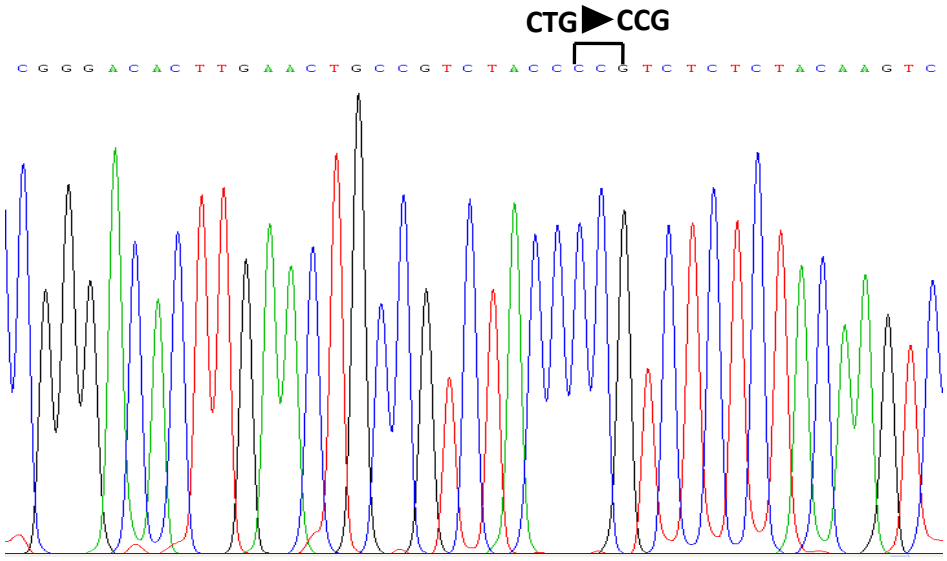


Figure 30: Electropherogram showing L 345 P novel mutation in exon 1 of AR gene in MAIS affected male 'H'. Where L = Leucine and P = Proline.

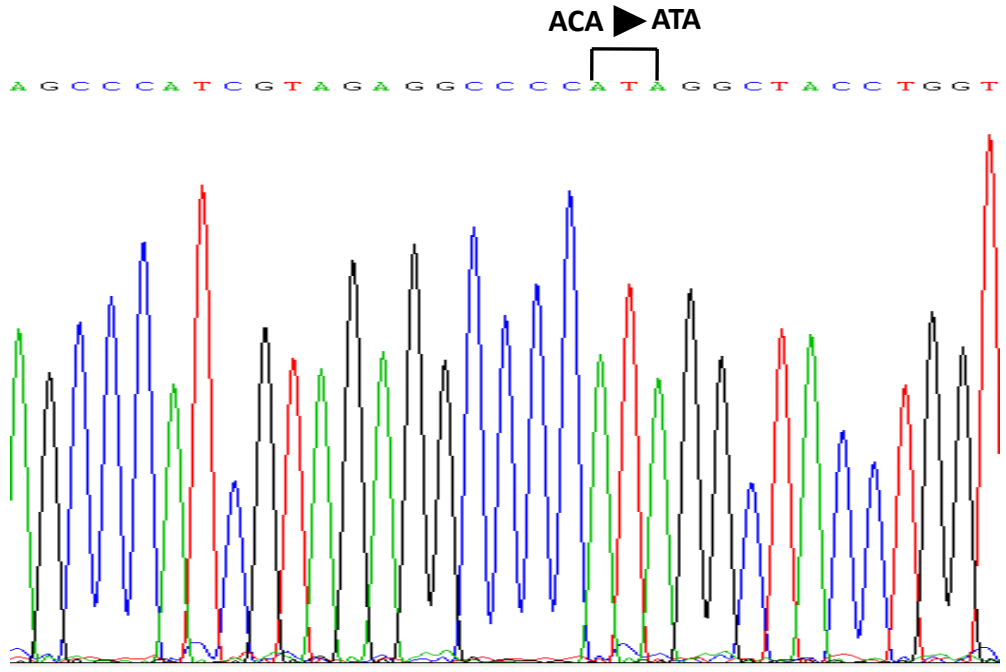


Figure 31: Electropherogram showing T 105 I novel mutation in exon 1 of AR gene in MAIS affected male 'I'. Where T = Threonine and I = Isoleucine.

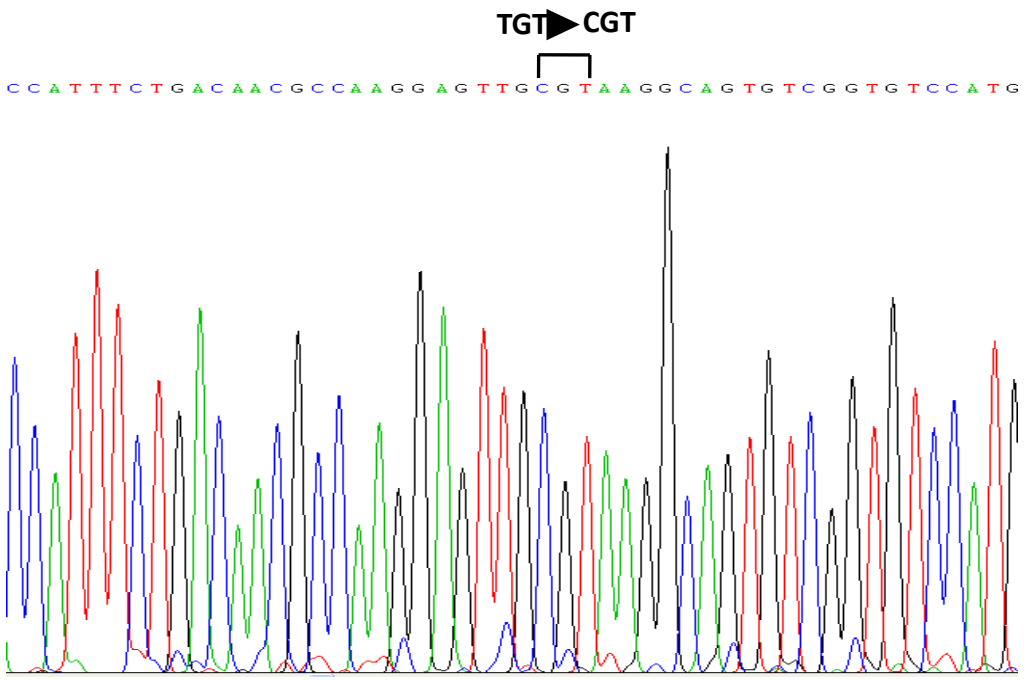


Figure 32: Electropherogram showing C 240 R novel mutation in exon 1 of AR gene in MAIS affected male 'I' Where C = Cysteine and R = Arginine.

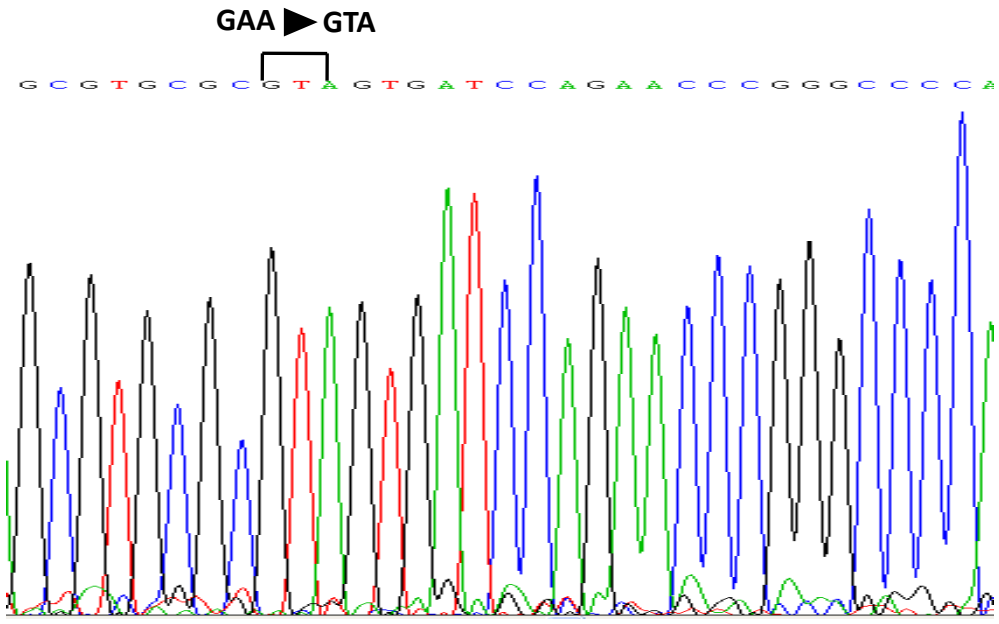


Figure 33: Electropherogram showing E 32 V novel mutation in exon 1 of AR gene in MAIS affected male 'K'. Where E = Glutamic acid and V = Valine.

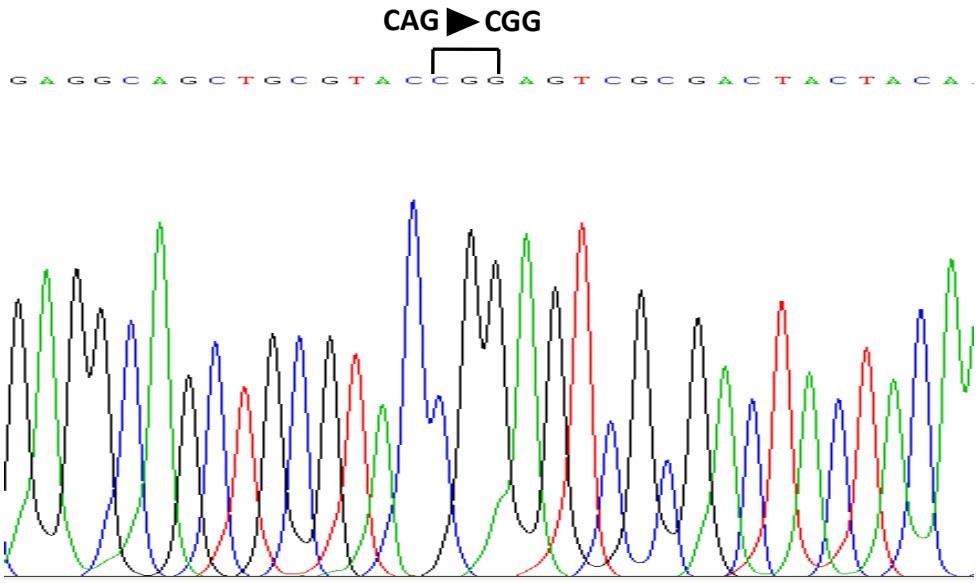


Figure 34: Electropherogram showing Q 360 R novel mutation in exon 1 of AR gene in MAIS affected male 'K' Where Q = Glutamine and R = Arginine.

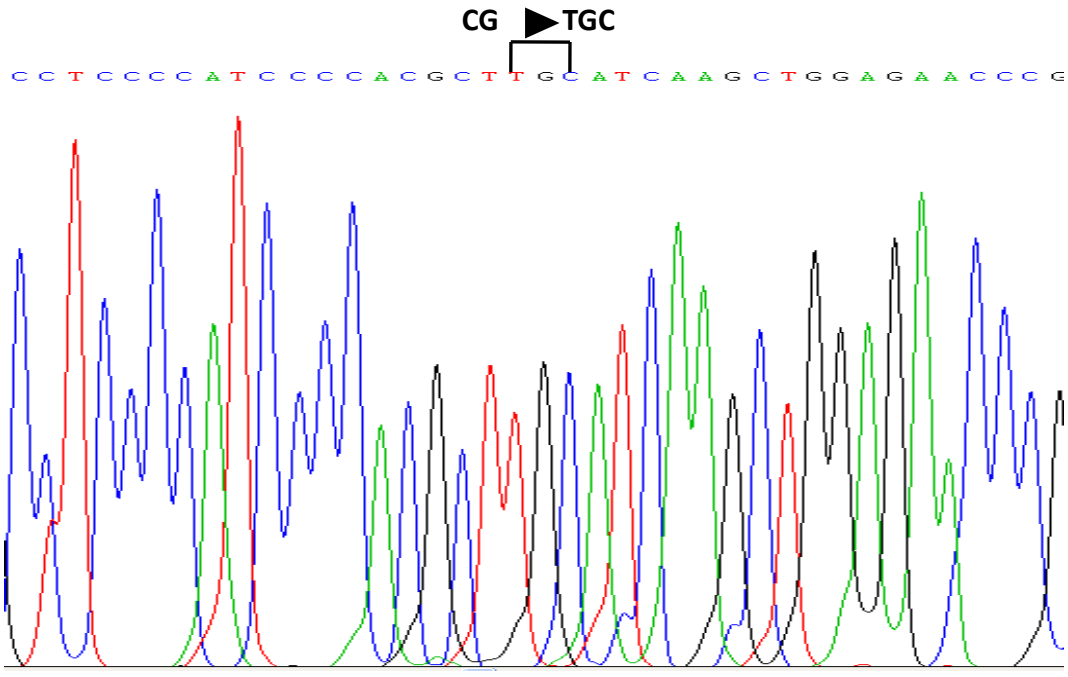


Figure 35: Electropherogram showing R 386 C novel mutation in exon 1 of AR gene in MAIS affected male 'K'. Where R = Arginine and C = Cysteine.

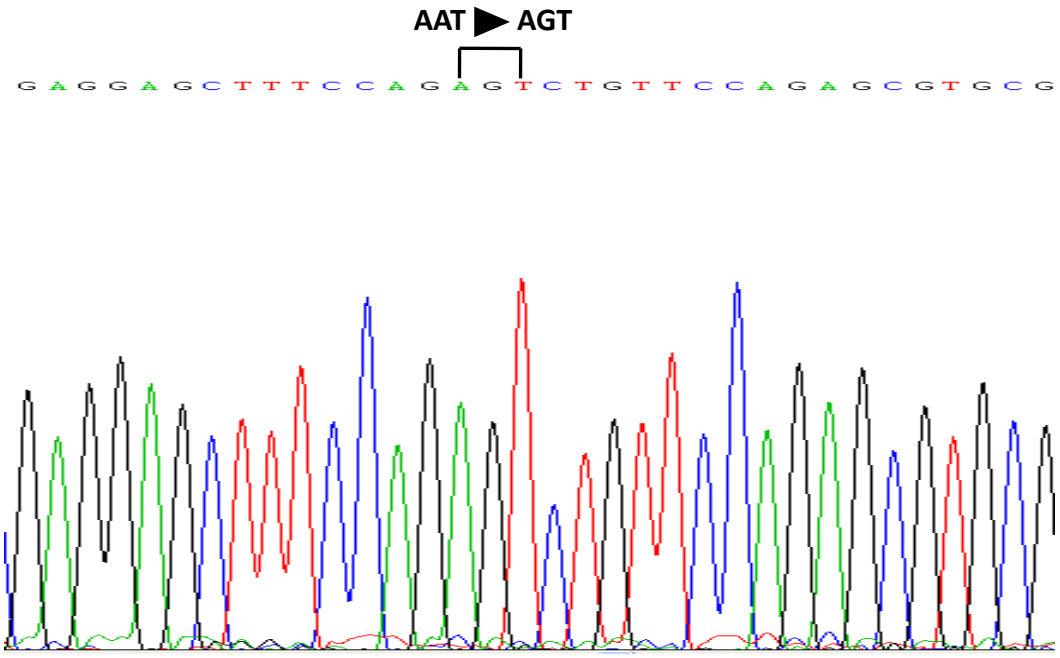


Figure 36: Electropherogram showing N 25 S novel mutation exon 1 of AR gene in MAIS affected male 'L'. Where N = Aspargine and S = Serine.

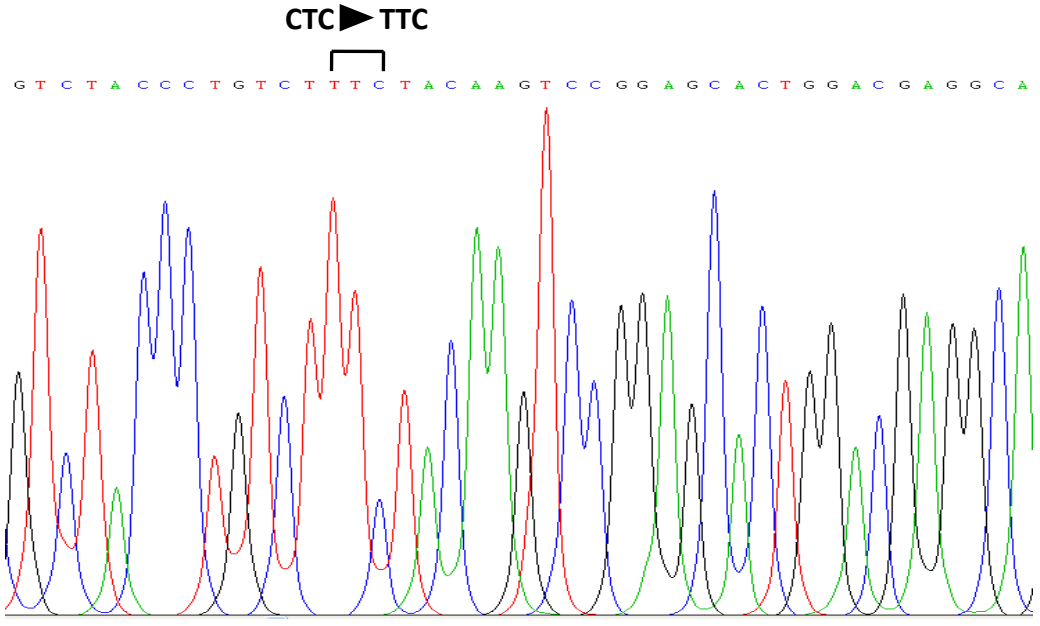


Figure 37: Electropherogram showing L 347 F novel mutation in exon 1 of AR gene in MAIS affected male 'L' Where L = Leucine and F = Phenylalanine.

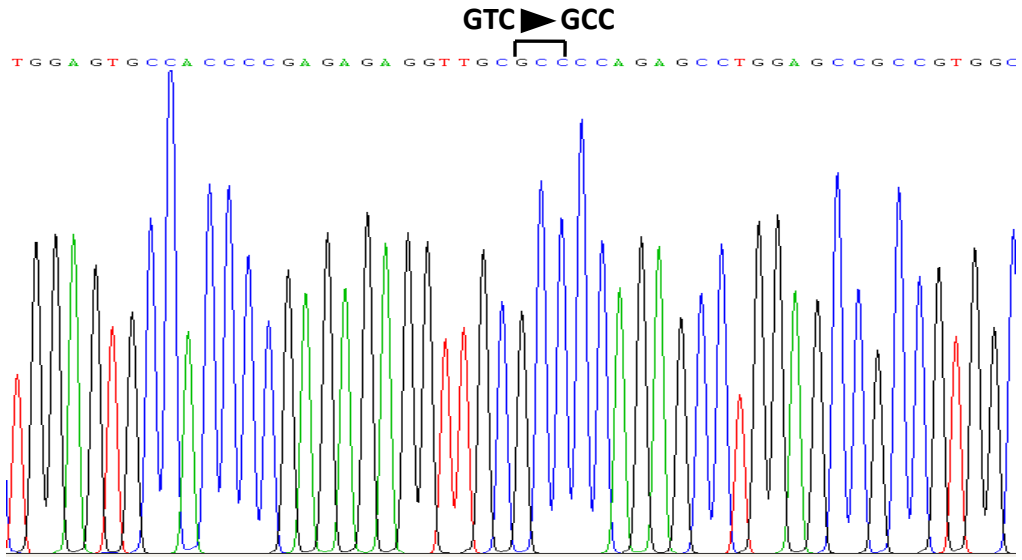


Figure 38: Electropherogram showing novel mutation V132A in exon 1 of AR gene in BC affected female 'd'. Where V = Valine and A = Alanine.

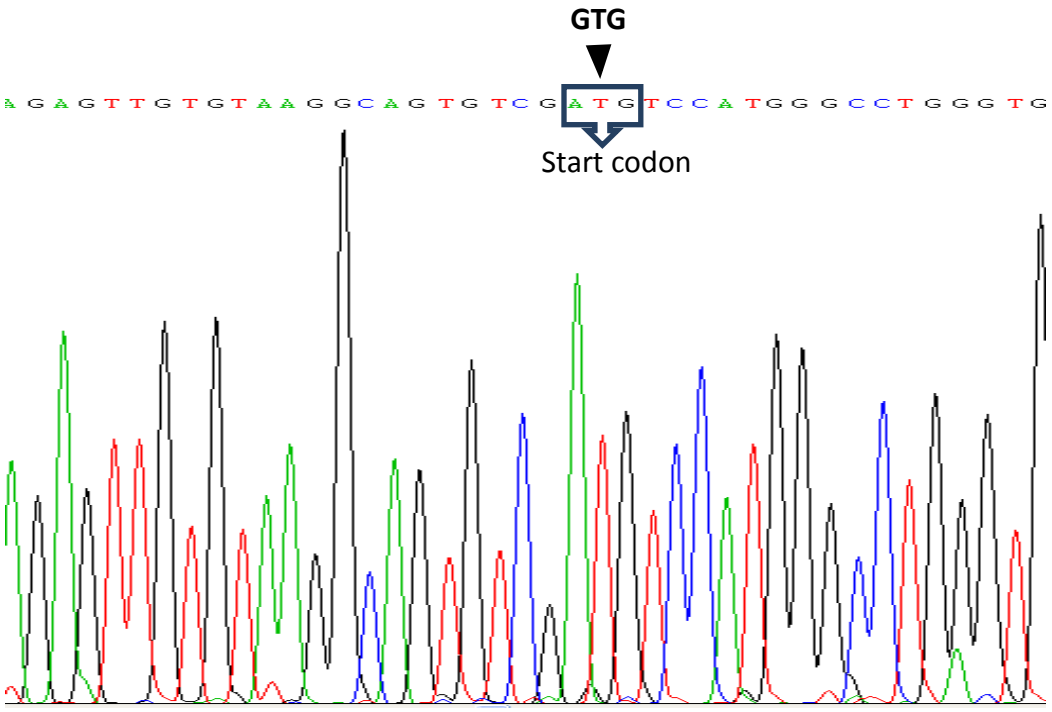


Figure 39: Electropherogram showing the novel mutation V 245 M in exon 1 of AR gene in BC affected female 'd', where V = Valine and M = Methionine.

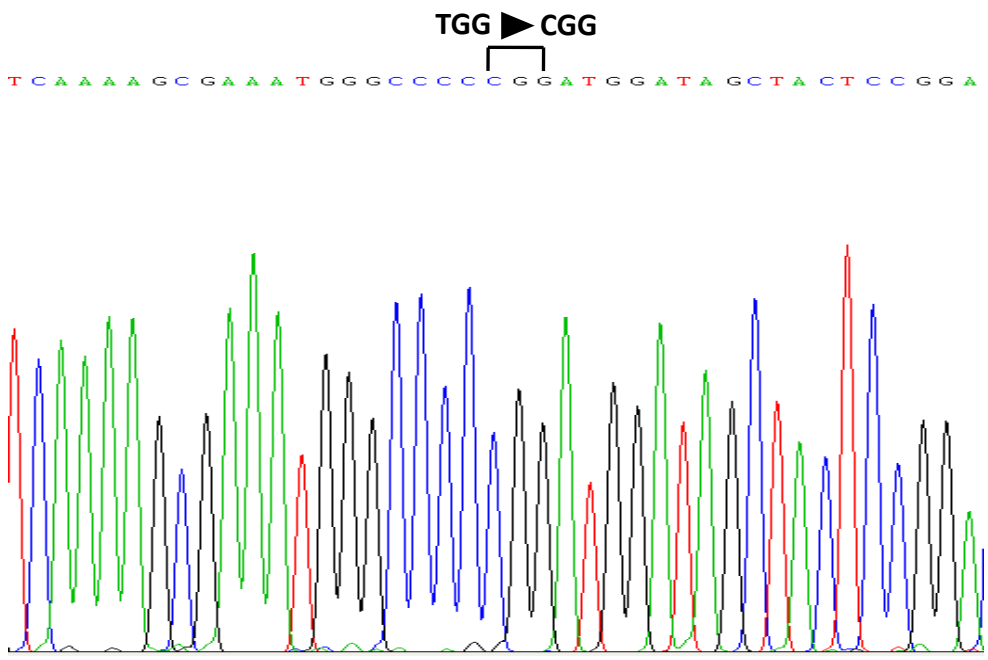


Figure 40: Electropherogram showing W 527 R novel mutation in exon 1 of AR gene in BC affected female ‘d’. Where W = Tryptone and R = Arginine.

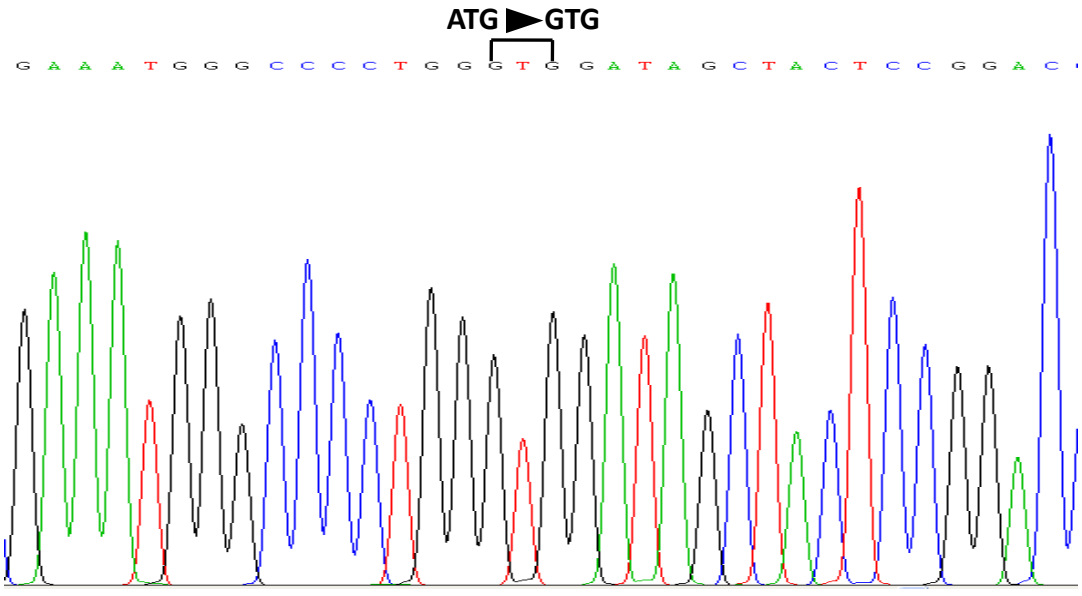


Figure 41: Electropherogram showing M 528 V novel mutation in exon 1 of AR gene in BC affected female ‘d’. Where M = Methionine and V = Valine.

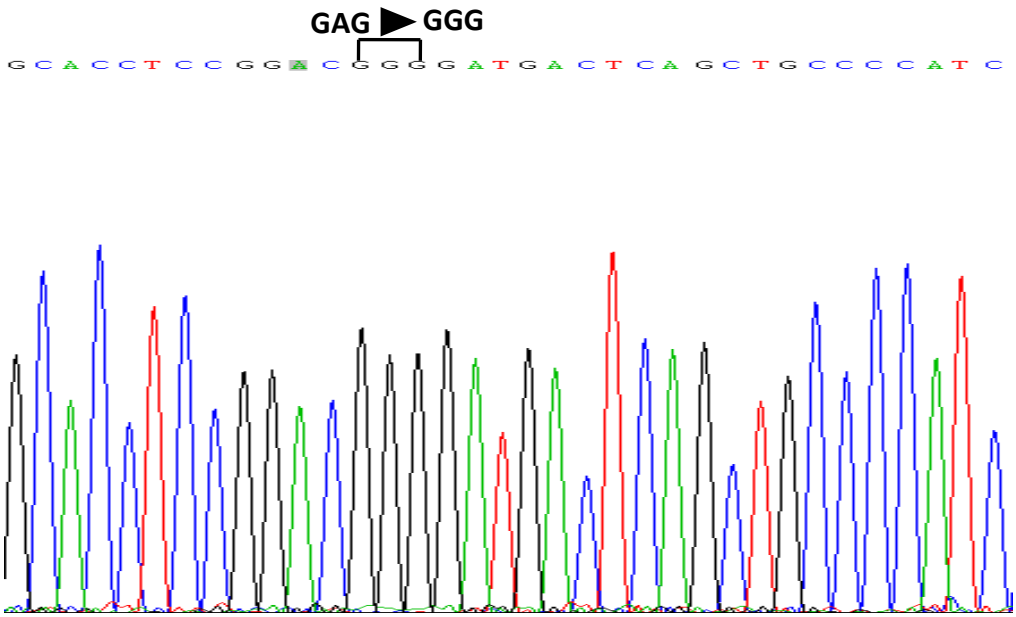


Figure 42: Electropherogram showing E 155 G reported mutation in exon 1 of AR gene in MAIS affected male 'C'. Where E = Glutamic acid and G = Glycine.

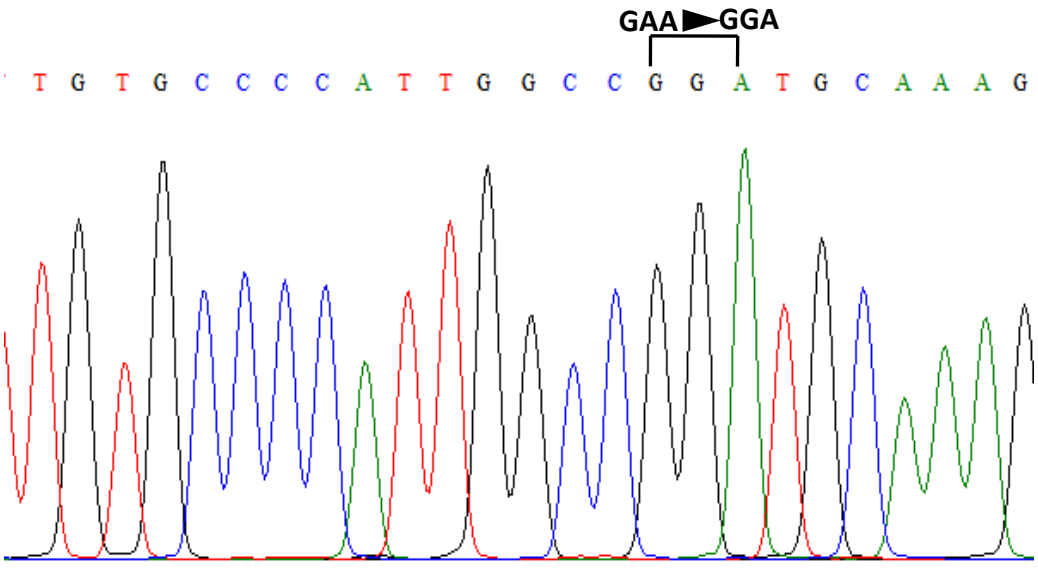


Figure 43: Electropherogram showing E 289 G reported mutation in exon 1 of AR gene in MAIS affected male 'C'. Where E = Glutamic acid and G = Glysine.

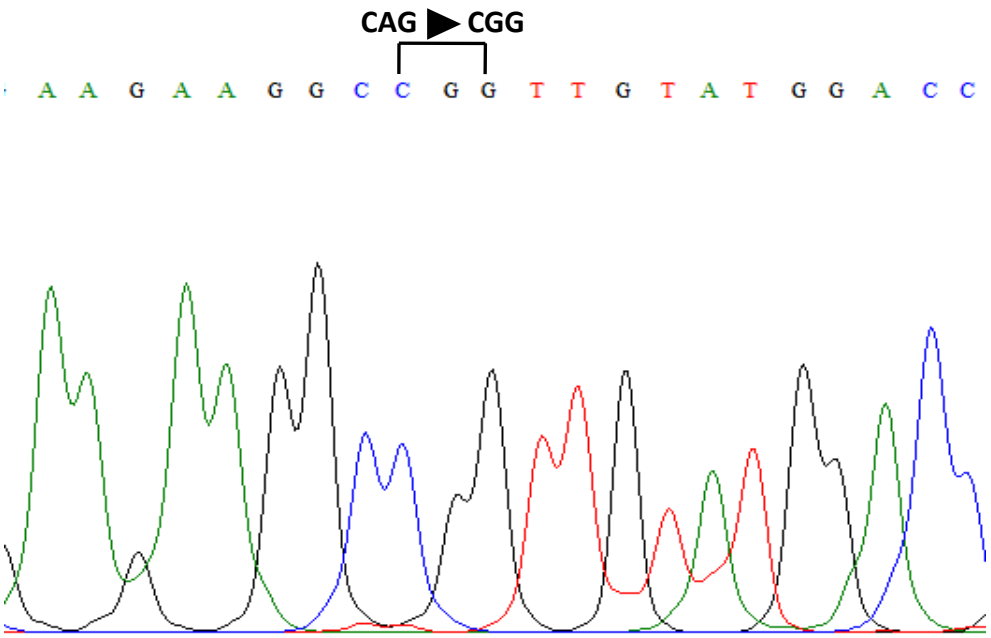


Figure 44: Electropherogram showing Q 445 R reported mutation in exon 1 of AR gene in MAIS affected male 'C'. Where Q = Glutamine and R = Arginine.

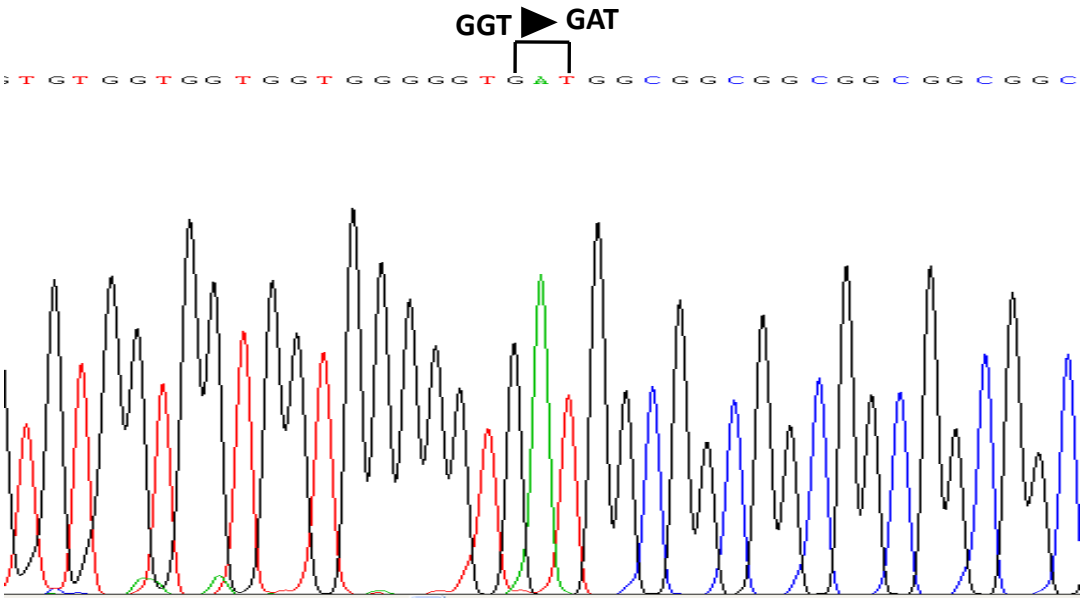


Figure 45: Electropherogram showing the G 456 D reported mutation in exon 1 of AR gene in MAIS affected male 'E'. Where G = Glycine and D = Aspartic acid.

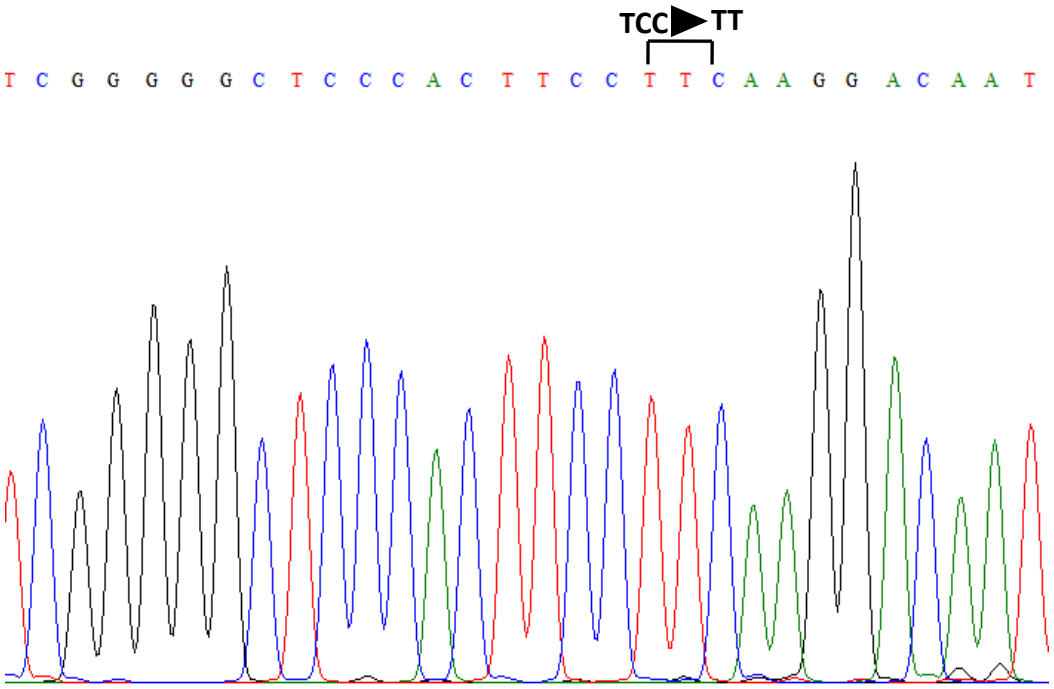


Figure 46: Electropherogram showing S 221 F reported mutation in exon 1 of AR gene in MAIS affected male 'K'. Where S = Serine and F = Phenylalanine.

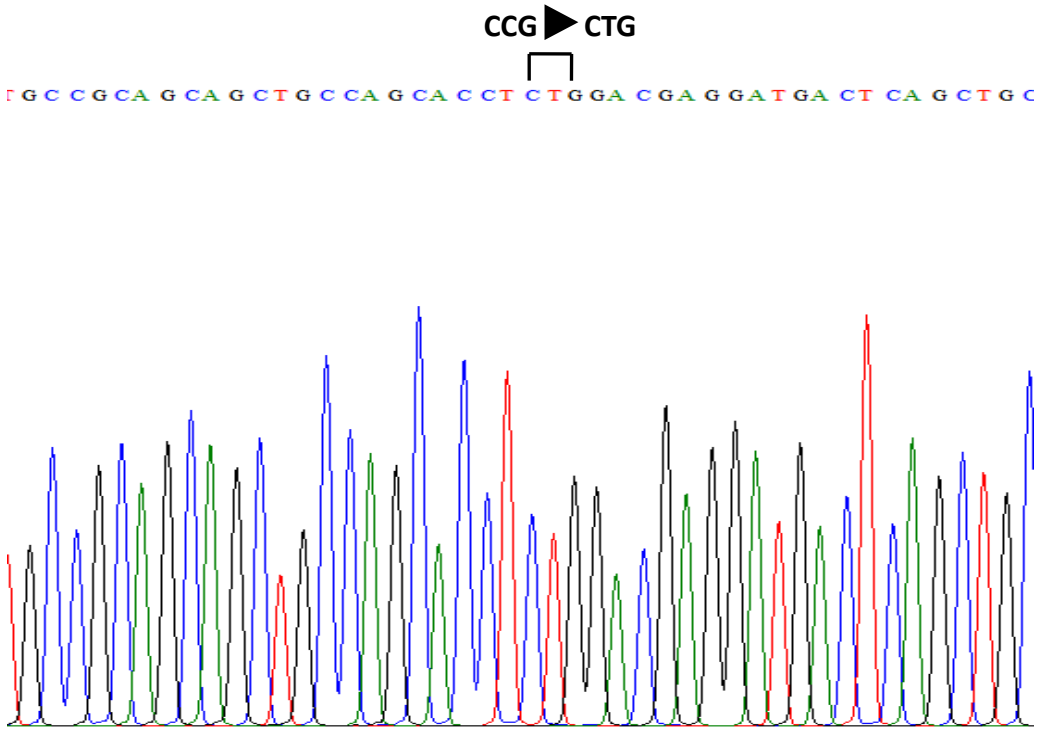


Figure 47: Electropherogram showing P 153 L novel mutation in exon 1 of AR gene in unaffected male 'B'. Where P = Proline and L = Leucine.

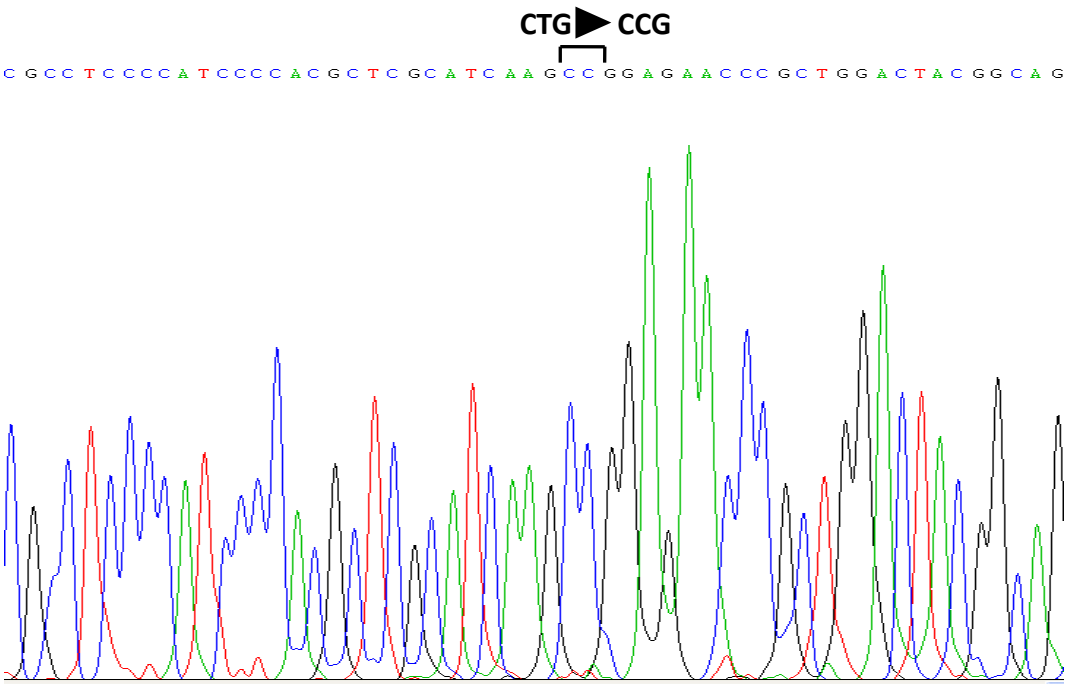


Figure 48: Electropherogram showing L 389 P novel mutation in exon 1 of AR gene in unaffected male 'B'. Where L = Leucine and P = Proline.

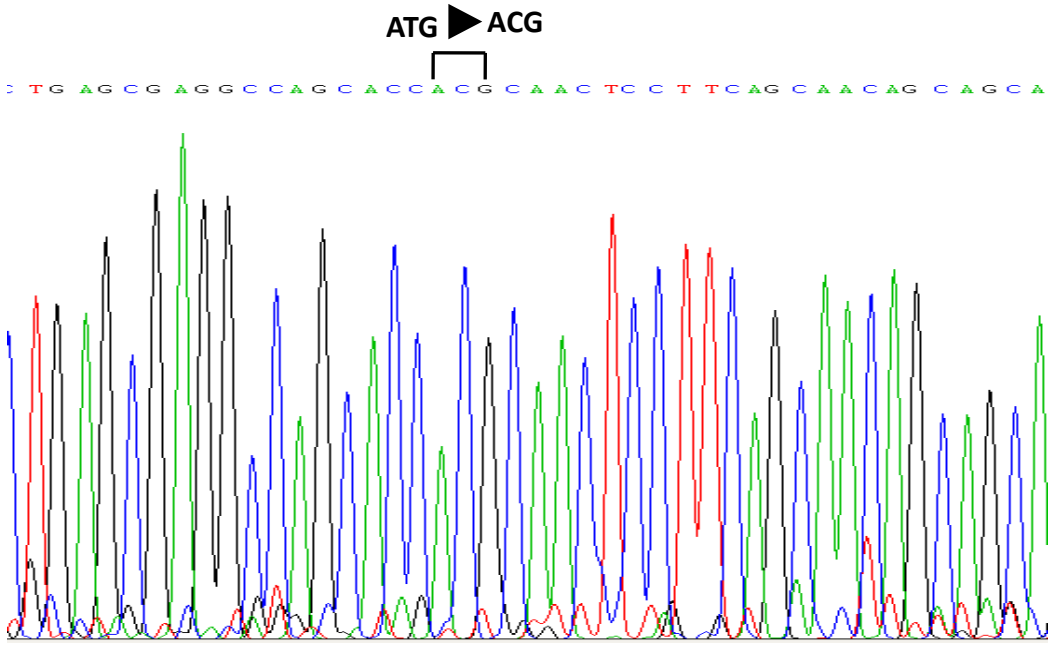


Figure 49: Electropherogram showing M 191 T novel mutation in exon 1 of AR gene in unaffected male 'b'. Where M = Methionine and T = Threonine.

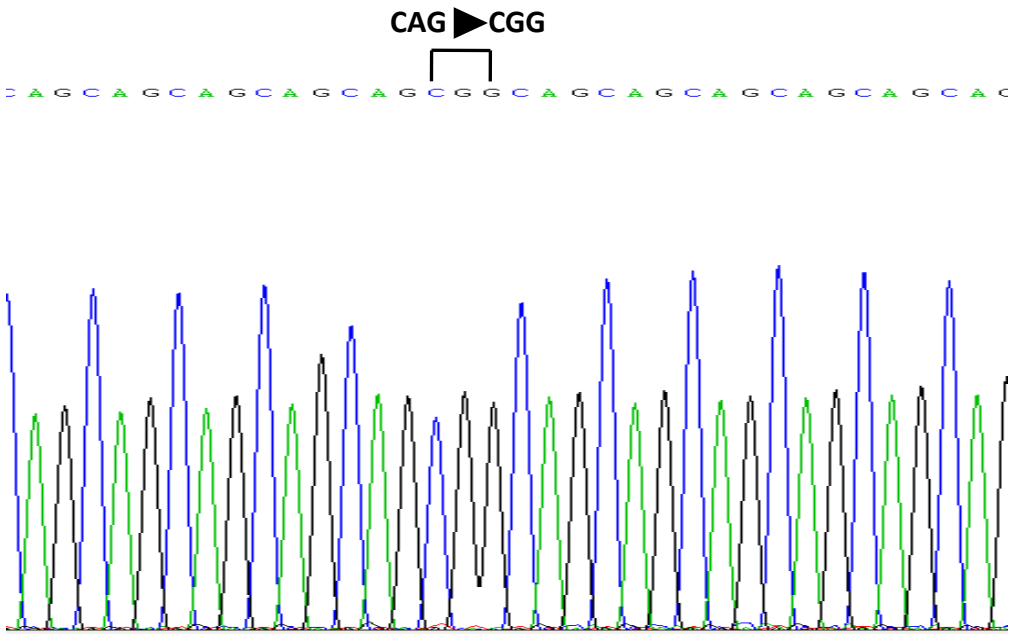


Figure 50: Electropherogram showing Q 72 R novel mutation in exon 1 of AR gene in unaffected female 'F'. Where Q = Glutamine and R = Arginine.

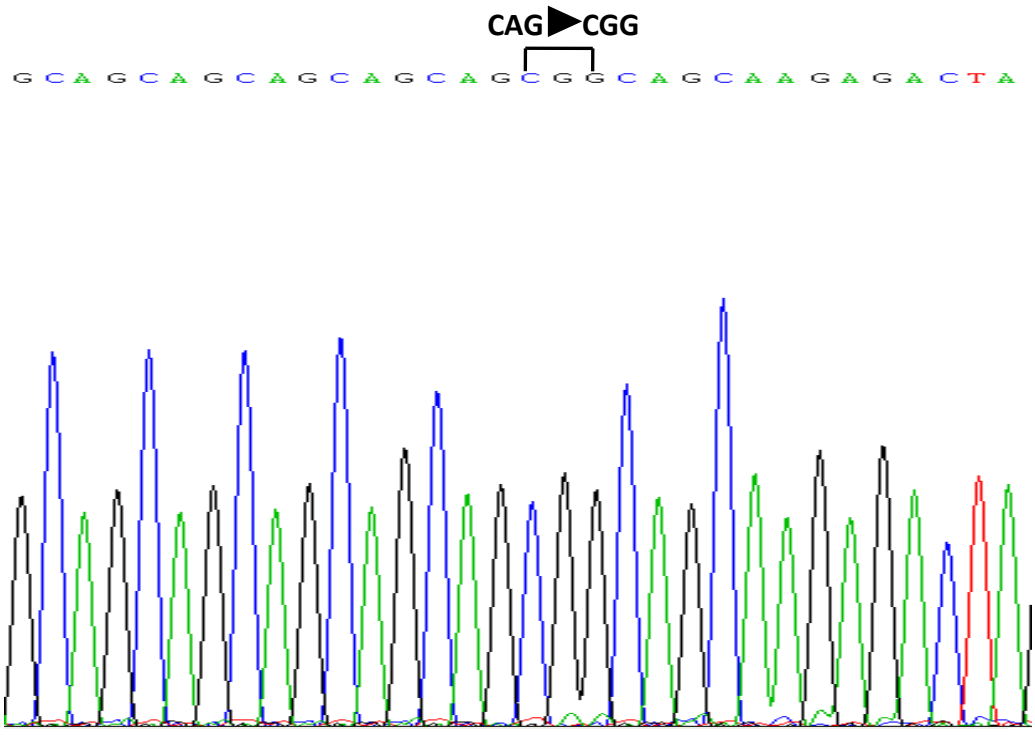


Figure 51: Electropherogram showing Q 78 R novel mutation in exon 1 of AR gene in unaffected female 'F'. Where Q = Glutamine and R = Arginine.

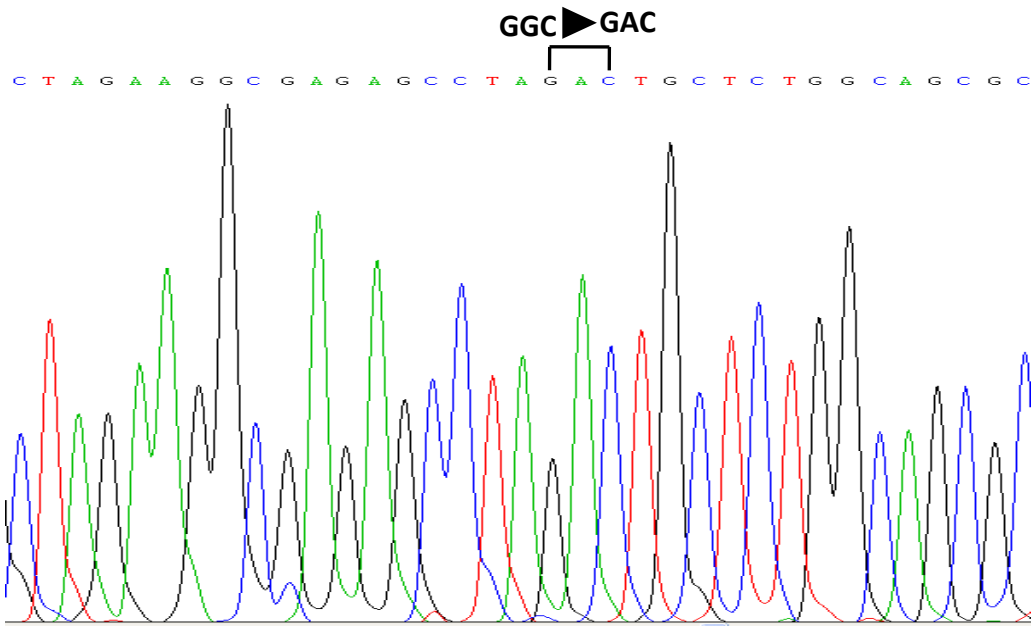


Figure 52: Electropherogram showing G 326 D novel mutation in exon 1 of AR gene in unaffected female 'F'. Where G = Glycine and D = Aspartic acid.

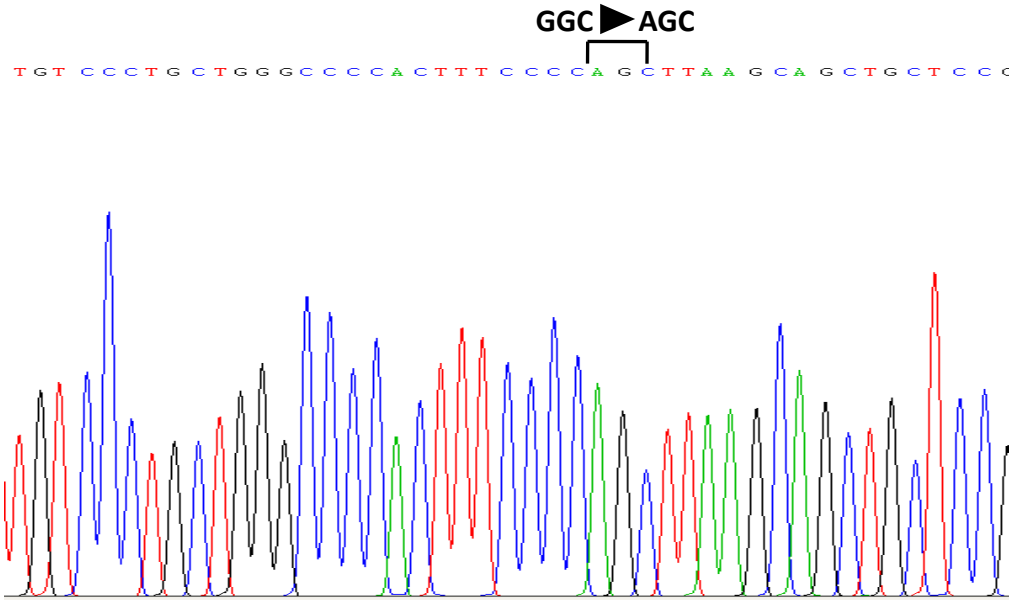


Figure 53: Electropherogram showing G 173 S novel mutation in exon 1 of AR gene in unaffected female 'c'. Where G = Glycine and S = Serine.

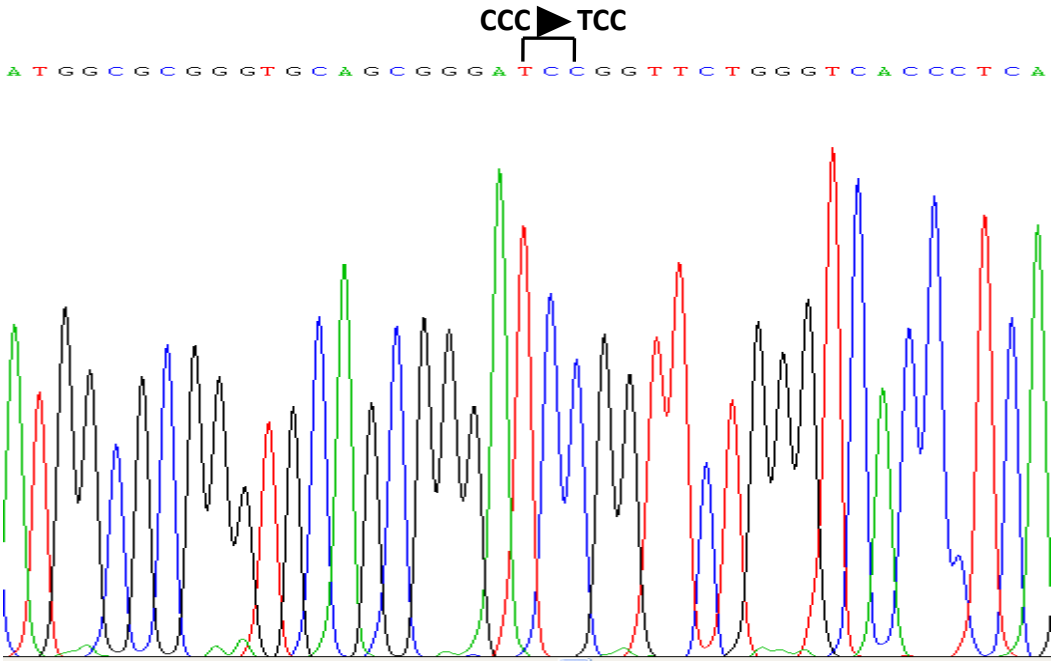


Figure 54: Electropherogram showing P 422 S novel mutation in exon 1 of AR gene in unaffected female 'c'. Where P = Proline and S = Serine.

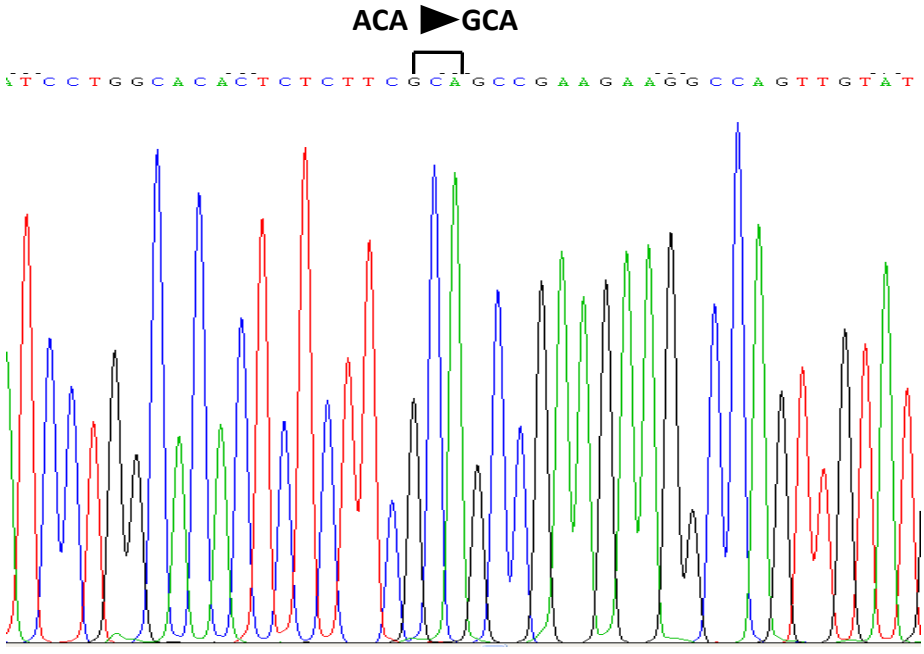


Figure 55: Electropherogram showing the T 440 A reported mutation in exon 1 of AR gene in unaffected female 'F'. Where T = Threonine and A = Alanine.

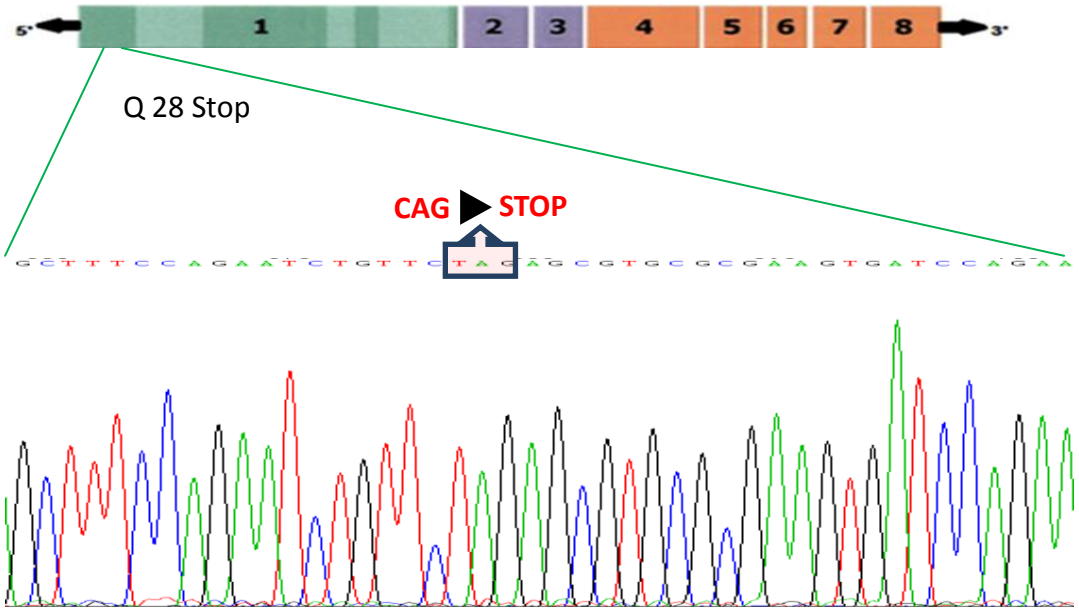


Figure 56: Electropherogram showing Q 28 Stop codon, reported mutation in exon 1 of AR gene in unaffected male 'b'. Where Q = glutamine.

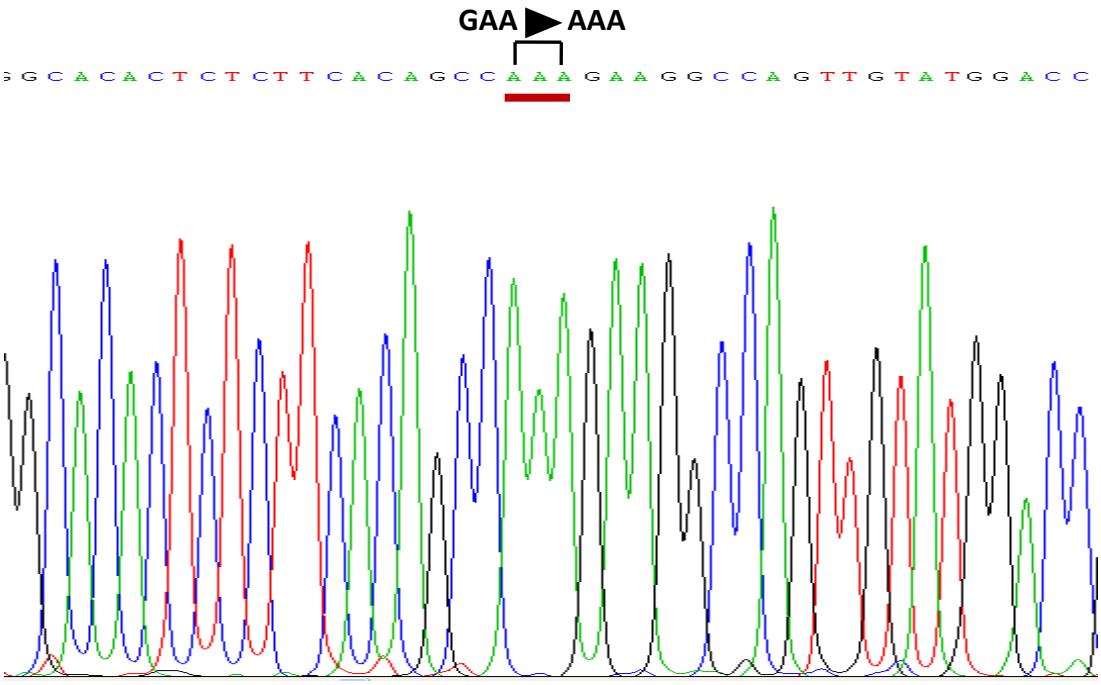


Figure 57: Electropherogram showing the E 442 K reported mutation in exon 1 of AR gene in unaffected male 'b'. Where E = Glutamic acid and K = Lysine.

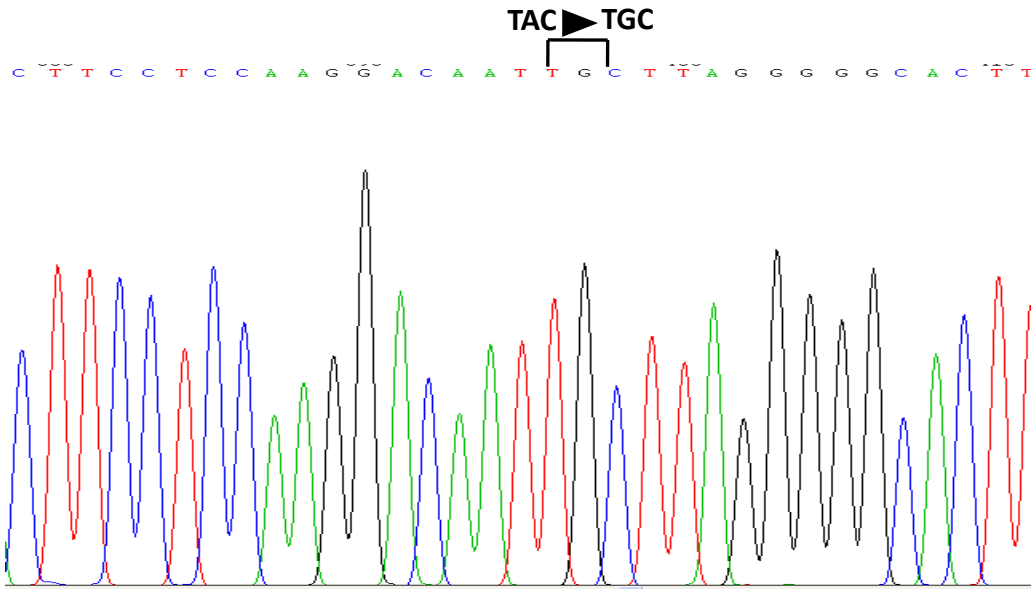


Figure 58: Electropherogram showing the Y 225 C reported mutation in exon 1 of AR gene in unaffected female 'c'. Where Y = Tyrosine and C = Cysteine.

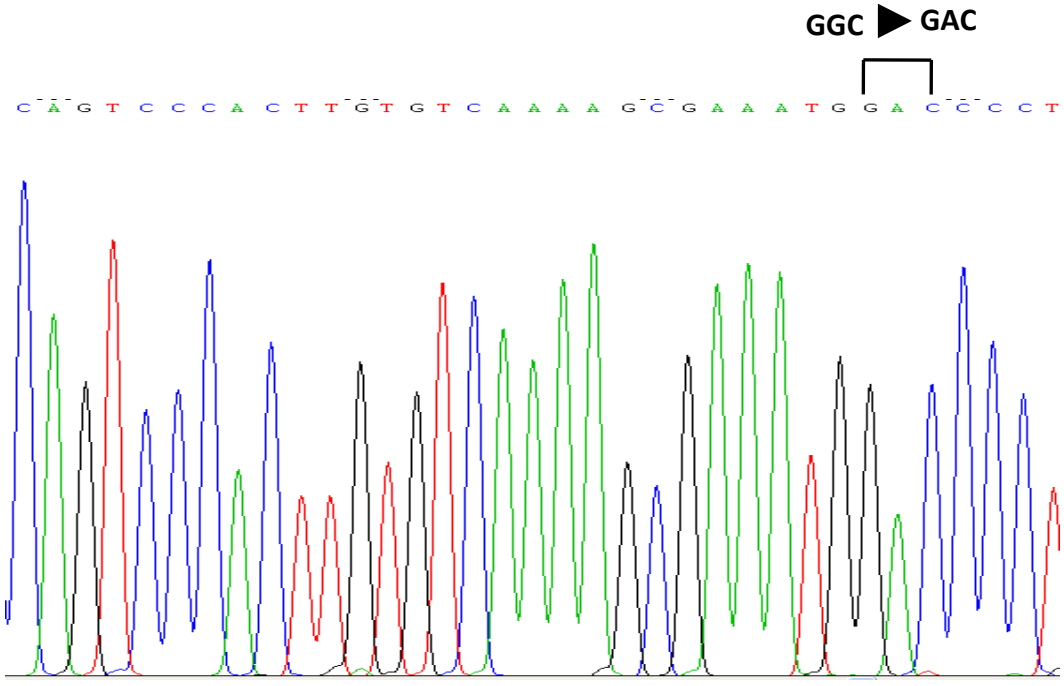


Figure 59: Electropherogram showing the G 525 D reported mutation in exon 1 of AR gene in unaffected female 'c'. Where G = Glycine and D = Aspartic acid.

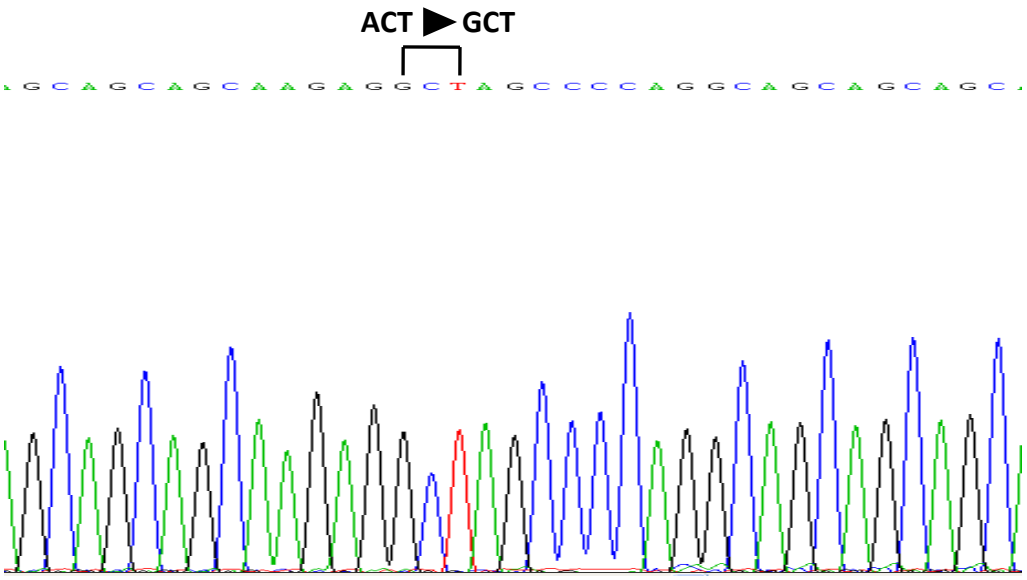


Figure 60: Electropherogram showing T 82 A novel mutation in exon 1 of AR gene in control female 'D'. Where T = Threonine and A = Alanine.

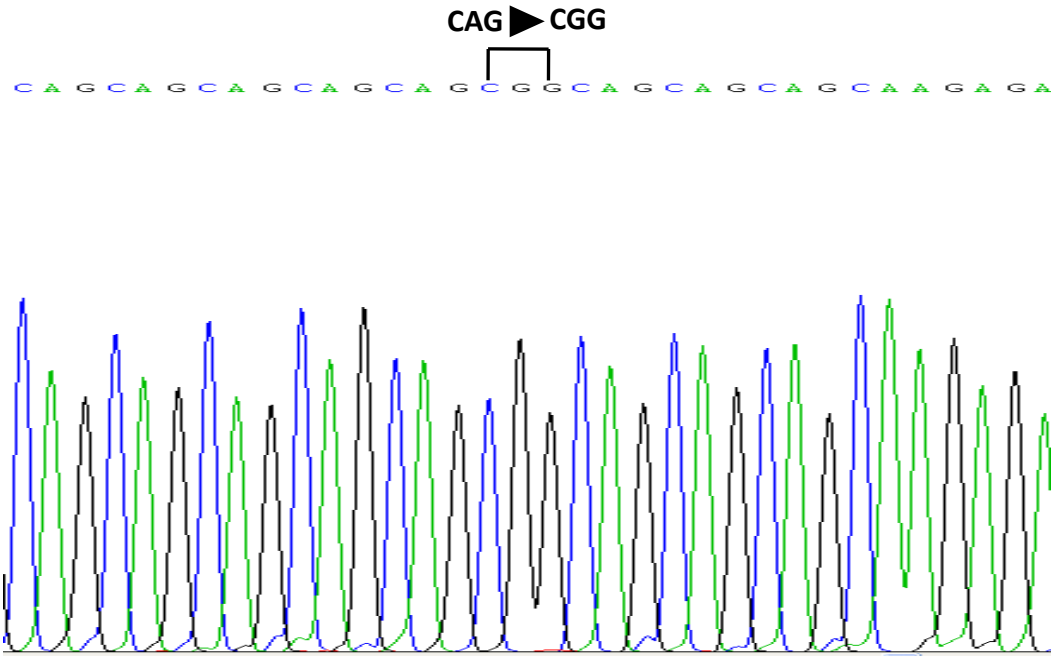


Figure 61: Electropherogram showing Q 79 R novel mutation in exon 1 of AR gene in control female 'G'. Where Q = Glutamine and R = Arginine.

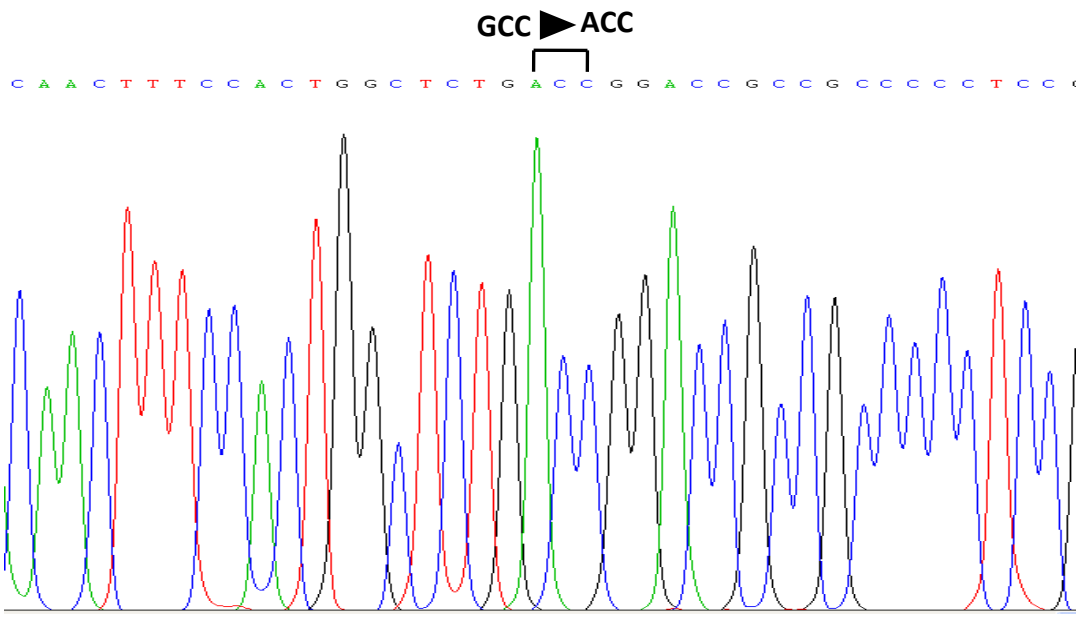


Figure 62: Electropherogram showing A 372 T novel mutation in exon 1 of AR gene in control female 'G'. Where A = Alanine and T = Threonine.

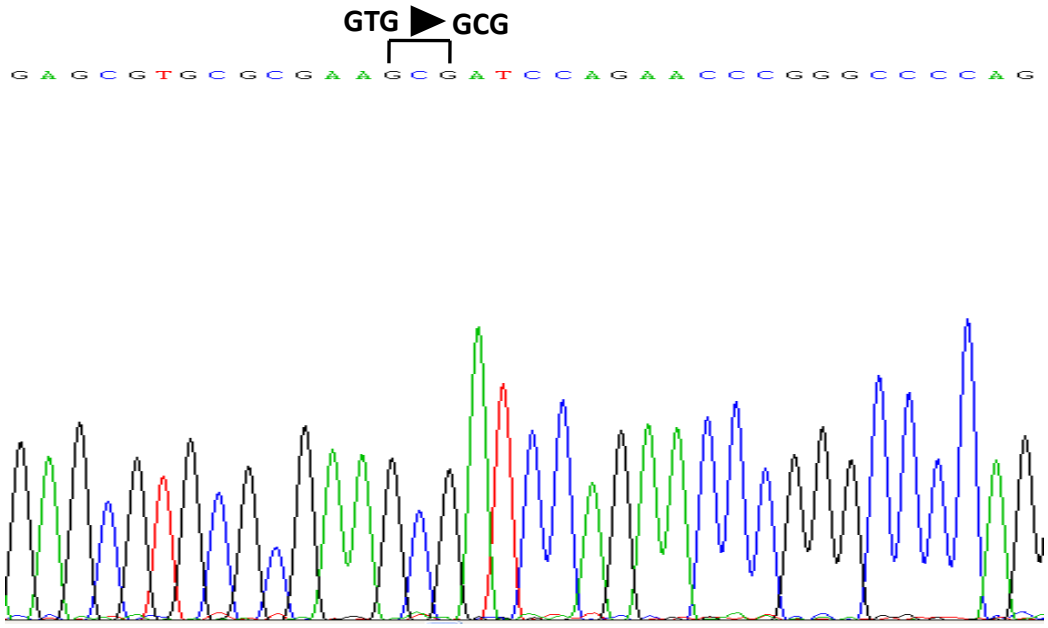


Figure 63: Electropherogram showing V 33 A novel mutation in exon 1 of AR gene in control female 'J'. Where V = Valine and A = Alanine.

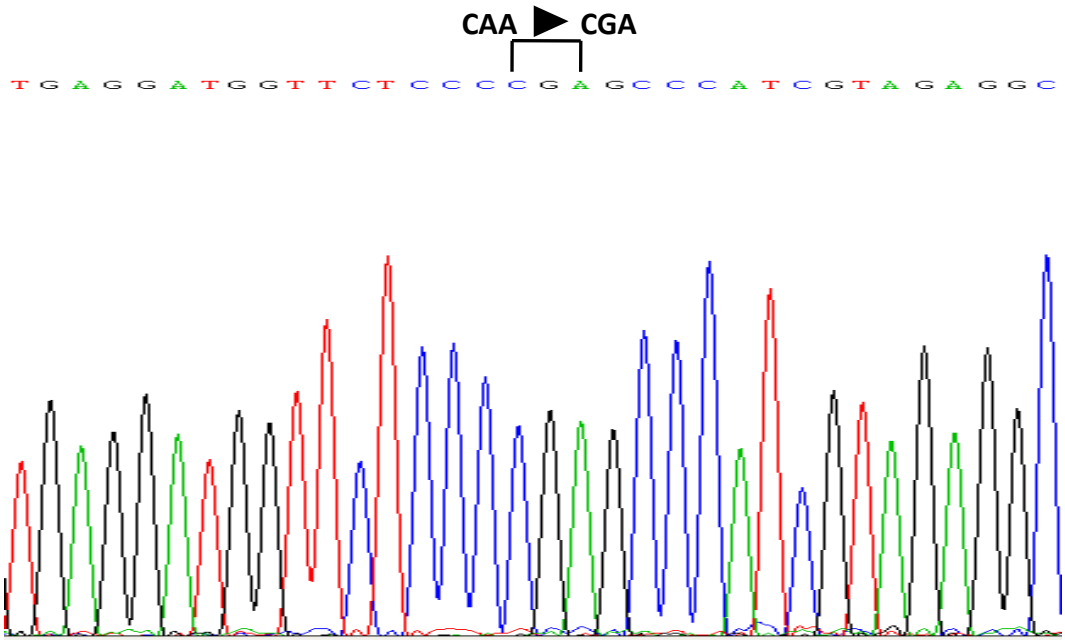


Figure 64: Electropherogram showing Q 98 R novel mutation in exon 1 of AR gene in control female 'J'. Where Q = Glutamine and R = Arginine.

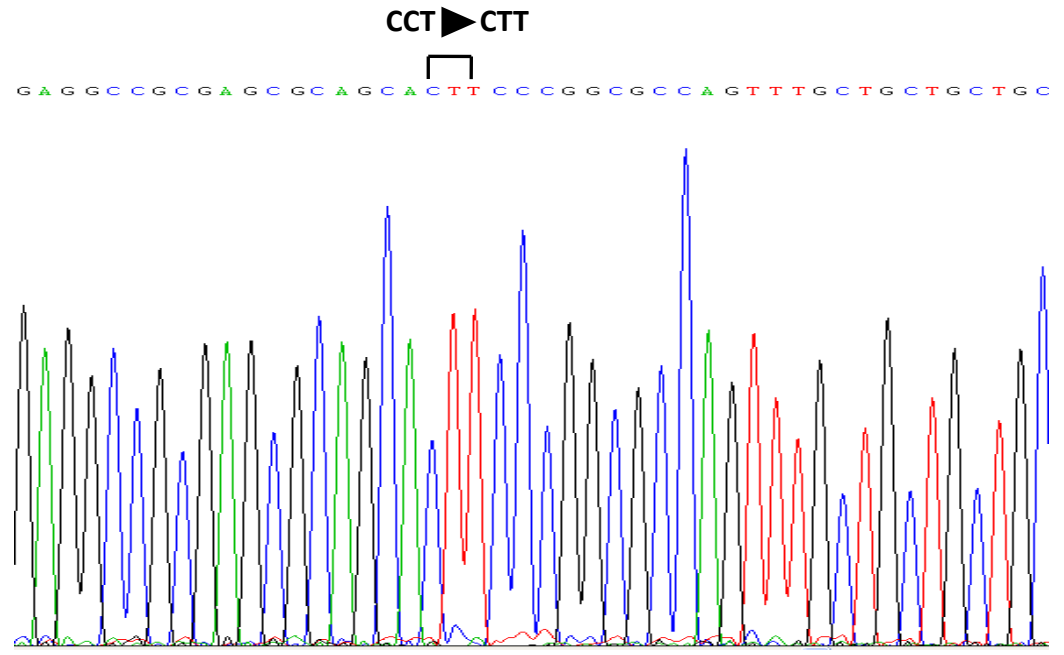


Figure 65: Electropherogram showing P 49 L novel mutation in exon 1 of AR gene in control male 'a'. Where P = Proline and L = Leucine.

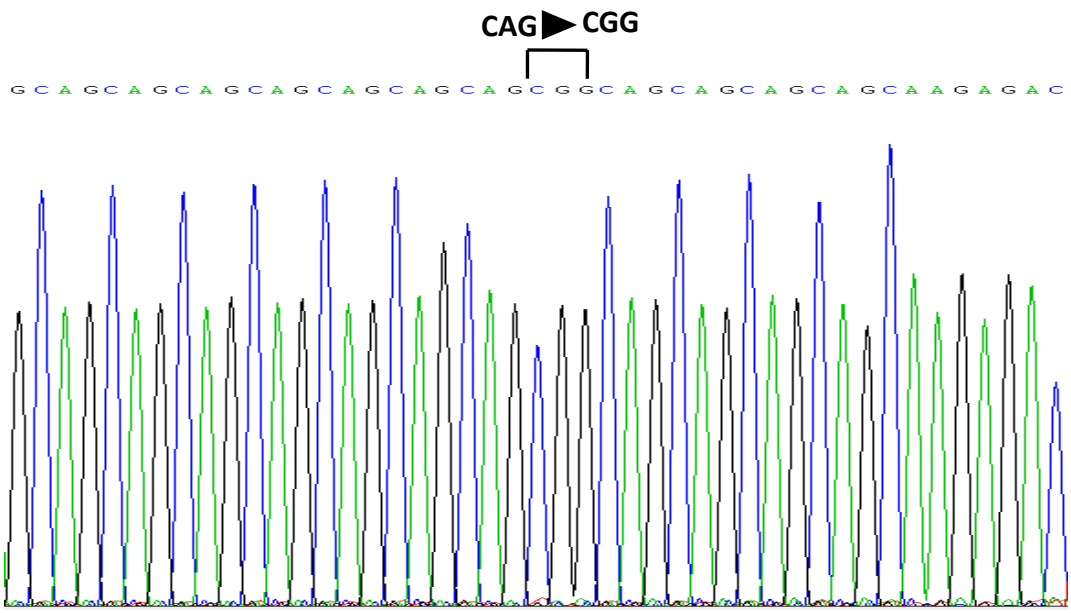


Figure 66: Electropherogram showing Q 79 R novel mutation in exon 1 of AR gene in control male 'a'. Where Q = Glutamine and R = Arginine.

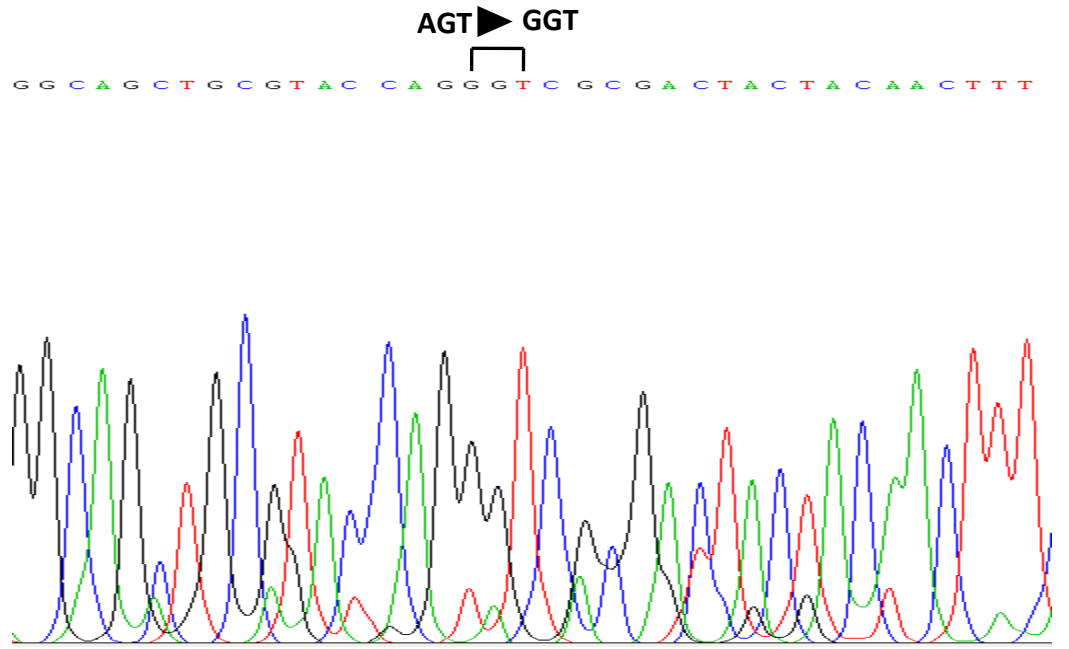


Figure 67: Electropherogram showing S 361 G novel mutation in exon 1 of AR gene in control male 'a'. Where S = Serine and G = Glycine.

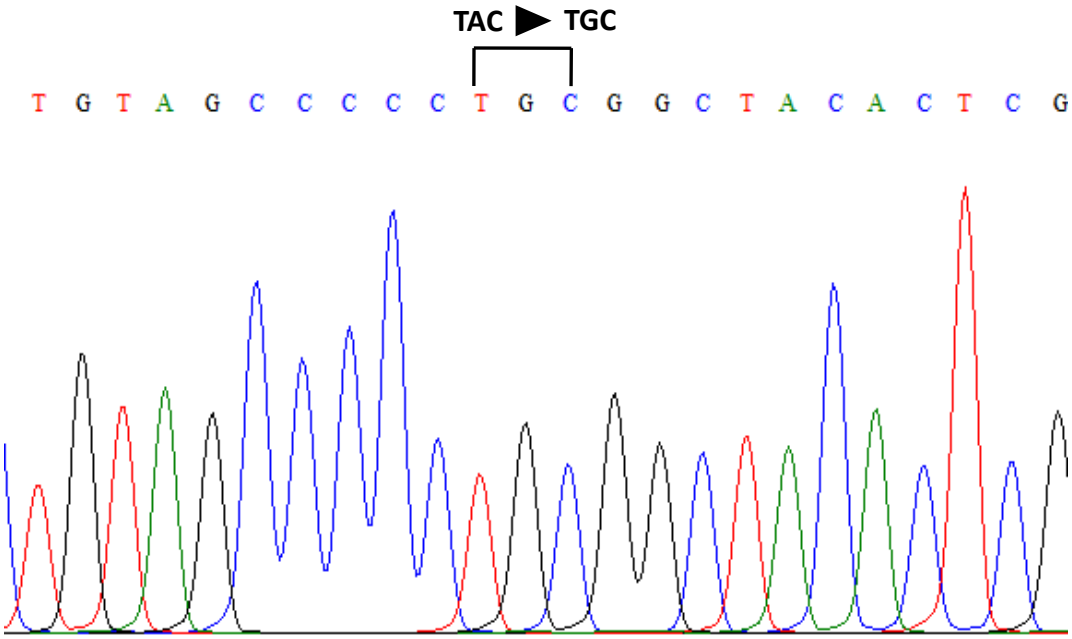


Figure 68: Electropherogram showing the Y 481 C reported mutation in exon 1 of AR gene in control female 'D'. Where Y = Tyrosine and C = Cysteine.

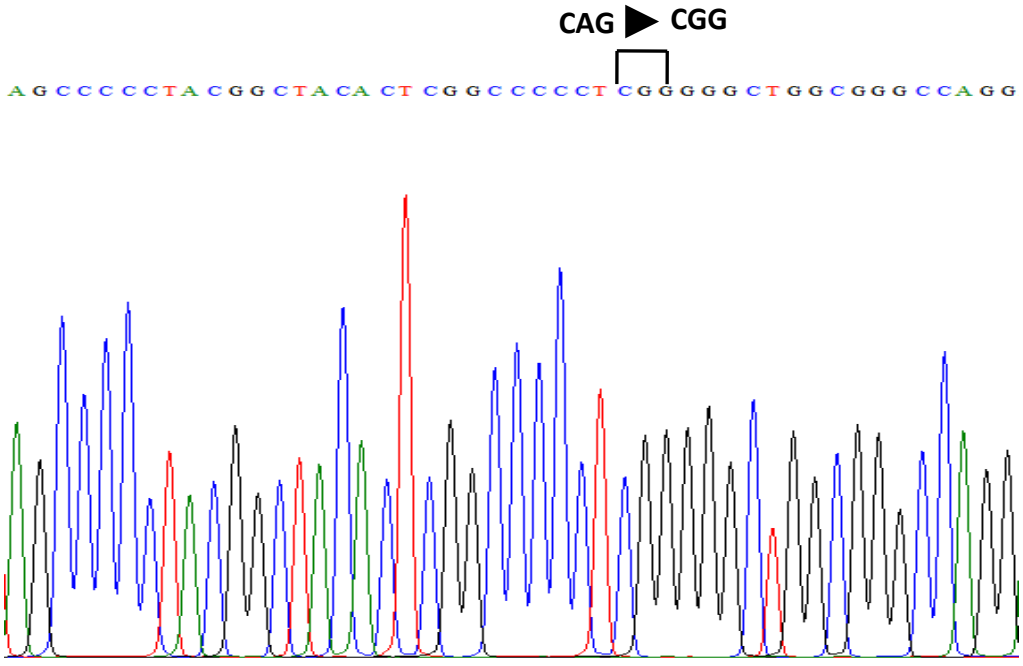


Figure 69: Electropherogram showing the Q 488 R reported mutation in exon 1 of AR gene in the control male 'a'. Where Q = Glutamine and R = Arginine.

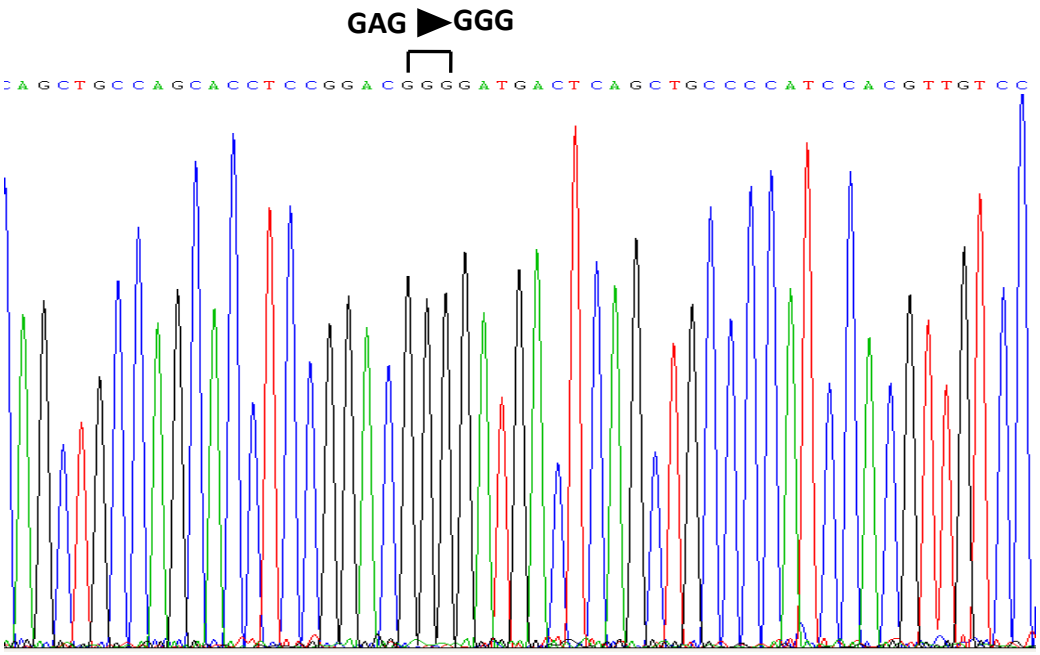


Figure 70: Electropherogram showing the E 155 G reported mutation in exon 1 of AR gene in the control male ‘a’. Where E = Glutamic acid and G = Glycine.

REFERENCES

ACKERMAN, M. J. 1998. The long QT syndrome. *Pediatr Rev,* 19**,** 232-8.

AGOFF, S. N., SWANSON, P. E., LINDEN, H., HAWES, S. E. & LAWTON, T. J. 2003. Androgen receptor expression in estrogen receptor-negative breast cancer. Immunohistochemical, clinical, and prognostic associations. *Am J Clin Pathol,* 120**,** 725-31.

AIMAN, J. & GRIFFIN, J. E. 1982. The frequency of androgen receptor deficiency in infertile men. *J Clin Endocrinol Metab,* 54**,** 725-32.

AIMAN, J., GRIFFIN, J. E., GAZAK, J. M., WILSON, J. D. & MACDONALD, P. C. 1979. Androgen insensitivity as a cause of infertility in otherwise normal men. *N Engl J Med,* 300**,** 223-7.

ALVAREZ-NAVA, F., GONZALEZ, S., SOTO, M., MARTINEZ, C. & PRIETO, M. 1997. Complete androgen insensitivity syndrome: clinical and anatomopathological findings in 23 patients. *Genet Couns,* 8**,** 7-12.

ANGHEL, A., RAICA, M., MARIAN, C., URSONIU, S. & MITRASCA, O. 2006. Combined profile of the tandem repeats CAG, TA and CA of the androgen and estrogen receptor genes in breast cancer. *J Cancer Res Clin Oncol,* 132**,** 727-33.

ARAUJO, A. B., ESCHE, G. R., KUPELIAN, V., O'DONNELL, A. B., TRAVISON, T. G., WILLIAMS, R. E., CLARK, R. V. & MCKINLAY, J. B. 2007. Prevalence of symptomatic androgen deficiency in men. *J Clin Endocrinol Metab,* 92**,** 4241-7.

AUDI, L., FERNANDEZ-CANCIO, M., CARRASCOSA, A., ANDALUZ, P., TORAN, N., PIRO, C., VILARO, E., VICENS-CALVET, E., GUSSINYE, M., ALBISU, M. A., YESTE, D., CLEMENTE, M., HERNANDEZ DE LA CALLE, I., DEL CAMPO, M., VENDRELL, T., BLANCO, A., MARTINEZ-MORA, J., GRANADA, M. L., SALINAS, I., FORN, J., CALAF, J., ANGERRI, O., MARTINEZ-SOPENA, M. J., DEL VALLE, J., GARCIA, E., GRACIA-BOUTHELIER, R., LAPUNZINA, P., MAYAYO, E., LABARTA, J. I., LLEDO, G., SANCHEZ DEL POZO, J., ARROYO, J., PEREZ-AYTES, A., BENEYTO, M., SEGURA, A., BORRAS, V., GABAU, E., CAIMARI, M., RODRIGUEZ, A., MARTINEZ-AEDO, M. J., CARRERA, M., CASTANO, L., ANDRADE, M., BERMUDEZ DE LA VEGA, J. A. & GRUPO DE APOYO AL SINDROME DE INSENSIBILIDAD A LOS, A. 2010. Novel (60%) and recurrent (40%) androgen receptor gene mutations in a series of 59 patients with a 46,XY disorder of sex development. *J Clin Endocrinol Metab,* 95**,** 1876-88.

AVILA, D. M., FUQUA, S. A., GEORGE, F. W. & MCPHAUL, M. J. 1998. Identification of genes expressed in the rat prostate that are modulated differently by castration and Finasteride treatment. *J Endocrinol,* 159**,** 403-11.

BAKAN, R. 1985. Queen Elizabeth I: a case of testicular feminization? *Med Hypotheses,* 17**,** 277-84.

BEATO, M. 1989. Gene regulation by steroid hormones. *Cell,* 56**,** 335-44.

BEATO, M., HERRLICH, P. & SCHUTZ, G. 1995. Steroid hormone receptors: many actors in search of a plot. *Cell,* 83**,** 851-7.

BECK, I. M., VANDEN BERGHE, W., VERMEULEN, L., YAMAMOTO, K. R., HAEGEMAN, G. & DE BOSSCHER, K. 2009. Crosstalk in inflammation: the interplay

of glucocorticoid receptor-based mechanisms and kinases and phosphatases. *Endocr Rev,* 30**,** 830-82.

BEMBO, S. A. AND H. E. CARLSON (2004). "Gynecomastia: its features, and when and how to treat it." Cleve Clin J Med **71**(6): 511-517.

BENNETT, N. C., GARDINER, R. A., HOOPER, J. D., JOHNSON, D. W. & GOBE, G. C. 2010. Molecular cell biology of androgen receptor signalling. *Int J Biochem Cell Biol,* 42**,** 813-27.

BENSON, J. R. AND I. JATOI (2012). "The global breast cancer burden." Future Oncol **8**(6): 697-702.

BERREVOETS, C. A., DOESBURG, P., STEKETEE, K., TRAPMAN, J. & BRINKMANN, A. O. 1998. Functional interactions of the AF-2 activation domain core region of the human androgen receptor with the amino-terminal domain and with the transcriptional coactivator TIF2 (transcriptional intermediary factor2). *Mol Endocrinol,* 12**,** 1172-83.

BETNEY, R. & MCEWAN, I. J. 2003. Role of conserved hydrophobic amino acids in androgen receptor AF-1 function. *J Mol Endocrinol,* 31**,** 427-39.

BEUSCHLEIN, F., BOULKROUN, S., OSSWALD, A., WIELAND, T., NIELSEN, H. N., LICHTENAUER, U. D., PENTON, D., SCHACK, V. R., AMAR, L., FISCHER, E., WALTHER, A., TAUBER, P., SCHWARZMAYR, T., DIENER, S., GRAF, E., ALLOLIO, B., SAMSON-COUTERIE, B., BENECKE, A., QUINKLER, M., FALLO, F., PLOUIN, P. F., MANTERO, F., MEITINGER, T., MULATERO, P., JEUNEMAITRE, X., WARTH, R., VILSEN, B., ZENNARO, M. C., STROM, T. M. & REINCKE, M. 2013. Somatic

mutations in ATP1A1 and ATP2B3 lead to aldosterone-producing adenomas and secondary hypertension. *Nat Genet,* 45**,** 440-4, 444e1-2.

BHASIN, S., P. ENZLIN, A. COVIELLO AND R. Basson (2007). "Sexual dysfunction in men and women with endocrine disorders." Lancet **369**(9561): 597-611.

BOAS, M., BOISEN, K. A., VIRTANEN, H. E., KALEVA, M., SUOMI, A. M., SCHMIDT, I. M., DAMGAARD, I. N., KAI, C. M., CHELLAKOOTY, M., SKAKKEBAEK, N. E., TOPPARI, J. & MAIN, K. M. 2006. Postnatal penile length and growth rate correlate to serum testosterone levels: a longitudinal study of 1962 normal boys. *Eur J Endocrinol,* 154**,** 125-9.

BOCKLANDT, S. & VILAIN, E. 2007. Sex differences in brain and behavior: hormones versus genes. *Adv Genet,* 59**,** 245-66.

BOEHMER, A. L., BRINKMANN, A. O., NIERMEIJER, M. F., BAKKER, L., HALLEY, D. J. & DROP, S. L. 1997. Germ-line and somatic mosaicism in the androgen insensitivity syndrome: implications for genetic counseling. *Am J Hum Genet,* 60**,** 1003-6.

BOEHMER, A. L., BRINKMANN, O., BRUGGENWIRTH, H., VAN ASSENDELFT, C., OTTEN, B. J., VERLEUN-MOOIJMAN, M. C., NIERMEIJER, M. F., BRUNNER, H. G., ROUWE, C. W., WAELKENS, J. J., OOSTDIJK, W., KLEIJER, W. J., VAN DER KWAST, T. H., DE VROEDE, M. A. & DROP, S. L. 2001. Genotype versus phenotype in families with androgen insensitivity syndrome. *J Clin Endocrinol Metab,* 86**,** 4151-60.

BORING, C. C., SQUIRES, T. S., TONG, T. & MONTGOMERY, S. 1994. Cancer statistics, 1994. *CA Cancer J Clin,* 44**,** 7-26.

BOYAR, R. M., MOORE, R. J., ROSNER, W., AIMAN, J., CHIPMAN, J., MADDEN, J. D., MARKS, J. F. & GRIFFIN, J. E. 1978. Studies of gonadotropin-gonadal dynamics in patients with androgen insensitivity. *J Clin Endocrinol Metab,* 47**,** 1116-22.

BRINKMANN, A. O., FABER, P. W., VAN ROOIJ, H. C., KUIPER, G. G., RIS, C., KLAASSEN, P., VAN DER KORPUT, J. A., VOORHORST, M. M., VAN LAAR, J. H., MULDER, E. & ET AL. 1989. The human androgen receptor: domain structure, genomic organization and regulation of expression. *J Steroid Biochem,* 34**,** 307-10.

BRINKMANN, A. O., JENSTER, G., KUIPER, G. G., RIS, C., VAN LAAR, J. H., VAN DER KORPUT, J. A., DEGENHART, H. J., TRIFIRO, M. A., PINSKY, L., ROMALO, G. & ET AL. 1992. The human androgen receptor: structure/function relationship in normal and pathological situations. *J Steroid Biochem Mol Biol,* 41**,** 361-8.

BRODSKY, J. L. & CHIOSIS, G. 2006. Hsp70 molecular chaperones: emerging roles in human disease and identification of small molecule modulators. *Curr Top Med Chem,* 6**,** 1215-25.

BROOKE, G. N., PARKER, M. G. & BEVAN, C. L. 2008. Mechanisms of androgen receptor activation in advanced prostate cancer: differential co-activator recruitment and gene expression. *Oncogene,* 27**,** 2941-50.

BROWN, C. J., GOSS, S. J., LUBAHN, D. B., JOSEPH, D. R., WILSON, E. M., FRENCH, F. S. & WILLARD, H. F. 1989. Androgen receptor locus on the human X

chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. *Am J Hum Genet,* 44**,** 264-9.

BRUCHOVSKY, N. & WILSON, J. D. 1968. The conversion of testosterone to 5-alpha-androstan-17-beta-ol-3-one by rat prostate in vivo and in vitro. *J Biol Chem,* 243**,** 2012-21.

BUCHANAN, G., BIRRELL, S. N., PETERS, A. A., BIANCO-MIOTTO, T., RAMSAY, K., COPS, E. J., YANG, M., HARRIS, J. M., SIMILA, H. A., MOORE, N. L., BENTEL, J. M., RICCIARDELLI, C., HORSFALL, D. J., BUTLER, L. M. & TILLEY, W. D. 2005. Decreased androgen receptor levels and receptor function in breast cancer contribute to the failure of response to medroxyprogesterone acetate. *Cancer Res,* 65**,** 8487-96.

BURD, C. J., MOREY, L. M. & KNUDSEN, K. E. 2006. Androgen receptor corepressors and prostate cancer. *Endocr Relat Cancer,* 13**,** 979-94.

CALDERWOOD, S. K., KHALEQUE, M. A., SAWYER, D. B. & CIOCCA, D. R. 2006. Heat shock proteins in cancer: chaperones of tumorigenesis. *Trends Biochem Sci,* 31**,** 164-72.

CALLEWAERT, L., VAN TILBORGH, N. & CLAESSENS, F. 2006. Interplay between two hormone-independent activation domains in the androgen receptor. *Cancer Res,* 66**,** 543-53.

CARDOZO, L., G. BACHMANN, D. MCCLISH, D. FONDA AND L. BIRGERSON (1998). "Meta-analysis of estrogen therapy in the management of urogenital atrophy in

postmenopausal women: second report of the Hormones and Urogenital Therapy Committee." Obstet Gynecol **92**(4 Pt 2): 722-727.

CHAMBERLAIN, N. L., DRIVER, E. D. & MIESFELD, R. L. 1994. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Res,* 22**,** 3181-6.

CHAMBERLAIN, N. L., WHITACRE, D. C. & MIESFELD, R. L. 1996. Delineation of two distinct type 1 activation functions in the androgen receptor amino-terminal domain. *J Biol Chem,* 271**,** 26772-8.

CHANG, B. L., ZHENG, S. L., HAWKINS, G. A., ISAACS, S. D., WILEY, K. E., TURNER, A., CARPTEN, J. D., BLEECKER, E. R., WALSH, P. C., TRENT, J. M., MEYERS, D. A., ISAACS, W. B. & XU, J. 2002. Polymorphic GGC repeats in the androgen receptor gene are associated with hereditary and sporadic prostate cancer risk. *Hum Genet,* 110**,** 122-9.

CHENG, S., BRZOSTEK, S., LEE, S. R., HOLLENBERG, A. N. & BALK, S. P. 2002. Inhibition of the dihydrotestosterone-activated androgen receptor by nuclear receptor corepressor. *Mol Endocrinol,* 16**,** 1492-501.

CHMELAR, R., BUCHANAN, G., NEED, E. F., TILLEY, W. & GREENBERG, N. M. 2007. Androgen receptor coregulators and their involvement in the development and progression of prostate cancer. *Int J Cancer,* 120**,** 719-33.

CLAESSENS, F., CELIS, L., PEETERS, B., HEYNS, W., VERHOEVEN, G. & ROMBAUTS, W. 1989. Functional characterization of an androgen response element in the

first intron of the C3(1) gene of prostatic binding protein. *Biochem Biophys Res Commun,* 164**,** 833-40.

CLAESSENS, F., VERRIJDT, G., SCHOENMAKERS, E., HAELENS, A., PEETERS, B., VERHOEVEN, G. & ROMBAUTS, W. 2001. Selective DNA binding by the androgen receptor as a mechanism for hormone-specific gene regulation. *J Steroid Biochem Mol Biol,* 76**,** 23-30.

CONCATO, J., HORWITZ, R. I., FEINSTEIN, A. R., ELMORE, J. G. & SCHIFF, S. F. 1992. Problems of comorbidity in mortality after prostatectomy. *JAMA,* 267**,** 1077-82.

COPELLI, S. B., LUMBROSO, S., AUDRAN, F., PELLIZZARI, E. H., HEINRICH, J. J., CIGORRAGA, S. B., SULTAN, C. & CHEMES, H. E. 1999. A novel E153X point mutation in the androgen receptor gene in a patient with complete androgen insensitivity syndrome. *Asian J Androl,* 1**,** 73-7.

CREIGHTON, T. E. 1990. Protein folding. *Biochem J,* 270**,** 1-16.

CULIG, Z., HOBISCH, A., BARTSCH, G. & KLOCKER, H. 2000a. Androgen receptor--an update of mechanisms of action in prostate cancer. *Urol Res,* 28**,** 211-9.

CULIG, Z., HOBISCH, A., BARTSCH, G. & KLOCKER, H. 2000b. Expression and function of androgen receptor in carcinoma of the prostate. *Microsc Res Tech,* 51**,** 447-55.

CUNDY, T. F., REES, M., EVANS, B. A., HUGHES, I. A., BUTLER, J. & WHEELER, M. J. 1986. Mild androgen insensitivity presenting with sexual dysfunction. *Fertil Steril,* 46**,** 721-3.

CUNHA, G. R. 1994. Role of mesenchymal-epithelial interactions in normal and abnormal development of the mammary gland and prostate. *Cancer,* 74**,** 1030-44.

DAGAN, E., FRIEDMAN, E., PAPERNA, T., CARMI, N. & GERSHONI-BARUCH, R. 2002. Androgen receptor CAG repeat length in Jewish Israeli women who are BRCA1/2 mutation carriers: association with breast/ovarian cancer phenotype. *Eur J Hum Genet,* 10**,** 724-8.

DARNE, C. H., MOREL, L., CLAESSENS, F., MANIN, M., FABRE, S., VEYSSIERE, G., ROMBAUTS, W. & JEAN, C. L. 1997. Ubiquitous transcription factors NF1 and Sp1 are involved in the androgen activation of the mouse vas deferens protein promoter. *Mol Cell Endocrinol,* 132**,** 13-23.

DAUGAARD, M., ROHDE, M. & JAATTELA, M. 2007. The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Lett,* 581**,** 3702-10.

DE ABREU, F. B., PIROLO, L. J., CANEVARI RDE, A., ROSA, F. E., MORAES NETO, F. A., CALDEIRA, J. R., RAINHO, C. A. & ROGATTO, S. R. 2007. Shorter CAG repeat in the AR gene is associated with atypical hyperplasia and breast carcinoma. *Anticancer Res,* 27**,** 1199-205.

DE MAIO, A. 1999. Heat shock proteins: facts, thoughts, and dreams. *Shock,* 11**,** 1-12.

DEBES, J. D. & TINDALL, D. J. 2002. The role of androgens and the androgen receptor in prostate cancer. *Cancer Lett,* 187**,** 1-7.

DERKACZ, M., CHMIEL-PERZYNSKA, I. & NOWAKOWSKI, A. 2011. Gynecomastia - a difficult diagnostic problem. *Endokrynol Pol,* 62**,** 190-202.

DESARNAUD, F., P. GECK, C. PARKIN, G. CARPINITO AND A. N. MAKAROVSKIY (2011). "Gene expression profiling of the androgen independent prostate cancer cells demonstrates complex mechanisms mediating resistance to docetaxel." Cancer Biol Ther **11**(2): 204-212.

DEVOS, A., CLAESSENS, F., ALEN, P., WINDERICKX, J., HEYNS, W., ROMBAUTS, W. & PEETERS, B. 1997. Identification of a functional androgen-response element in the exon 1-coding sequence of the cystatin-related protein gene crp2. *Mol Endocrinol,* 11**,** 1033-43.

DOANE, A. S., DANSO, M., LAL, P., DONATON, M., ZHANG, L., HUDIS, C. & GERALD, W. L. 2006. An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen. *Oncogene,* 25**,** 3994-4008.

DUNNING, A. M., MCBRIDE, S., GREGORY, J., DUROCHER, F., FOSTER, N. A., HEALEY, C. S., SMITH, N., PHAROAH, P. D., LUBEN, R. N., EASTON, D. F. & PONDER, B. A. 1999. No association between androgen or vitamin D receptor gene polymorphisms and risk of breast cancer. *Carcinogenesis,* 20**,** 2131-5.

EDWARDS, S. M., BADZIOCH, M. D., MINTER, R., HAMOUDI, R., COLLINS, N., ARDERN-JONES, A., DOWE, A., OSBORNE, S., KELLY, J., SHEARER, R., EASTON, D. F., SAUNDERS, G. F., DEARNALEY, D. P. & EELES, R. A. 1999. Androgen receptor

polymorphisms: association with prostate cancer risk, relapse and overall survival. *Int J Cancer,* 84**,** 458-65.

ELHAJI, Y. A., GOTTLIEB, B., LUMBROSO, R., BEITEL, L. K., FOULKES, W. D., PINSKY, L. & TRIFIRO, M. A. 2001. The polymorphic CAG repeat of the androgen receptor gene: a potential role in breast cancer in women over 40. *Breast Cancer Res Treat,* 70**,** 109-16.

FABER, P. W., VAN ROOIJ, H. C., VAN DER KORPUT, H. A., BAARENDS, W. M., BRINKMANN, A. O., GROOTEGOED, J. A. & TRAPMAN, J. 1991. Characterization of the human androgen receptor transcription unit. *J Biol Chem,* 266**,** 10743-9.

FARLA, P., HERSMUS, R., GEVERTS, B., MARI, P. O., NIGG, A. L., DUBBINK, H. J., TRAPMAN, J. & HOUTSMULLER, A. B. 2004. The androgen receptor ligand-binding domain stabilizes DNA binding in living cells. *J Struct Biol,* 147**,** 50-61.

FERLIN, A., F. RAICU, V. GATTA, D. ZUCCARELLO, G. PALKA AND C. FORESTA (2007). "Male infertility: role of genetic background." Reprod Biomed Online **14**(6): 734-745.

FRANK, S. A. & NOWAK, M. A. 2004. Problems of somatic mutation and cancer. *Bioessays,* 26**,** 291-9.

FREEDMAN, L. P. 1992. Anatomy of the steroid receptor zinc finger region. *Endocr Rev,* 13**,** 129-45.

FREEDMAN, M. L., C. L. PEARCE, K. L. PENNEY, J. N. HIRSCHHORN, L. N. KOLONEL, B. E. HENDERSON AND D. ALTSHULER (2005). "Systematic evaluation of genetic

variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study." Am J Hum Genet **76**(1): 82-90.

FREEMAN, E. R., BLOOM, D. A. & MCGUIRE, E. J. 2001. A brief history of testosterone. *J Urol,* 165**,** 371-3.

FUJIMOTO, N., MIZOKAMI, A., HARADA, S. & MATSUMOTO, T. 2001. Different expression of androgen receptor coactivators in human prostate. *Urology,* 58**,** 289-94.

FUJIMOTO, N., YEH, S., KANG, H. Y., INUI, S., CHANG, H. C., MIZOKAMI, A. & CHANG, C. 1999. Cloning and characterization of androgen receptor coactivator, ARA55, in human prostate. *J Biol Chem,* 274**,** 8316-21.

GALANI, A., KITSIOU-TZELI, S., SOFOKLEOUS, C., KANAVAKIS, E. & KALPINI-MAVROU, A. 2008. Androgen insensitivity syndrome: clinical features and molecular defects. *Hormones (Athens),* 7**,** 217-29.

GANS, S. L. & RUBIN, C. L. 1962. Apparent female infants with hernias and testes. *Am J Dis Child,* 104**,** 82-6.

GAO, T., MARCELLI, M. & MCPHAUL, M. J. 1996. Transcriptional activation and transient expression of the human androgen receptor. *J Steroid Biochem Mol Biol,* 59**,** 9-20.

GAO, W., BOHL, C. E. & DALTON, J. T. 2005. Chemistry and structural biology of androgen receptor. *Chem Rev,* 105**,** 3352-70.

GAO, X., LOGGIE, B. W. & NAWAZ, Z. 2002. The roles of sex steroid receptor coregulators in cancer. *Mol Cancer,* 1**,** 7.

GARDNER DG, S. D. 2011. Green Span's basic and clinical endocrinology. chapter 12, Testes.

GELMANN, E. P. 2002. Molecular biology of the androgen receptor. *J Clin Oncol,* 20**,** 3001-15.

GHADESSY, F. J., J. LIM, A. A. ABDULLAH, V. PANET-RAYMOND, C. K. CHOO, R. LUMBROSO, T. G. TUT, B. GOTTLIEB, L. PINSKY, M. A. TRIFIRO AND E. L. YONG (1999). "Oligospermic infertility associated with an androgen receptor mutation that disrupts interdomain and coactivator (TIF2) interactions." J Clin Invest **103**(11): 1517-1525.

GIAMMANCO, M., TABACCHI, G., GIAMMANCO, S., DI MAJO, D. & LA GUARDIA, M. 2005. Testosterone and aggressiveness. *Med Sci Monit,* 11**,** RA136-45.

GIGUERE, Y., DEWAILLY, E., BRISSON, J., AYOTTE, P., LAFLAMME, N., DEMERS, A., FOREST, V. I., DODIN, S., ROBERT, J. & ROUSSEAU, F. 2001. Short polyglutamine tracts in the androgen receptor are protective against breast cancer in the general population. *Cancer Res,* 61**,** 5869-74.

GIWERCMAN, Y. L., NORDENSKJOLD, A., RITZEN, E. M., NILSSON, K. O., IVARSSON, S. A., GRANDELL, U. & WEDELL, A. 2002. An androgen receptor gene mutation (E653K) in a family with congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency as well as in partial androgen insensitivity. *J Clin Endocrinol Metab,* 87**,** 2623-8.

GLASS, C. K. & ROSENFELD, M. G. 2000. The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev,* 14**,** 121-41.

GONZALEZ-PEREZ, E., ESTEBAN, E., VIA, M., GARCIA-MORO, C., HERNANDEZ, M. & MORAL, P. 2006. Genetic change in the polynesian population of Easter Island: evidence from Alu insertion polymorphisms. *Ann Hum Genet,* 70**,** 829-40.

GONZALEZ, A., JAVIER DORTA, F., RODRIGUEZ, G., BRITO, B., RODRIGUEZ, M. A., CABRERA, A., DIAZ-CHICO, J. C., REYES, R., AGUIRRE-JAIME, A. & NICOLAS DIAZ-CHICO, B. 2007. Increased risk of breast cancer in women bearing a combination of large CAG and GGN repeats in the exon 1 of the androgen receptor gene. *Eur J Cancer,* 43**,** 2373-80.

GOTTLIEB, B., BEITEL, L. K., NADARAJAH, A., PALIOURAS, M. & TRIFIRO, M. 2012. The androgen receptor gene mutations database: 2012 update. *Hum Mutat,* 33**,** 887-94.

GOTTLIEB, B., BEITEL, L. K. & TRIFIRO, M. A. 2001a. Somatic mosaicism and variable expressivity. *Trends Genet,* 17**,** 79-82.

GOTTLIEB, B., BEITEL, L. K. & TRIFIRO, M. A. 2001b. Variable expressivity and mutation databases: The androgen receptor gene mutations database. *Hum Mutat,* 17**,** 382-8.

GOTTLIEB, B., BEITEL, L. K., WU, J. H. & TRIFIRO, M. 2004. The androgen receptor gene mutations database (ARDB): 2004 update. *Hum Mutat,* 23**,** 527-33.

GOTTLIEB, B., LOMBROSO, R., BEITEL, L. K. & TRIFIRO, M. A. 2005. Molecular pathology of the androgen receptor in male (in)fertility. *Reprod Biomed Online,* 10**,** 42-8.

GOTTLIEB, B., PINSKY, L., BEITEL, L. K. & TRIFIRO, M. 1999a. Androgen insensitivity. *Am J Med Genet,* 89**,** 210-7.

GOTTLIEB, B., VASILIOU, D. M., LUMBROSO, R., BEITEL, L. K., PINSKY, L. & TRIFIRO, M. A. 1999b. Analysis of exon 1 mutations in the androgen receptor gene. *Hum Mutat,* 14**,** 527-39.

GRACE, H. J., BERGE, J. E. & OSBORNE, J. 1970. Testicular feminization in a Bantu subject. *S Afr Med J,* 44**,** 19-20.

GRIFFIN, J. E. AND J. D. WILSON (1980). "The syndromes of androgen resistance." N Engl J Med **302**(4): 198-209.

GRINO, P. B., GRIFFIN, J. E., CUSHARD, W. G., JR. & WILSON, J. D. 1988. A mutation of the androgen receptor associated with partial androgen resistance, familial gynecomastia, and fertility. *J Clin Endocrinol Metab,* 66**,** 754-61.

GRONEMEYER, H., GUSTAFSSON, J. A. & LAUDET, V. 2004. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov,* 3**,** 950-64.

GROSSMANN, M. E., HUANG, H. & TINDALL, D. J. 2001. Androgen receptor signaling in androgen-refractory prostate cancer. *J Natl Cancer Inst,* 93**,** 1687-97.

GUPTA, D. 1977. Hypothalamic control of the mammalian sexual maturation. *Padiatr Padol Suppl***,** 83-102.

HAELENS, A., VERRIJDT, G., SCHOENMAKERS, E., ALEN, P., PEETERS, B., ROMBAUTS, W. & CLAESSENS, F. 1999. The first exon of the human sc gene contains an androgen responsive unit and an interferon regulatory factor element. *Mol Cell Endocrinol,* 153**,** 91-102.

HAELENS, A., T. TANNER, S. DENAYER, L. CALLEWAERT AND F. CLAESSENS 2007. "The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor." Cancer Res **67**(9): 4514-4523.

HAIMAN, C. A., BROWN, M., HANKINSON, S. E., SPIEGELMAN, D., COLDITZ, G. A., WILLETT, W. C., KANTOFF, P. W. & HUNTER, D. J. 2002. The androgen receptor CAG repeat polymorphism and risk of breast cancer in the Nurses' Health Study. *Cancer Res,* 62**,** 1045-9.

HAKIMI, J. M., SCHOENBERG, M. P., RONDINELLI, R. H., PIANTADOSI, S. & BARRACK, E. R. 1997. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin Cancer Res,* 3**,** 1599-608.

HAN, S. J., LONARD, D. M. & O'MALLEY, B. W. 2009. Multi-modulation of nuclear receptor coactivators through posttranslational modifications. *Trends Endocrinol Metab,* 20**,** 8-15.

HANNEMA, S. E., SCOTT, I. S., HODAPP, J., MARTIN, H., COLEMAN, N., SCHWABE, J. W. & HUGHES, I. A. 2004. Residual activity of mutant androgen receptors explains wolffian duct development in the complete androgen insensitivity syndrome. *J Clin Endocrinol Metab,* 89**,** 5815-22.

HE, B., MINGES, J. T., LEE, L. W. & WILSON, E. M. 2002. The FXXLF motif mediates androgen receptor-specific interactions with coregulators. *J Biol Chem,* 277**,** 10226-35.

HE, B. & WILSON, E. M. 2002. The NH(2)-terminal and carboxyl-terminal interaction in the human androgen receptor. *Mol Genet Metab,* 75**,** 293-8.

HEEMERS, H. V. & TINDALL, D. J. 2007. Androgen receptor (AR) coregulators: a diversity of functions converging on and regulating the AR transcriptional complex. *Endocr Rev,* 28**,** 778-808.

HEERY, D. M., KALKHOVEN, E., HOARE, S. & PARKER, M. G. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature,* 387**,** 733-6.

HEINLEIN, C. A. & CHANG, C. 2002. Androgen receptor (AR) coregulators: an overview. *Endocr Rev,* 23**,** 175-200.

HEINLEIN, C. A. & CHANG, C. 2004. Androgen receptor in prostate cancer. *Endocr Rev,* 25**,** 276-308.

HIORT, O., SINNECKER, G. H., HOLTERHUS, P. M., NITSCHE, E. M. & KRUSE, K. 1998. Inherited and de novo androgen receptor gene mutations: investigation of single-case families. *J Pediatr,* 132**,** 939-43.

HIRSCHHORN, R. 2003. In vivo reversion to normal of inherited mutations in humans. *J Med Genet,* 40**,** 721-8.

HOLTERHUS, P. M., BRUGGENWIRTH, H. T., HIORT, O., KLEINKAUF-HOUCKEN, A., KRUSE, K., SINNECKER, G. H. & BRINKMANN, A. O. 1997. Mosaicism due to a somatic mutation of the androgen receptor gene determines phenotype in androgen insensitivity syndrome. *J Clin Endocrinol Metab,* 82**,** 3584-9.

HOLTERHUS, P. M., HIORT, O., DEMETER, J., BROWN, P. O. & BROOKS, J. D. 2003. Differential gene-expression patterns in genital fibroblasts of normal males and 46,XY females with androgen insensitivity syndrome: evidence for early programming involving the androgen receptor. *Genome Biol,* 4**,** R37.

HOLTERHUS, P. M., SINNECKER, G. H. & HIORT, O. 2000. Phenotypic diversity and testosterone-induced normalization of mutant L712F androgen receptor function in a kindred with androgen insensitivity. *J Clin Endocrinol Metab,* 85**,** 3245-50.

HOLTERHUS, P. M., WIEBEL, J., SINNECKER, G. H., BRUGGENWIRTH, H. T., SIPPELL, W. G., BRINKMANN, A. O., KRUSE, K. & HIORT, O. 1999. Clinical and molecular spectrum of somatic mosaicism in androgen insensitivity syndrome. *Pediatr Res,* 46**,** 684-90.

HOLTERHUS, P. M., H. T. BRUGGENWIRTH, O. HIORT, A. KLEINKAUF-HOUCKEN, K. KRUSE, G. H. SINNECKER AND A. O. BRINKMANN 1997. "Mosaicism due to a somatic mutation of the androgen receptor gene determines phenotype in androgen insensitivity syndrome." J Clin Endocrinol Metab **82**(11): 3584-3589.

HORWITZ, K. B., JACKSON, T. A., BAIN, D. L., RICHER, J. K., TAKIMOTO, G. S. & TUNG, L. 1996. Nuclear receptor coactivators and corepressors. *Mol Endocrinol,* 10**,** 1167-77.

HOULSTON, R. S. & PETO, J. 2004. The search for low-penetrance cancer susceptibility alleles. *Oncogene,* 23**,** 6471-6.

HSIANG, Y. H., BERKOVITZ, G. D., BROWN, T. R., MIGEON, C. J. & BRODIE, A. M. 1987. The influence of 4-hydroxy-4-androstene-3,17-dione on androgen metabolism and action in cultured human foreskin fibroblasts. *J Steroid Biochem,* 26**,** 131-5.

HSING, A. W., GAO, Y. T., WU, G., WANG, X., DENG, J., CHEN, Y. L., SESTERHENN, I. A., MOSTOFI, F. K., BENICHOU, J. & CHANG, C. 2000. Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case-control study in China. *Cancer Res,* 60**,** 5111-6.

HU, R., DAWOOD, S., HOLMES, M. D., COLLINS, L. C., SCHNITT, S. J., COLE, K., MAROTTI, J. D., HANKINSON, S. E., COLDITZ, G. A. & TAMIMI, R. M. 2011. Androgen receptor expression and breast cancer survival in postmenopausal women. *Clin Cancer Res,* 17**,** 1867-74.

HUDSON, T. S., S. N. PERKINS, S. D. HURSTING, H. A. YOUNG, Y. S. KIM, T. C. WANG AND T. T. WANG 2012. "Inhibition of androgen-responsive LNCaP prostate cancer cell tumor xenograft growth by dietary phenethyl isothiocyanate correlates with decreased angiogenesis and inhibition of cell attachment." Int J Oncol **40**(4): 1113-1121.

HUGHES, I. A. & DEEB, A. 2006. Androgen resistance. *Best Pract Res Clin Endocrinol Metab,* 20**,** 577-98.

HYYTINEN, E. R., HAAPALA, K., THOMPSON, J., LAPPALAINEN, I., ROIHA, M., RANTALA, I., HELIN, H. J., JANNE, O. A., VIHINEN, M., PALVIMO, J. J. & KOIVISTO,

P. A. 2002. Pattern of somatic androgen receptor gene mutations in patients with hormone-refractory prostate cancer. *Lab Invest,* 82**,** 1591-8.

IKEZOE, K., YOSHIMURA, T., TANIWAKI, T., MATSUURA, E., FURUYA, H., YAMADA, T., NAGAMATSU, K. & KIRA, J. 1999. Autosomal dominant familial spinal and bulbar muscular atrophy with gynecomastia. *Neurology,* 53**,** 2187-9.

IKONEN, T., PALVIMO, J. J. & JANNE, O. A. 1997. Interaction between the amino- and carboxyl-terminal regions of the rat androgen receptor modulates transcriptional activity and is influenced by nuclear receptor coactivators. *J Biol Chem,* 272**,** 29821-8.

IRVINE, R. A., YU, M. C., ROSS, R. K. & COETZEE, G. A. 1995. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res,* 55**,** 1937-40.

JAKUBOWSKA, A., GRONWALD, J., MENKISZAK, J., GORSKI, B., HUZARSKI, T., BYRSKI, T., TOLOCZKO-GRABAREK, A., GILBERT, M., EDLER, L., ZAPATKA, M., EILS, R., LUBINSKI, J., SCOTT, R. J. & HAMANN, U. 2010. BRCA1-associated breast and ovarian cancer risks in Poland: no association with commonly studied polymorphisms. *Breast Cancer Res Treat,* 119**,** 201-11.

JANNE, O. A., PALVIMO, J. J., KALLIO, P. & MEHTO, M. 1993. Androgen receptor and mechanism of androgen action. *Ann Med,* 25**,** 83-9.

JANNE, O. A. & SHAN, L. X. 1991. Structure and function of the androgen receptor. *Ann N Y Acad Sci,* 626**,** 81-91.

JENSTER, G., VAN DER KORPUT, H. A., TRAPMAN, J. & BRINKMANN, A. O. 1995. Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *J Biol Chem,* 270**,** 7341-6.

JENSTER, G., VAN DER KORPUT, H. A., VAN VROONHOVEN, C., VAN DER KWAST, T. H., TRAPMAN, J. & BRINKMANN, A. O. 1991. Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization. *Mol Endocrinol,* 5**,** 1396-404.

JOHN, H. AND C. SCHMID 2000. "Kallmann's syndrome: clues to clinical diagnosis." Int J Impot Res **12**(5): 269-271.

KAAKS, R., BERRINO, F., KEY, T., RINALDI, S., DOSSUS, L., BIESSY, C., SECRETO, G., AMIANO, P., BINGHAM, S., BOEING, H., BUENO DE MESQUITA, H. B., CHANG-CLAUDE, J., CLAVEL-CHAPELON, F., FOURNIER, A., VAN GILS, C. H., GONZALEZ, C. A., GURREA, A. B., CRITSELIS, E., KHAW, K. T., KROGH, V., LAHMANN, P. H., NAGEL, G., OLSEN, A., ONLAND-MORET, N. C., OVERVAD, K., PALLI, D., PANICO, S., PEETERS, P., QUIROS, J. R., RODDAM, A., THIEBAUT, A., TJONNELAND, A., CHIRLAQUE, M. D., TRICHOPOULOU, A., TRICHOPOULOS, D., TUMINO, R., VINEIS, P., NORAT, T., FERRARI, P., SLIMANI, N. & RIBOLI, E. 2005a. Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC). *J Natl Cancer Inst,* 97**,** 755-65.

KAAKS, R., RINALDI, S., KEY, T. J., BERRINO, F., PEETERS, P. H., BIESSY, C., DOSSUS, L., LUKANOVA, A., BINGHAM, S., KHAW, K. T., ALLEN, N. E., BUENO-DE-

MESQUITA, H. B., VAN GILS, C. H., GROBBEE, D., BOEING, H., LAHMANN, P. H., NAGEL, G., CHANG-CLAUDE, J., CLAVEL-CHAPELON, F., FOURNIER, A., THIEBAUT, A., GONZALEZ, C. A., QUIROS, J. R., TORMO, M. J., ARDANAZ, E., AMIANO, P., KROGH, V., PALLI, D., PANICO, S., TUMINO, R., VINEIS, P., TRICHOPOULOU, A., KALAPOTHAKI, V., TRICHOPOULOS, D., FERRARI, P., NORAT, T., SARACCI, R. & RIBOLI, E. 2005b. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. *Endocr Relat Cancer,* 12**,** 1071-82.

KADOURI, L., EASTON, D. F., EDWARDS, S., HUBERT, A., KOTE-JARAI, Z., GLASER, B., DUROCHER, F., ABELIOVICH, D., PERETZ, T. & EELES, R. A. 2001. CAG and GGC repeat polymorphisms in the androgen receptor gene and breast cancer susceptibility in BRCA1/2 carriers and non-carriers. *Br J Cancer,* 85**,** 36-40.

KAMISCHKE, A., BAUMGARDT, A., HORST, J. & NIESCHLAG, E. 2003. Clinical and diagnostic features of patients with suspected Klinefelter syndrome. *J Androl,* 24**,** 41-8.

KATAYAMA, Y., TRAN, V. K., HOAN, N. T., ZHANG, Z., GOJI, K., YAGI, M., TAKESHIMA, Y., SAIKI, K., NHAN, N. T. & MATSUO, M. 2006. Co-occurrence of mutations in both dystrophin- and androgen-receptor genes is a novel cause of female Duchenne muscular dystrophy. *Hum Genet,* 119**,** 516-9.

KAZEMI-ESFARJANI, P., TRIFIRO, M. A. & PINSKY, L. 1995. Evidence for a repressive function of the long polyglutamine tract in the human androgen

receptor: possible pathogenetic relevance for the (CAG)n-expanded neuronopathies. *Hum Mol Genet,* 4**,** 523-7.

KEENAN, B. S., MEYER, W. J., 3RD, HADJIAN, A. J., JONES, H. W. & MIGEON, C. J. 1974. Syndrome of androgen insensitivity in man: absence of 5 alpha-dihydrotestosterone binding protein in skin fibroblasts. *J Clin Endocrinol Metab,* 38**,** 1143-6.

KELLER, E. T., ERSHLER, W. B. & CHANG, C. 1996. The androgen receptor: a mediator of diverse responses. *Front Biosci,* 1**,** d59-71.

KICMAN, A. T. 2010. Biochemical and physiological aspects of endogenous androgens. *Handb Exp Pharmacol***,** 25-64.

KIDWAI, N., Y. GONG, X. SUN, C. G. DESHPANDE, A. V. YELDANDI, M. S. RAO AND S. BADVE 2004. "Expression of androgen receptor and prostate-specific antigen in male breast carcinoma." Breast Cancer Res **6**(1): R18-23.

KITTLES, R. A., YOUNG, D., WEINRICH, S., HUDSON, J., ARGYROPOULOS, G., UKOLI, F., ADAMS-CAMPBELL, L. & DUNSTON, G. M. 2001. Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: implications for prostate cancer risk. *Hum Genet,* 109**,** 253-61.

KO, T. M., Y. S. YANG, M. Y. WU, C. H. KAO, P. M. HSU, S. M. CHUANG AND T. Y. LEE 1997. "Complete androgen insensitivity syndrome. Molecular characterization in two Chinese women." J Reprod Med **42**(7): 424-428.

KOMORI, S., KASUMI, H., SAKATA, K. & KOYAMA, K. 2007. The role of androgens in spermatogenesis. *Soc Reprod Fertil Suppl,* 63**,** 25-30.

KRISTIANSEN, M., LANGEROD, A., KNUDSEN, G. P., WEBER, B. L., BORRESEN-DALE, A. L. & ORSTAVIK, K. H. 2002. High frequency of skewed X inactivation in young breast cancer patients. *J Med Genet,* 39**,** 30-3.

KRSMANOVIC, L. Z., HU, L., LEUNG, P. K., FENG, H. & CATT, K. J. 2009. The hypothalamic GnRH pulse generator: multiple regulatory mechanisms. *Trends Endocrinol Metab,* 20**,** 402-8.

KUENEN-BOUMEESTER, V., VAN DER KWAST, T. H., CLAASSEN, C. C., LOOK, M. P., LIEM, G. S., KLIJN, J. G. & HENZEN-LOGMANS, S. C. 1996. The clinical significance of androgen receptors in breast cancer and their relation to histological and cell biological parameters. *Eur J Cancer,* 32A**,** 1560-5.

KUIPER, G. G., FABER, P. W., VAN ROOIJ, H. C., VAN DER KORPUT, J. A., RIS-STALPERS, C., KLAASSEN, P., TRAPMAN, J. & BRINKMANN, A. O. 1989. Structural organization of the human androgen receptor gene. *J Mol Endocrinol,* 2**,** R1-4.

LABRIE, F., LUU-THE, V., LABRIE, C., BELANGER, A., SIMARD, J., LIN, S. X. & PELLETIER, G. 2003. Endocrine and intracrine sources of androgens in women: inhibition of breast cancer and other roles of androgens and their precursor dehydroepiandrosterone. *Endocr Rev,* 24**,** 152-82.

LAMONT, K. R. & TINDALL, D. J. 2010. Androgen regulation of gene expression. *Adv Cancer Res,* 107**,** 137-62.

LANGLEY, E., ZHOU, Z. X. & WILSON, E. M. 1995. Evidence for an anti-parallel orientation of the ligand-activated human androgen receptor dimer. *J Biol Chem,* 270**,** 29983-90.

LAREYRE, J. J., CLAESSENS, F., ROMBAUTS, W., DUFAURE, J. P. & DREVET, J. R. 1997. Characterization of an androgen response element within the promoter of the epididymis-specific murine glutathione peroxidase 5 gene. *Mol Cell Endocrinol,* 129**,** 33-46.

LATCHMAN, D. S. 1991. Heat-shock proteins. *Lancet,* 337**,** 424.

LEE, H. J. & CHANG, C. 2003. Recent advances in androgen receptor action. *Cell Mol Life Sci,* 60**,** 1613-22.

LI, J. & AL-AZZAWI, F. 2009. Mechanism of androgen receptor action. *Maturitas,* 63**,** 142-8.

LIAO, G., CHEN, L. Y., ZHANG, A., GODAVARTHY, A., XIA, F., GHOSH, J. C., LI, H. & CHEN, J. D. 2003. Regulation of androgen receptor activity by the nuclear receptor corepressor SMRT. *J Biol Chem,* 278**,** 5052-61.

LIEDE, A., ZHANG, W., DE LEON MATSUDA, M. L., TAN, A. & NAROD, S. A. 2003. Androgen receptor gene polymorphism and breast cancer susceptibility in The Philippines. *Cancer Epidemiol Biomarkers Prev,* 12**,** 848-52.

LILLIE, E. O., BERNSTEIN, L. & URSIN, G. 2003. The role of androgens and polymorphisms in the androgen receptor in the epidemiology of breast cancer. *Breast Cancer Res,* 5**,** 164-73.

LINDQUIST, S. 1986. The heat-shock response. *Annu Rev Biochem,* 55**,** 1151-91.

LINDQUIST, S. & CRAIG, E. A. 1988. The heat-shock proteins. *Annu Rev Genet,* 22**,** 631-77.

LINDZEY, J., KUMAR, M. V., GROSSMAN, M., YOUNG, C. & TINDALL, D. J. 1994. Molecular mechanisms of androgen action. *Vitam Horm,* 49**,** 383-432.

LIU, J., TSANG, S. & WONG, T. M. 2006. Testosterone is required for delayed cardioprotection and enhanced heat shock protein 70 expression induced by preconditioning. *Endocrinology,* 147**,** 4569-77.

LOBACCARO, J. M., LUMBROSO, S., BELON, C., GALTIER-DEREURE, F., BRINGER, J., LESIMPLE, T., NAMER, M., CUTULI, B. F., PUJOL, H. & SULTAN, C. 1993. Androgen receptor gene mutation in male breast cancer. *Hum Mol Genet,* 2**,** 1799-802.

LORENI, F., STAVENHAGEN, J., KALFF, M. & ROBINS, D. M. 1988. A complex androgen-responsive enhancer resides 2 kilobases upstream of the mouse Slp gene. *Mol Cell Biol,* 8**,** 2350-60.

LUBAHN, D. B., BROWN, T. R., SIMENTAL, J. A., HIGGS, H. N., MIGEON, C. J., WILSON, E. M. & FRENCH, F. S. 1989. 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. *Proc Natl Acad Sci U S A,* 86**,** 9534-8.

LUBAHN, D. B., JOSEPH, D. R., SAR, M., TAN, J., HIGGS, H. N., LARSON, R. E., FRENCH, F. S. & WILSON, E. M. 1988a. The human androgen receptor:

complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate. *Mol Endocrinol,* 2**,** 1265-75.

LUBAHN, D. B., JOSEPH, D. R., SULLIVAN, P. M., WILLARD, H. F., FRENCH, F. S. & WILSON, E. M. 1988b. Cloning of human androgen receptor complementary DNA and localization to the X chromosome. *Science,* 240**,** 327-30.

LUNDIN, K. B., GIWERCMAN, Y. L., RYLANDER, L., HAGMAR, L. & GIWERCMAN, A. 2006. Androgen receptor gene GGN repeat length and reproductive characteristics in young Swedish men. *Eur J Endocrinol,* 155**,** 347-54.

LUO, F. W., WU, W. Q., GENG, Q., LI, F., CHEN, W. B., GAN, W. X. & XIE, J. S. 2011. [A novel mutation Glu441stop (GAA to TAA) of androgen receptor gene resulting in complete androgen insensitivity syndrome]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi,* 28**,** 176-9.

MA, Z. Q., LIU, Z., NGAN, E. S. & TSAI, S. Y. 2001. Cdc25B functions as a novel coactivator for the steroid receptors. *Mol Cell Biol,* 21**,** 8056-67.

MACLEAN, H. E., WARNE, G. L. & ZAJAC, J. D. 1997. Localization of functional domains in the androgen receptor. *J Steroid Biochem Mol Biol,* 62**,** 233-42.

MANJILI, M. H., K. NAJARIAN AND X. Y. WANG 2012. "Signatures of tumor-immune interactions as biomarkers for breast cancer prognosis." Future Oncol **8**(6): 703-711.

MATIAS, P. M., DONNER, P., COELHO, R., THOMAZ, M., PEIXOTO, C., MACEDO, S., OTTO, N., JOSCHKO, S., SCHOLZ, P., WEGG, A., BASLER, S., SCHAFER, M., EGNER, U. & CARRONDO, M. A. 2000. Structural evidence for ligand specificity in the

binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *J Biol Chem,* 275**,** 26164-71.

MATZKIN, H. & SOLOWAY, M. S. 1992. Immunohistochemical evidence of the existence and localization of aromatase in human prostatic tissues. *Prostate,* 21**,** 309-14.

MAYER, M. P. & BUKAU, B. 2005. Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci,* 62**,** 670-84.

MAZEN, I., EL-RUBY, M., KAMAL, R., EL-NEKHELY, I., EL-GHANDOUR, M., TANTAWY, S. & EL-GAMMAL, M. 2010. Screening of genital anomalies in newborns and infants in two egyptian governorates. *Horm Res Paediatr,* 73**,** 438-42.

MCEWAN, I. J. 2004. Molecular mechanisms of androgen receptor-mediated gene regulation: structure-function analysis of the AF-1 domain. *Endocr Relat Cancer,* 11**,** 281-93.

MCKENNA, N. J., XU, J., NAWAZ, Z., TSAI, S. Y., TSAI, M. J. & O'MALLEY, B. W. 1999. Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. *J Steroid Biochem Mol Biol,* 69**,** 3-12.

MCKUSICK, V. A. 1985. Marcella O'Grady Boveri (1865-1950) and the chromosome theory of cancer. *J Med Genet,* 22**,** 431-40.

MCPHAUL, M. J., MARCELLI, M., TILLEY, W. D., GRIFFIN, J. E. & WILSON, J. D. 1991. Androgen resistance caused by mutations in the androgen receptor gene. *FASEB J,* 5**,** 2910-5.

MELO, K. F., MENDONCA, B. B., BILLERBECK, A. E., COSTA, E. M., INACIO, M., SILVA, F. A., LEAL, A. M., LATRONICO, A. C. & ARNHOLD, I. J. 2003. Clinical, hormonal, behavioral, and genetic characteristics of androgen insensitivity syndrome in a Brazilian cohort: five novel mutations in the androgen receptor gene. *J Clin Endocrinol Metab,* 88**,** 3241-50.

MENIN, C., BANNA, G. L., DE SALVO, G., LAZZAROTTO, V., DE NICOLO, A., AGATA, S., MONTAGNA, M., SORDI, G., NICOLETTO, O., CHIECO-BIANCHI, L. & D'ANDREA, E. 2001. Lack of association between androgen receptor CAG polymorphism and familial breast/ovarian cancer. *Cancer Lett,* 168**,** 31-6.

MIGEON, C. J., BROWN, T. R., LANES, R., PALACIOS, A., AMRHEIN, J. A. & SCHOEN, E. J. 1984. A clinical syndrome of mild androgen insensitivity. *J Clin Endocrinol Metab,* 59**,** 672-8.

MILLER, S. A., DYKES, D. D. & POLESKY, H. F. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res,* 16**,** 1215.

MOINFAR, F., OKCU, M., TSYBROVSKYY, O., REGITNIG, P., LAX, S. F., WEYBORA, W., RATSCHEK, M., TAVASSOLI, F. A. & DENK, H. 2003. Androgen receptors frequently are expressed in breast carcinomas: potential relevance to new therapeutic strategies. *Cancer,* 98**,** 703-11.

MONTGOMERY, J. S., PRICE, D. K. & FIGG, W. D. 2001. The androgen receptor gene and its influence on the development and progression of prostate cancer. *J Pathol,* 195**,** 138-46.

MOORADIAN, A. D., MORLEY, J. E. & KORENMAN, S. G. 1987. Biological actions of androgens. *Endocr Rev,* 8**,** 1-28.

MORALES, A., HEATON, J. P. & CARSON, C. C., 3RD 2000. Andropause: a misnomer for a true clinical entity. *J Urol,* 163**,** 705-12.

MORRIS, J. M. 1953. The syndrome of testicular feminization in male pseudohermaphrodites. *Am J Obstet Gynecol,* 65**,** 1192-1211.

MURONO, K., MENDONCA, B. B., ARNHOLD, I. J., RIGON, A. C., MIGEON, C. J. & BROWN, T. R. 1995. Human androgen insensitivity due to point mutations encoding amino acid substitutions in the androgen receptor steroid-binding domain. *Hum Mutat,* 6**,** 152-62.

NARULA, H. S. AND H. E. CARLSON 2014. "Gynaecomastia-pathophysiology, diagnosis and treatment." Nat Rev Endocrinol.

NESSLER-MENARDI, C., JOTOVA, I., CULIG, Z., EDER, I. E., PUTZ, T., BARTSCH, G. & KLOCKER, H. 2000. Expression of androgen receptor coregulatory proteins in prostate cancer and stromal-cell culture models. *Prostate,* 45**,** 124-31.

NICOLAS DIAZ-CHICO, B., GERMAN RODRIGUEZ, F., GONZALEZ, A., RAMIREZ, R., BILBAO, C., CABRERA DE LEON, A., AGUIRRE JAIME, A., CHIRINO, R., NAVARRO, D. & DIAZ-CHICO, J. C. 2007. Androgens and androgen receptors in breast cancer. *J Steroid Biochem Mol Biol,* 105**,** 1-15.

NIESCHLAG, E. 1997. Andrology at the end of the twentieth century: from spermatology to male reproductive health. Inaugural Address at the VIth

International Congress of Andrology, Salzburg, 25 May 1997. *Int J Androl,* 20**,** 129-31.

O'CONNOR, A. E. & DE KRETSER, D. M. 2004. Inhibins in normal male physiology. *Semin Reprod Med,* 22**,** 177-85.

O'MALLEY, B. W. & KUMAR, R. 2009. Nuclear receptor coregulators in cancer biology. *Cancer Res,* 69**,** 8217-22.

O'MALLEY, B. W., QIN, J. & LANZ, R. B. 2008. Cracking the coregulator codes. *Curr Opin Cell Biol,* 20**,** 310-5.

OZANNE, D. M., BRADY, M. E., COOK, S., GAUGHAN, L., NEAL, D. E. & ROBSON, C. N. 2000. Androgen receptor nuclear translocation is facilitated by the f-actin cross-linking protein filamin. *Mol Endocrinol,* 14**,** 1618-26.

PAJONK, F., A. VAN OPHOVEN AND W. H. MCBRIDE 2005. "Hyperthermia-induced proteasome inhibition and loss of androgen receptor expression in human prostate cancer cells." Cancer Res **65**(11): 4836-4843.

PERKINS, P., COOKSLEY, C. D. & COX, J. D. 1996. Breast cancer. Is ethnicity an independent prognostic factor for survival? *Cancer,* 78**,** 1241-7.

PHILIBERT, P., AUDRAN, F., PIENKOWSKI, C., MORANGE, I., KOHLER, B., FLORI, E., HEINRICH, C., DACOU-VOUTETAKIS, C., JOSEPH, M. G., GUEDJ, A. M., JOURNEL, H., HECART-BRUNA, A. C., KHOTCHALI, I., TEN, S., BOUCHARD, P., PARIS, F. & SULTAN, C. 2010. Complete androgen insensitivity syndrome is frequently due to premature stop codons in exon 1 of the androgen receptor gene: an international collaborative report of 13 new mutations. *Fertil Steril,* 94**,** 472-6.

PINSKY, L., KAUFMAN, M. & KILLINGER, D. W. 1989. Impaired spermatogenesis is not an obligate expression of receptor-defective androgen resistance. *Am J Med Genet,* 32**,** 100-4.

PINSKY, L., KAUFMAN, M., KILLINGER, D. W., BURKO, B., SHATZ, D. & VOLPE, R. 1984. Human minimal androgen insensitivity with normal dihydrotestosterone-binding capacity in cultured genital skin fibroblasts: evidence for an androgen-selective qualitative abnormality of the receptor. *Am J Hum Genet,* 36**,** 965-78.

PLATZ, E. A., GIOVANNUCCI, E., DAHL, D. M., KRITHIVAS, K., HENNEKENS, C. H., BROWN, M., STAMPFER, M. J. & KANTOFF, P. W. 1998. The androgen receptor gene GGN microsatellite and prostate cancer risk. *Cancer Epidemiol Biomarkers Prev,* 7**,** 379-84.

QUIGLEY, C. A., DE BELLIS, A., MARSCHKE, K. B., EL-AWADY, M. K., WILSON, E. M. & FRENCH, F. S. 1995. Androgen receptor defects: historical, clinical, and molecular perspectives. *Endocr Rev,* 16**,** 271-321.

QUIGLEY, C. A., EVANS, B. A., SIMENTAL, J. A., MARSCHKE, K. B., SAR, M., LUBAHN, D. B., DAVIES, P., HUGHES, I. A., WILSON, E. M. & FRENCH, F. S. 1992. Complete androgen insensitivity due to deletion of exon C of the androgen receptor gene highlights the functional importance of the second zinc finger of the androgen receptor in vivo. *Mol Endocrinol,* 6**,** 1103-12.

RAJENDER, S., SINGH, L. & THANGARAJ, K. 2007a. L859F mutation in androgen receptor gene results in complete loss of androgen binding to the receptor. *J Androl,* 28**,** 772-6.

RAJENDER, S., SINGH, L. & THANGARAJ, K. 2007b. Phenotypic heterogeneity of mutations in androgen receptor gene. *Asian J Androl,* 9**,** 147-79.

REBBECK, T. R., KANTOFF, P. W., KRITHIVAS, K., NEUHAUSEN, S., BLACKWOOD, M. A., GODWIN, A. K., DALY, M. B., NAROD, S. A., GARBER, J. E., LYNCH, H. T., WEBER, B. L. & BROWN, M. 1999. Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. *Am J Hum Genet,* 64**,** 1371-7.

RENNIE, P. S., BRUCHOVSKY, N., LECO, K. J., SHEPPARD, P. C., MCQUEEN, S. A., CHENG, H., SNOEK, R., HAMEL, A., BOCK, M. E., MACDONALD, B. S. & ET AL. 1993. Characterization of two cis-acting DNA elements involved in the androgen regulation of the probasin gene. *Mol Endocrinol,* 7**,** 23-36.

RIEGMAN, P. H., VLIETSTRA, R. J., VAN DER KORPUT, J. A., BRINKMANN, A. O. & TRAPMAN, J. 1991. The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. *Mol Endocrinol,* 5**,** 1921-30.

ROBINS, D. M. (005. "Androgen receptor and molecular mechanisms of male-specific gene expression." Novartis Found Symp **268**: 42-52; discussion 53-46, 96-49.

ROBYR, D., WOLFFE, A. P. & WAHLI, W. 2000. Nuclear hormone receptor coregulators in action: diversity for shared tasks. *Mol Endocrinol,* 14**,** 329-47.

RODRIGUEZ, G., BILBAO, C., RAMIREZ, R., FALCON, O., LEON, L., CHIRINO, R., FALCON, O., JR., DIAZ, B. P., RIVERO, J. F., PERUCHO, M., DIAZ-CHICO, B. N. & DIAZ-CHICO, J. C. 2006. Alleles with short CAG and GGN repeats in the androgen

receptor gene are associated with benign endometrial cancer. *Int J Cancer,* 118**,** 1420-5.

ROYCHOUDHURY, A., BASU, S. & SENGUPTA, D. N. 2009. Analysis of comparative efficiencies of different transformation methods of E. coli using two common plasmid vectors. *Indian J Biochem Biophys,* 46**,** 395-400.

SACK, J. S., KISH, K. F., WANG, C., ATTAR, R. M., KIEFER, S. E., AN, Y., WU, G. Y., SCHEFFLER, J. E., SALVATI, M. E., KRYSTEK, S. R., JR., WEINMANN, R. & EINSPAHR, H. M. 2001. Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proc Natl Acad Sci U S A,* 98**,** 4904-9.

SARPEL, U., PALMER, S. K. & DOLGIN, S. E. 2005. The incidence of complete androgen insensitivity in girls with inguinal hernias and assessment of screening by vaginal length measurement. *J Pediatr Surg,* 40**,** 133-6; discussion 136-7.

SASAKI, M., KANEUCHI, M., SAKURAGI, N., FUJIMOTO, S., CARROLL, P. R. & DAHIYA, R. 2003. The polyglycine and polyglutamine repeats in the androgen receptor gene in Japanese and Caucasian populations. *Biochem Biophys Res Commun,* 312**,** 1244-7.

SECRETO, G. AND B. ZUMOFF 2012. "Role of androgen excess in the development of estrogen receptor-positive and estrogen receptor-negative breast cancer." Anticancer Res **32**(8): 3223-3228.

SCHILDKRAUT, J. M., MURPHY, S. K., PALMIERI, R. T., IVERSEN, E., MOORMAN, P. G., HUANG, Z., HALABI, S., CALINGAERT, B., GUSBERG, A., MARKS, J. R. &

BERCHUCK, A. 2007. Trinucleotide repeat polymorphisms in the androgen receptor gene and risk of ovarian cancer. *Cancer Epidemiol Biomarkers Prev,* 16**,** 473-80.

SEGAWA, N., NAKAMURA, M., SHAN, L., UTSUNOMIYA, H., NAKAMURA, Y., MORI, I., KATSUOKA, Y. & KAKUDO, K. 2002. Expression and somatic mutation on androgen receptor gene in prostate cancer. *Int J Urol,* 9**,** 545-53.

SHANG, Y., MYERS, M. & BROWN, M. 2002. Formation of the androgen receptor transcription complex. *Mol Cell,* 9**,** 601-10.

SHATKINA, L., MINK, S., ROGATSCH, H., KLOCKER, H., LANGER, G., NESTL, A. & CATO, A. C. 2003. The cochaperone Bag-1L enhances androgen receptor action via interaction with the NH2-terminal region of the receptor. *Mol Cell Biol,* 23**,** 7189-97.

SHECKTER, C. B., MATSUMOTO, A. M. & BREMNER, W. J. 1989. Testosterone administration inhibits gonadotropin secretion by an effect directly on the human pituitary. *J Clin Endocrinol Metab,* 68**,** 397-401.

SHIINA, H., MATSUMOTO, T., SATO, T., IGARASHI, K., MIYAMOTO, J., TAKEMASA, S., SAKARI, M., TAKADA, I., NAKAMURA, T., METZGER, D., CHAMBON, P., KANNO, J., YOSHIKAWA, H. & KATO, S. 2006. Premature ovarian failure in androgen receptor-deficient mice. *Proc Natl Acad Sci U S A,* 103**,** 224-9.

SIMENTAL, J. A., SAR, M., LANE, M. V., FRENCH, F. S. & WILSON, E. M. 1991. Transcriptional activation and nuclear targeting signals of the human androgen receptor. *J Biol Chem,* 266**,** 510-8.

SINGH, R., ARTAZA, J. N., TAYLOR, W. E., BRAGA, M., YUAN, X., GONZALEZ-CADAVID, N. F. & BHASIN, S. 2006. Testosterone inhibits adipogenic differentiation in 3T3-L1 cells: nuclear translocation of androgen receptor complex with beta-catenin and T-cell factor 4 may bypass canonical Wnt signaling to down-regulate adipogenic transcription factors. *Endocrinology,* 147**,** 141-54.

SLATTERY, M. L., SWEENEY, C., HERRICK, J., WOLFF, R., BAUMGARTNER, K., GIULIANO, A. & BYERS, T. 2007. ESR1, AR, body size, and breast cancer risk in Hispanic and non-Hispanic white women living in the Southwestern United States. *Breast Cancer Res Treat,* 105**,** 327-35.

SPEROFF, L. 2003. "The Million Women Study and breast cancer." Maturitas **46**(1): 1-6.

STEGEMAN, I. AND P. M. BOSSUYT 2012. "Cancer risk models and preselection for screening." Cancer Epidemiol **36**(5): 461-469.

SPURDLE, A. B., ANTONIOU, A. C., DUFFY, D. L., PANDEYA, N., KELEMEN, L., CHEN, X., PEOCK, S., COOK, M. R., SMITH, P. L., PURDIE, D. M., NEWMAN, B., DITE, G. S., APICELLA, C., SOUTHEY, M. C., GILES, G. G., HOPPER, J. L., CHENEVIX-TRENCH, G., EASTON, D. F. & COLLABORATORS, E. S. 2005. The androgen receptor CAG repeat polymorphism and modification of breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res,* 7**,** R176-83.

SPURDLE, A. B., DITE, G. S., CHEN, X., MAYNE, C. J., SOUTHEY, M. C., BATTEN, L. E., CHY, H., TRUTE, L., MCCREDIE, M. R., GILES, G. G., ARMES, J., VENTER, D. J., HOPPER, J. L. & CHENEVIX-TRENCH, G. 1999. Androgen receptor exon 1 CAG repeat length and breast cancer in women before age forty years. *J Natl Cancer Inst,* 91**,** 961-6.

SPURDLE, A. B., WEBB, P. M., CHEN, X., MARTIN, N. G., GILES, G. G., HOPPER, J. L. & CHENEVIX-TRENCH, G. 2000. Androgen receptor exon 1 CAG repeat length and risk of ovarian cancer. *Int J Cancer,* 87**,** 637-43.

SRIVASTAVA, P. 2004. Heat shock proteins and immune response: methods to madness. *Methods,* 32**,** 1-2.

STANFORD, J. L., JUST, J. J., GIBBS, M., WICKLUND, K. G., NEAL, C. L., BLUMENSTEIN, B. A. & OSTRANDER, E. A. 1997. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res,* 57**,** 1194-8.

STEINKAMP, M. P., O'MAHONY, O. A., BROGLEY, M., REHMAN, H., LAPENSEE, E. W., DHANASEKARAN, S., HOFER, M. D., KUEFER, R., CHINNAIYAN, A., RUBIN, M. A., PIENTA, K. J. & ROBINS, D. M. 2009. Treatment-dependent androgen receptor mutations in prostate cancer exploit multiple mechanisms to evade therapy. *Cancer Res,* 69**,** 4434-42.

STEKETEE, K., BERREVOETS, C. A., DUBBINK, H. J., DOESBURG, P., HERSMUS, R., BRINKMANN, A. O. & TRAPMAN, J. 2002. 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**,** 5780-91.

SULTAN, C., PARIS, F., JEANDEL, C., LUMBROSO, S. & GALIFER, R. B. 2002. Ambiguous genitalia in the newborn. *Semin Reprod Med,* 20**,** 181-8.

SUN, S., LUO, F., ZHOU, Z. & WU, W. 2010. A novel androgen receptor gene mutation in a Chinese patient with complete androgen insensitivity syndrome. *Eur J Obstet Gynecol Reprod Biol,* 153**,** 173-5.

SUTER, N. M., MALONE, K. E., DALING, J. R., DOODY, D. R. & OSTRANDER, E. A. 2003. Androgen receptor (CAG)n and (GGC)n polymorphisms and breast cancer risk in a population-based case-control study of young women. *Cancer Epidemiol Biomarkers Prev,* 12**,** 127-35.

SZILARD, L. 1959. On the Nature of the Aging Process. *Proc Natl Acad Sci U S A,* 45**,** 30-45.

THOMSON, R. L., G. D. BRINKWORTH, M. NOAKES, P. M. CLIFTON, R. J. NORMAN AND J. D. BUCKLEY 2012. "The effect of diet and exercise on markers of endothelial function in overweight and obese women with polycystic ovary syndrome." Hum Reprod **27**(7): 2169-2176.

TIEFENBACHER, K. & DAXENBICHLER, G. 2008. The Role of Androgens in Normal and Malignant Breast Tissue. *Breast Care (Basel),* 3**,** 325-331.

TRIEZENBERG, S. J. 1995. Structure and function of transcriptional activation domains. *Curr Opin Genet Dev,* 5**,** 190-6.

TSEZOU, A., TZETIS, M., GENNATAS, C., GIANNATOU, E., PAMPANOS, A., MALAMIS, G., KANAVAKIS, E. & KITSIOU, S. 2008. Association of repeat polymorphisms in the estrogen receptors alpha, beta (ESR1, ESR2) and androgen receptor (AR) genes with the occurrence of breast cancer. *Breast,* 17**,** 159-66.

TSUGAYA, M., HARADA, N., TOZAWA, K., YAMADA, Y., HAYASHI, Y., TANAKA, S., MARUYAMA, K. & KOHRI, K. 1996. Aromatase mRNA levels in benign prostatic hyperplasia and prostate cancer. *Int J Urol,* 3**,** 292-6.

TSUJIMOTO, Y., TAKAKUWA, T., TAKAYAMA, H., NISHIMURA, K., OKUYAMA, A., AOZASA, K. & NONOMURA, N. 2004. In situ shortening of CAG repeat length within the androgen receptor gene in prostatic cancer and its possible precursors. *Prostate,* 58**,** 283-90.

TUT, T. G., GHADESSY, F. J., TRIFIRO, M. A., PINSKY, L. & YONG, E. L. 1997. Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. *J Clin Endocrinol Metab,* 82**,** 3777-82.

UMAR, A., C. A. BERREVOETS, N. M. VAN, M. VAN LEEUWEN, M. VERBIEST, W. J. KLEIJER, D. DOOIJES, J. A. GROOTEGOED, S. L. DROP AND A. O. BRINKMANN 2005. "Functional analysis of a novel androgen receptor mutation, Q902K, in an individual with partial androgen insensitivity." J Clin Endocrinol Metab **90**(1): 507-515.

VAN DE WIJNGAART, D. J., VAN ROYEN, M. E., HERSMUS, R., PIKE, A. C., HOUTSMULLER, A. B., JENSTER, G., TRAPMAN, J. & DUBBINK, H. J. 2006. Novel

FXXFF and FXXMF motifs in androgen receptor cofactors mediate high affinity and specific interactions with the ligand-binding domain. *J Biol Chem,* 281**,** 19407-16.

VELDSCHOLTE, J., BERREVOETS, C. A., BRINKMANN, A. O., GROOTEGOED, J. A. & MULDER, E. 1992. Anti-androgens and the mutated androgen receptor of LNCaP cells: differential effects on binding affinity, heat-shock protein interaction, and transcription activation. *Biochemistry,* 31**,** 2393-9.

VERRIJDT, G., HAELENS, A. & CLAESSENS, F. 2003. Selective DNA recognition by the androgen receptor as a mechanism for hormone-specific regulation of gene expression. *Mol Genet Metab,* 78**,** 175-85.

VERRIJDT, G., SCHOENMAKERS, E., ALEN, P., HAELENS, A., PEETERS, B., ROMBAUTS, W. & CLAESSENS, F. 1999. Androgen specificity of a response unit upstream of the human secretory component gene is mediated by differential receptor binding to an essential androgen response element. *Mol Endocrinol,* 13**,** 1558-70.

VINCENT, A., HEITZ, D., PETIT, C., KRETZ, C., OBERLE, I. & MANDEL, J. L. 1991. Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis. *Nature,* 349**,** 624-6.

VINER, R. M., TEOH, Y., WILLIAMS, D. M., PATTERSON, M. N. & HUGHES, I. A. 1997. Androgen insensitivity syndrome: a survey of diagnostic procedures and management in the UK. *Arch Dis Child,* 77**,** 305-9.

VOEGEL, J. J., HEINE, M. J., ZECHEL, C., CHAMBON, P. & GRONEMEYER, H. 1996. TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. *EMBO J,* 15**,** 3667-75.

WADLER, B. M., JUDGE, C. M., PROUT, M., ALLEN, J. D. & GELLER, A. C. 2011. Improving Breast Cancer Control via the Use of Community Health Workers in South Africa: A Critical Review. *J Oncol,* 2011.

WANG, L., HSU, C. L. & CHANG, C. 2005a. Androgen receptor corepressors: an overview. *Prostate,* 63**,** 117-30.

WANG, W., JOHN, E. M. & INGLES, S. A. 2005b. Androgen receptor and prostate-specific antigen gene polymorphisms and breast cancer in African-American women. *Cancer Epidemiol Biomarkers Prev,* 14**,** 2990-4.

WHITE, P. C. & SPEISER, P. W. 2000. Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Endocr Rev,* 21**,** 245-91.

WIENER, J. S., TEAGUE, J. L., ROTH, D. R., GONZALES, E. T., JR. & LAMB, D. J. 1997. Molecular biology and function of the androgen receptor in genital development. *J Urol,* 157**,** 1377-86.

WONG, C. I., ZHOU, Z. X., SAR, M. & WILSON, E. M. 1993. Steroid requirement for androgen receptor dimerization and DNA binding. Modulation by intramolecular interactions between the NH2-terminal and steroid-binding domains. *J Biol Chem,* 268**,** 19004-12.

WOOSTER, R., MANGION, J., EELES, R., SMITH, S., DOWSETT, M., AVERILL, D., BARRETT-LEE, P., EASTON, D. F., PONDER, B. A. & STRATTON, M. R. 1992. A

germline mutation in the androgen receptor gene in two brothers with breast cancer and Reifenstein syndrome. *Nat Genet,* 2**,** 132-4.

WRIGHT, A. S., DOUGLAS, R. C., THOMAS, L. N., LAZIER, C. B. & RITTMASTER, R. S. 1999. Androgen-induced regrowth in the castrated rat ventral prostate: role of 5alpha-reductase. *Endocrinology,* 140**,** 4509-15.

WU, C. 1995. Heat shock transcription factors: structure and regulation. *Annu Rev Cell Dev Biol,* 11**,** 441-69.

XIAO, N. & DEFRANCO, D. B. 1997. Overexpression of unliganded steroid receptors activates endogenous heat shock factor. *Mol Endocrinol,* 11**,** 1365-74.

XU, J., WU, R. C. & O'MALLEY, B. W. 2009. Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat Rev Cancer,* 9**,** 615-30.

YAMAMOTO, A., HASHIMOTO, Y., KOHRI, K., OGATA, E., KATO, S., IKEDA, K. & NAKANISHI, M. 2000. Cyclin E as a coactivator of the androgen receptor. *J Cell Biol,* 150**,** 873-80.

YEH, S., KANG, H. Y., MIYAMOTO, H., NISHIMURA, K., CHANG, H. C., TING, H. J., RAHMAN, M., LIN, H. K., FUJIMOTO, N., HU, Y. C., MIZOKAMI, A., HUANG, K. E. & CHANG, C. 1999. Differential induction of androgen receptor transactivation by different androgen receptor coactivators in human prostate cancer DU145 cells. *Endocrine,* 11**,** 195-202.

YEH, S., SAMPSON, E. R., LEE, D. K., KIM, E., HSU, C. L., CHEN, Y. L., CHANG, H. C., ALTUWAIJRI, S., HUANG, K. E. & CHANG, C. 2000. Functional analysis of androgen

receptor N-terminal and ligand binding domain interacting coregulators in prostate cancer. *J Formos Med Assoc,* 99**,** 885-94.

YEH, S. H., CHIU, C. M., CHEN, C. L., LU, S. F., HSU, H. C., CHEN, D. S. & CHEN, P. J. 2007. Somatic mutations at the trinucleotide repeats of androgen receptor gene in male hepatocellular carcinoma. *Int J Cancer,* 120**,** 1610-7.

YONG, E. L., CHUA, K. L., YANG, M., ROY, A. & RATNAM, S. 1994. Complete androgen insensitivity due to a splice-site mutation in the androgen receptor gene and genetic screening with single-stranded conformation polymorphism. *Fertil Steril,* 61**,** 856-62.

YONG, E. L., LOY, C. J. & SIM, K. S. 2003. Androgen receptor gene and male infertility. *Hum Reprod Update,* 9**,** 1-7.

YOUNG, I. E., KURIAN, K. M., MACKENZIE, M. A., KUNKLER, I. H., COHEN, B. B., HOOPER, M. L., WYLLIE, A. H. & STEEL, C. M. 2000. The CAG repeat within the androgen receptor gene in male breast cancer patients. *J Med Genet,* 37**,** 139-40.

YOUSSOUFIAN, H. & PYERITZ, R. E. 2002. Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet,* 3**,** 748-58.

YU, H., BHARAJ, B., VASSILIKOS, E. J., GIAI, M. & DIAMANDIS, E. P. 2000. Shorter CAG repeat length in the androgen receptor gene is associated with more aggressive forms of breast cancer. *Breast Cancer Res Treat,* 59**,** 153-61.

ZAVA, D. T. & MCGUIRE, W. L. 1977. Estrogen receptors in androgen-induced breast tumor regression. *Cancer Res,* 37**,** 1608-10. ZAVA, D. T. AND W. L.

MCGUIRE 1977. "Estrogen receptors in androgen-induced breast tumor regression." Cancer Res **37**(6): 1608-1610.

ZEGARRA-MORO, O. L., L. J. SCHMIDT, H. HUANG AND D. J. TINDALL 2002. "Disruption of androgen receptor function inhibits proliferation of androgen-refractory prostate cancer cells." Cancer Res **62**(4): 1008-1013.

ZHOU, Z. X., WONG, C. I., SAR, M. & WILSON, E. M. 1994. The androgen receptor: an overview. *Recent Prog Horm Res,* 49**,** 249-74.

ZILLIACUS, J., A. P. WRIGHT, U. NORINDER, J. A. GUSTAFSSON AND J. CARLSTEDT-DUKE 1992. "Determinants for DNA-binding site recognition by the glucocorticoid receptor." J Biol Chem **267**(35): 24941-24947.

ZINN, A. R., RAMOS, P., ELDER, F. F., KOWAL, K., SAMANGO-SPROUSE, C. & ROSS, J. L. 2005. Androgen receptor CAGn repeat length influences phenotype of 47,XXY (Klinefelter) syndrome. *J Clin Endocrinol Metab,* 90**,** 5041-6.

ZURRIDA, S., F. NOLE, B. BONANNI, M. G. MASTROPASQUA, P. ARNONE, O. GENTILINI AND A. LATRONICO 2010. "Male breast cancer." Future Oncol **6**(6): 985-991.