MOLECULAR MECHANISMS IN DISORDERS OF GONADAL DEVELOPMENT

Ameli Norling

Stockholm 2014
"As we enter the era of Big Data, with information and processing power increasing at exponential rates, it may be time to develop a healthier attitude towards computers and what they might accomplish for us. Technology is beneficial as a labor-saving device, but we should not expect our machines to do our thinking for us."

“The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning.”

From “The Signal and the Noise: why so many predictions fail - but some don’t”

Nate Silver
ABSTRACT

"Hur många barn får jag när jag blir stor?"
Controlling fertility, or rather infertility, is a common problem for couples nowadays where approximately 10% experience infertility. This thesis focuses on the rare conditions leading to infertility in which the gonads have either failed to develop properly, or prematurely lost their function. The overall aim of this thesis was to identify and study candidate genes in gonadal disorders of sex development (DSD) and primary ovarian insufficiency (POI) to achieve a better understanding of the underlying molecular mechanisms and to improve available diagnostic tools and genetic counseling for affected patients and family members.

The process of sex determination is sensitive to gene dosage, and genes affecting gonadal development are located on sex chromosomes and autosomes. Attempting to identify new candidate regions for testicular development, nine patients with 46,XY complete gonadal dysgenesis were investigated with a customized 1M array-CGH (comparative genomic hybridization) platform with whole-genome coverage and probe enrichment targeting 78 genes involved in DSD. Aberrations affecting SUPT3H and C2ORF80 were identified and characterized in two affected sisters. Sequence analysis of these genes in all patients revealed no additional mutations. PIP5K1B and PRKACG were also identified as candidate genes in another patient. All four genes are expressed in testicular tissues, and one is shown to cause gonadal DSD in mice (Paper I).

Based on the report of one patient with 46,XY ovarian DSD due to CBX2 mutations and the corresponding knock out mouse model, we considered CBX2 as a candidate gene for 46,XY/46,XX gonadal DSD and POI. By sequencing and MLPA a cohort of 47 patients with gonadal DSD or POI was investigated for sequence alterations and copy number changes in CBX2. No causative mutation was detected. Our study does therefore not support CBX2 mutations as a common cause of gonadal DSD (Paper II).

DNA samples from 26 patients with POI were analyzed by our customized array-CGH platform to identify novel candidate genes in POI. Eleven unique copy number changes were identified in a total of 13 patients. We identified the first mutation affecting the regulatory region of GDF9, a 475bp duplication containing three NOBOX binding elements and a regulatory E-box. In addition, we found the second DNAH6 deletion, corroborating its potential role in causing in POI. Also TSPYL6, SMARCC1, CSPG5, SH3GL3 and ZFR2 were identified as candidate genes in POI (Paper III).

PSMC3IP mutations have been described causing autosomal recessive POI in a consanguineous family with several affected members. We are the first to have investigated the PSMC3IP gene in a group of unrelated patients with POI. No mutation was detected and we therefore conclude that PSMC3IP mutations are not a common cause of POI in a Swedish patient cohort (Paper IV).

In conclusion, there is accumulating evidence that gonadal DSD and POI are highly genetically heterogeneous. We recommend that the genetic investigation of patients with gonadal DSD should be centralized to a specialized DSD unit. Patients with POI should be handled by subspecialists within reproductive endocrinology and clinical genetic investigation can require genetic counseling. A suggested clinical investigation procedure for POI is presented. Ideally, genetic investigation of patients with gonadal DSD or POI should include investigation for gene dosage aberrations as well sequencing of several candidate genes.
LIST OF PUBLICATIONS

This thesis is based on the following articles, referred to by roman numerals in the text.


III. Norling A, Hirschberg AL, Rodriguez-Wallberg K, Iwarsson E, Wedell A, Barbaro M. Identification of a duplication within the GDF9 gene and novel candidate genes for primary ovarian insufficiency (POI) by a customized high resolution array-CGH platform. *Submitted*

IV. Norling A, Hirschberg AL, Karlsson L, Rodriguez-Wallberg K, Iwarsson E, Wedell A, Barbaro M. No mutations in the PSMC3IP gene identified in a Swedish cohort of women with Primary Ovarian Insufficiency (POI). *Accepted for publication in Sexual Development*
UNRELATED PUBLICATIONS BY THE AUTHOR

I. Johansson M, Norda A, Karlsson A.
Conserved gene structure and transcription factor sites in the human and mouse deoxycytidine kinase genes.

II. Van Rompay AR, Norda A, Lindén K, Johansson M, Karlsson A.
Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases.

III. Schneider B, Norda A, Karlsson A, Veron M, Deville-Bonne D.
Nucleotide affinity for a stable phosphorylated intermediate of nucleoside diphosphate kinase.
*Protein Science.* 2002 Jul;11(7):1648-56.
## CONTENTS

**INTRODUCTION** .................................................................................................................. 1

- Sex development .................................................................................................................. 1
  - Sex determination .............................................................................................................. 1
  - Sex differentiation ............................................................................................................ 3
- Internal genitalia .................................................................................................................... 4
- External genitalia .................................................................................................................. 4
- Fertility potential .................................................................................................................. 4
- Gender .................................................................................................................................. 4

**Disorders of gonadal development** .................................................................................. 5

- Complete 46,XY Gonadal Dysgenesis ................................................................................ 6
- Other 46,XY gonadal DSD .................................................................................................. 6
- 46,XX gonadal DSD ............................................................................................................ 7

**Primary Ovarian Insufficiency** ......................................................................................... 8

**Genetics** ............................................................................................................................. 9

- Genes in 46,XY gonadal DSD ............................................................................................. 10
- Genes in 46,XX gonadal DSD ............................................................................................. 12
  - *NR5A1* .......................................................................................................................... 14
  - *CBX2* ............................................................................................................................ 14
- Candidate genes in gonadal DSD ........................................................................................ 15
- POI ...................................................................................................................................... 18
- Candidate genes in POI ....................................................................................................... 21
- Tilting the balance ............................................................................................................... 28

**AIMS** .................................................................................................................................. 29

**MATERIAL AND METHODS** .......................................................................................... 30

- Patients ................................................................................................................................ 30
  - Clinical characterization .................................................................................................... 31
  - Clinical genetic analysis ................................................................................................... 31
- Controls ................................................................................................................................. 32
- DNA Sequencing .................................................................................................................. 32
- Reverse Transcriptase-PCR ................................................................................................. 33
- Array comparative genomic hybridization ........................................................................ 33
  - Procedure ........................................................................................................................ 34
  - Interpretation ...................................................................................................................... 35
  - Copy number variants ....................................................................................................... 35
  - Limitations ......................................................................................................................... 35
- Multiplex ligation-dependent probe amplification ............................................................. 36
  - MLPA probes .................................................................................................................... 36
  - Procedure ........................................................................................................................ 37
  - Controls ............................................................................................................................ 38
  - Interpretation ...................................................................................................................... 38

*In silico* analysis, online resources and tools ....................................................................... 38

Statistical analysis .................................................................................................................. 39
RESULTS AND DISCUSSION ................................................................. 40
   Clinical analyses (unpublished data) ........................................... 40
   XY gonadal DSD ......................................................................... 40
   POI ............................................................................................. 41
   Genome wide approach to identify novel candidate regions in 46,XY
gonadal dysgenesis (Paper I) .......................................................... 43
      SUPT3H on 6p21.1 ..................................................................... 43
      C2ORF80 on 2q34 ..................................................................... 44
      PIP5K1B and PRKACG on 9q21.11 ............................................. 45
      Conclusion .................................................................................. 45
   CBX2 candidate gene study (Paper II) ............................................. 46
      Sequence alterations ................................................................. 46
      Copy number changes ............................................................... 47
      Expression ................................................................................ 48
      Conclusion ................................................................................ 48
   Genome wide approach to identify novel candidate regions in primary
ovarian insufficiency (Paper III) ......................................................... 48
      GDF9 on 5q31.1 ......................................................................... 49
      DNAH6 on 2p11.2 ..................................................................... 50
      TSPYL6 on 2p16.2 ..................................................................... 51
      SMARCC1 and CSPG5 on 3p21.31 ............................................ 51
      SH3GL3 on 15q25.2 ................................................................. 51
      Del 19p13.3 ............................................................................... 52
      KRTAP2-3 and KRTAP2-4 on 17q21.2 ....................................... 52
      Conclusion ................................................................................ 52
   PSMC3IP gene study (paper IV) ......................................................... 53
      Sequencing results ..................................................................... 53
      Conclusion ................................................................................ 53
   POI gene sequencing (unpublished data) .......................................... 54
      FIGLA ...................................................................................... 54
      GDF9 ...................................................................................... 55
      NOBOX ................................................................................... 55

CONCLUDING REMARKS - FUTURE PERSPECTIVE ......................... 60
POPULÄRVETENSKAPLIG SAMMANFATTNING PÅ SVENSKA ............ 63
APPENDIX 1. Suggested clinical investigation of women with POI ....... 65
   Clinical presentation: .................................................................... 65
   Basic investigation ......................................................................... 65
   Specialist investigation ................................................................... 66
   Gonadal biopsy investigation ......................................................... 66
APPENDIX II. In silico analysis, programs, resources and tools ............ 67
ACKNOWLEDGEMENTS .................................................................. 69
REFERENCES .................................................................................. 73
LITERATURE RECOMMENDATIONS ............................................... 83
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>autosomal dominant</td>
</tr>
<tr>
<td>AMH</td>
<td>Anti Müllerian hormone</td>
</tr>
<tr>
<td>AR</td>
<td>autosomal recessive</td>
</tr>
<tr>
<td>array-CGH</td>
<td>array comparative genomic hybridization</td>
</tr>
<tr>
<td>ARX</td>
<td>Aristaless related homeobox, X-linked gene</td>
</tr>
<tr>
<td>ATRX</td>
<td>Alpha thalassemia/mental retardation syndrome X linked gene</td>
</tr>
<tr>
<td>BMP15</td>
<td>Bone morphogenetic protein 15 gene</td>
</tr>
<tr>
<td>bp</td>
<td>basepair</td>
</tr>
<tr>
<td>BPES</td>
<td>Blepharophimosis, ptosis and epicanthus inversus syndrome</td>
</tr>
<tr>
<td>C2ORF80</td>
<td>Chromosome 2 open reading frame 80 gene</td>
</tr>
<tr>
<td>CBX2</td>
<td>Chromobox homolog 2 gene</td>
</tr>
<tr>
<td>CD</td>
<td>Campomelic Dysplasia</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CNV</td>
<td>copy number variation</td>
</tr>
<tr>
<td>CSPG5</td>
<td>Chondroitin sulfate proteoglycan 5 (neuroglycan C) gene</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>β-catenin gene</td>
</tr>
<tr>
<td>CYP11A1</td>
<td>Cytochrome P450 side-chain cleavage enzyme</td>
</tr>
<tr>
<td>dbSNP</td>
<td>database for single nucleotide polymorphisms</td>
</tr>
<tr>
<td>dNTP</td>
<td>dideoxynucleotides</td>
</tr>
<tr>
<td>DECIPHER</td>
<td>Database of chromosomal imbalances and phenotype in humans using Ensembl resources</td>
</tr>
<tr>
<td>DGV</td>
<td>Database of genomic variants</td>
</tr>
<tr>
<td>DHH</td>
<td>Desert hedgehog gene</td>
</tr>
<tr>
<td>DHT</td>
<td>Dihydrotestosterone</td>
</tr>
<tr>
<td>DMRT1</td>
<td>Doublesex- and MAB3-related transcription factor 1 gene</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNAH5</td>
<td>Dynein, axonemal, heavy chain 5 gene</td>
</tr>
<tr>
<td>DNAH6</td>
<td>Dynein, axonemal, heavy chain 6 gene</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleotides</td>
</tr>
<tr>
<td>DSD</td>
<td>Disorders of Sex Development</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein Barr virus</td>
</tr>
<tr>
<td>ESR1</td>
<td>Estrogen receptor 1 gene</td>
</tr>
<tr>
<td>FIGLA</td>
<td>Folliculogenesis specific basic helix-loop-helix gene</td>
</tr>
<tr>
<td>FMR1</td>
<td>Fragile X mental retardation 1 gene</td>
</tr>
<tr>
<td>FOG2</td>
<td>Friend of GATA2 gene</td>
</tr>
<tr>
<td>FOXL2</td>
<td>Forkhead box L2 gene</td>
</tr>
<tr>
<td>FSH</td>
<td>Follicle stimulating hormone</td>
</tr>
<tr>
<td>FSHR</td>
<td>Follicle stimulating hormone receptor gene</td>
</tr>
<tr>
<td>FXTAS</td>
<td>Fragile X-associated tremor/ataxia syndrome</td>
</tr>
<tr>
<td>GATA4</td>
<td>GATA-binding protein 4 gene</td>
</tr>
<tr>
<td>GD</td>
<td>Gonadal Dysgenesis</td>
</tr>
<tr>
<td>GDF9</td>
<td>Growth differentiation factor 9 gene</td>
</tr>
<tr>
<td>GF</td>
<td>growth factor</td>
</tr>
<tr>
<td>GnRH</td>
<td>Gonadotropin releasing hormone</td>
</tr>
</tbody>
</table>
GOF  gain of function
GWAS  genome wide association study
INHA  Inhibin-alpha gene
KANSL1L  KAT8 regulatory NSL complex subunit 1-like gene
kb  kilobase (1000 base pairs)
KO  knock out
KRTAP2-3  Keratin associated protein 2-3 gene
KRTAP2-4  Keratin associated protein 2-4 gene
LH  Leutenizing hormone
LHCGR3  Leutenizing hormone receptor gene
LOF  loss of function
MAF  minor allele frequency
MAP3K1  mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase gene
MAPK  mitogen-activated protein kinase gene
MIM  OMIM number
MLPA  multiplex ligation-dependent probe amplification
NAP  nucleosome assembly protein
NBE  NOBOX binding element
NCBI  National Center for Biotechnology Information
NOBOX  Nobox oogenesis homeobox gene
NR  nuclear receptor
NR0B1  Nuclear receptor subfamily 0, group B, member 1 gene
NR5A1  Nuclear receptor subfamily 5, group A, member 1 gene
OMIM  Online Mendelian Inheritance in Man, database
OHSS  ovarian hyperstimulation syndrome
P450scc  P450 side-chain cleavage enzyme
PA  primary amenorrhea
pc  post conception
PCD  Primary ciliary dyskinesia
PCR  polymerase chain reaction
PEOA1  Progressive external ophthalmoplegia with mitochondrial DNA deletions
PIP5K1B  Phosphatidylinositol-4-phosphate 5-kinase, type 1, beta gene
POI  Primary Ovarian Insufficiency
PRKACG  Homo sapiens protein kinase, cAMP-dependent, catalytic, gamma gene
PSMC3IP  PSMC3 interacting protein gene
RNA  ribonucleid acid
RSP01  R-spondin family, member 1 gene
SA  secondary amenorrhea
SD  standard deviation
SERKAL  Female to male Sex Reversal and Kidney, Adrenal and Lung dysgenesis
SF  signaling factor
SF-1  Steroidogenic factor 1 gene
SH3GL3  SH3-domain GRB2-like 3 gene
SIDDT  Sudden infant death with dysgenesis of the testes syndrome
**SMARCC1**  SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily c member1 gene

**SNP**  single nucleotide polymorphism

**SOX9**  SRY-box 9 gene

**SRD5A2**  Steroid-5-alpha-reductase gene

**SRY**  Sex determining region Y gene

**SUPT3H**  Suppressor of Ty 3 homologue S. cerevisiae gene

**TF**  transcription factor

**TGFB1**  Transforming growth factor beta super family

**TPO**  thyroid peroxidase

**TSPYL1**  TSPY-like 1 gene

**TSPYL6**  TSPY-like 6 gene

**UCSC**  University of California Santa Cruz

**UTR**  untranslated region

**WNT4**  Wingless-type MMTV Integration site family, member 4 gene

**WT1**  Wilms tumor 1 gene

**ZFPM2**  zinc finger protein, FOG family member 2 gene

**ZFR2**  zinc finger RNA binding protein 2 gene

Human gene names are given in capital italics (*NR5A1*).

Human proteins are given in capital letters (*NR5A1*).

Mouse gene names are given in italics, first letter capitalized (*Nr5a1*).

For better legibility full gene names are only given first time the gene is discussed in depth.

For better legibility full gene names for candidate genes discussed in the introduction are not given. However, OMIM number is given for each gene in the corresponding table to facilitate further reading.
INTRODUCTION

“How many children will I have when I grow up?”

A popular children’s game at daycare when I was a child, was to grab some gravel, and throw it in the air while saying “How many children will I have when I grow up?”. The number of small rocks you would catch was how many children you would have as an adult. Sometimes “fate” would give you ten, and sometimes none. You could always try to catch the desired number, but most often, you got what you got. As silly as this game would seem growing up, controlling your fertility, or rather the unwanted infertility, is a common problem for couples today. Approximately 10% of all couples trying to conceive will experience infertility [1]. There are many different factors that can cause infertility, both in men and women. This thesis will only focus on the rare conditions leading to infertility in which the gonads have either failed to develop properly, or prematurely lost their function. To study these patients, and understand the investigations undertaken, an understanding of the normal process of sex development is necessary.

SEX DEVELOPMENT

Human sex development is a continuous process, starting at fertilization, and continuing through puberty when the adult sex appearance is acquired along with fertility potential. This process is governed by genetic and hormonal factors.

At fertilization, the genetic sex is determined by the sex chromosome complement. Normally, men have one X and one Y chromosome (46,XY), and women two X chromosomes (46,XX). The next step is called sex determination, with the differentiation of the bipotential gonads into either ovaries or testes, followed by sex differentiation, the differentiation of the internal and external genitalia into either a male or female phenotype.

Sex determination

During embryonic development, the gonads are initially formed as bipotential and can develop into either ovaries or testes. The gonadal anlagen appear at 32 days post conception (pc), arising from the mesoderm. The anlagen consist of three different bipotential somatic cell lineages, the supporting, steroidogenic and stromal cells. By the 5th week pc, the primordial germ cells, migrating from the yolk sac, start to colonize the gonads [2]. At week 6-8, sex determination occurs, when the somatic cells and
Introduction

Table 1. Gonadal cell lineages.

<table>
<thead>
<tr>
<th>Testis</th>
<th>Bipotential gonad</th>
<th>Ovary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spermatogonia</td>
<td>Germ cells</td>
<td>Oogonia</td>
</tr>
<tr>
<td>Sertoli cells</td>
<td>Supporting cells</td>
<td>Granulosa cells*</td>
</tr>
<tr>
<td>Leydig cells</td>
<td>Steroidogenic cells</td>
<td>Theca cells</td>
</tr>
<tr>
<td>Peritubular cells</td>
<td>Stromal cells</td>
<td>Stromal cells</td>
</tr>
</tbody>
</table>

*Granulosa cells also have a steroidogenic function with progesterone production and aromatase conversion of androgens to estradiol.

primordial germ cells of the gonad differentiate into ovarian or testicular cell types (table 1) depending on the predominant activation of the SRY/SOX9 or RSPO1/WNT4 signaling pathway, respectively (figure 1) [3].

Testes

The differentiation towards testes starts around week 6 pc. The SRY gene, located on the male Y chromosome, initiates the differentiation towards testicular development by promoting SOX9 expression in pre-Sertoli cells [4, 5]. A positive feed-back loop stabilizes SOX9 expression, while WNT4 expression is suppressed [6]. The germ cells increase rapidly by mitosis and do not enter meiosis, but lie resting until puberty. The germ cells are nursed by Sertoli cells that express differentiation factors, paracrine signals and Anti Müllerian Hormone (AMH) necessary for male sex differentiation. The expression of DHH and FGF9 by Sertoli cells is essential for differentiation of Leydig cells, which in turn express NR5A1 necessary for Sertoli cell AMH production [7]. The Leydig cells are the steroidogenic cells of the testes producing androgens from the 8th week pc. The development of the somatic cell lineages to form the morphological testicle is not dependent on the presence of germ cells, but occurs independently of these [2].

Ovaries

Around 7 weeks pc, the differentiation towards ovaries starts in 46,XX embryos. It is not, as previously believed, a passive process in the absence of SRY, but an active, genetically determined process where RSPO1 and WNT4 upregulation is important [8]. However, no “ovary determining factor” comparable to SRY has so far been identified.
**Figure 1.** Sex determination.

The ovarian development of follicle organization is dependent on correct interaction of germ cells and granulosa cells [2, 9]. The interaction process is not clearly understood, but close cell organization and paracrine signaling are important. Germ cell specific expression of *FIGLA* and *NOBOX* [9], along with selective *FOXL2* expression in developing granulosa cells, has been shown [2]. *NR5A1* is expressed by both granulosa and theca cells, and is a major regulator for the steroidogenesis except for corpus luteum progesterone production [10].

The germ cells proliferate by mitotic division, and before birth, enter meiosis, where they arrest in the diplotene stage. Around week 12 pc the germ cells are surrounded by somatic granulosa-cell precursors, and these develop to form primordial follicles. During gestation, they develop into primary and later secondary follicles. However, many undergo programmed cell death, and at birth approximately 2 million follicles remain [2].

**Sex differentiation**

The differentiation of the internal and external genitalia into male or female phenotype is termed sex differentiation.
**Introduction**

**Internal genitalia**

Both the Wolffian ducts, that give rise to the male internal genitalia, and the Müllerian ducts, that establish the female internal genitalia, are formed during early gestation. AMH production by Sertoli cells in the testis causes regression of the Müllerian ducts. Androgens, produced by the Leydig cells around gestational week 8, are important for the development of the Wolffian ducts into the epididymis, ductus deferens and seminal vesicles [2, 11]. In the absence of AMH, the Müllerian ducts persist and form the uterus, fallopian tubes and the upper part of the vagina. In the absence of androgens, the Wolffian ducts regress and form the suspensory ligament of the ovary and the Garner’s ducts. AMH is also produced in low levels by primordial follicles, but at a later stage and this does not affect development of the Müllerian ducts [2].

**External genitalia**

The bipotential external genitalia start differentiating at gestational week 8 into a male phenotype in response to androgens and dihydrotestosterone (DHT) in particular. In the absence of androgen action, due to failure to respond to or synthesize androgens (testosterone and/or DHT), the external genitalia develop in a female pattern. [12]

**Fertility potential**

At puberty the normal sex development culminates with the activation of the gonads by the pituitary-hypothalamic-gonadal hormone axis. The surge of FSH (follicle stimulating hormone) and LH (luteinizing hormone) from the pituitary glands in response to gonadotropin releasing hormone (GnRH) causes the supporting cells of the gonads to produce male and female sex steroids, and activate the germ cells. In men the spermatogenesis is initiated, and in women the menstrual cycle starts with menarche. Womens’ fertility peaks between the age of 25 and 30 after which it steadily declines, along with decreasing follicle quality and quantity. The loss of fertility often precedes the loss of menstruation. The overall international median age of natural menopause is estimated to 50 years of age with a normal distribution range between 40 and 60 years and a standard deviation (SD) estimate of +/- 4 years [13, 14].

**Gender**

Gender can be considered as a part of the developmental process. It comprises gender identity, how one perceives one self, and gender role, how one is perceived by others/society. Sexual behavior and orientation are other components of the overall sexual identity of an individual. These issues will not be discussed in this thesis.
DISORDERS OF GONADAL DEVELOPMENT

Impaired sex development can cause Disorders of Sex Development (DSD), defined as “congenital conditions where development of chromosomal, gonadal or anatomical sex is atypical” [15]. DSD can be categorized into chromosomal DSD, caused by chromosomal aberrations including conditions with mixed and chimeric DSD, gonadal DSD, comprising disorders of gonad development and sex determination, and DSD caused by disorders of sex steroid synthesis and action. A revised nomenclature has been implemented since 2006 (table 2) [15]. This thesis focuses on gonadal DSD.

Table 2. DSD classification and terminology.

<table>
<thead>
<tr>
<th>Chromosomal DSD</th>
<th>46,XY DSD</th>
<th>46,XX DSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 45,X</td>
<td>A. Gonadal DSD</td>
<td>A. Gonadal DSD</td>
</tr>
<tr>
<td>Turner syndrome and variants</td>
<td>1. Complete gonadal dysgenesis (GD)</td>
<td>1. Gonadal dysgenesis (GD)*</td>
</tr>
<tr>
<td>B. 47,XXY</td>
<td>2. Partial GD</td>
<td>2. Ovotesticular DSD</td>
</tr>
<tr>
<td>Klinefelter syndrome and variants</td>
<td>3. Gonadal regression</td>
<td>3. Testicular DSD</td>
</tr>
<tr>
<td>C. 45,X/46,XY</td>
<td>4. Ovotesticular DSD</td>
<td></td>
</tr>
<tr>
<td>Mixed GD, ovotesticular DSD</td>
<td>5. Ovarian DSD*</td>
<td></td>
</tr>
<tr>
<td>D. 46,XX/46,XY</td>
<td>B. Disorders of androgen synthesis and action</td>
<td>B. Disorders caused by androgen excess</td>
</tr>
<tr>
<td>Chimeric, ovotesticular DSD</td>
<td>E.g. AIS, CAH, 5α-reductase deficiency, CYP11A1 deficiency</td>
<td>E.g. CAH, POR mutation</td>
</tr>
<tr>
<td></td>
<td>C. Other DSD</td>
<td>C. Other DSD</td>
</tr>
<tr>
<td></td>
<td>E.g. hypospadias, persistent Müllerian duct syndrome</td>
<td>E.g. MURCS, MRKH</td>
</tr>
</tbody>
</table>

* In this thesis 46,XX GD is discussed as primary amenorrhea in primary ovarian insufficiency (POI)

a Only one patient reported so far.

DSD, Disorders of sex development; AIS, androgen insensitivity syndrome; CAH, congenital adrenal hyperplasia; MURCS, Müllerian duct aplasia, unilateral renal agenesis and cervicothoracic somite anomalies; MRKH, Mayer-Rokitansky-Küster-Hauser syndrome.
Complete 46,XY Gonadal Dysgenesis

One form of gonadal DSD is gonadal dysgenesis (GD) where the gonads have failed to differentiate. In complete GD, only streak gonads are present. Despite a 46,XY karyotype, as there are no Sertoli or Leydig cells secreting AMH and androgens in the developing embryo, the internal and external genitalia develop along the female pathway. These girls are often diagnosed in their late teens when they have not achieved menarche and a 46,XY karyotype is discovered. Patients can also be diagnosed at birth if a girl is born where a prenatal test has shown a 46,XY karyotype.

As these patients lack sex steroid producing gonads, hormonal substitution is necessary even though low levels of sex steroids are secreted from the adrenal glands. Sometimes, these levels are sufficient for development of secondary sexual characteristics such as breast development and genital hair. If not, puberty can be induced by transdermal low-dose estrogen substitution. Often a combined oral contraceptive is used as replacement therapy post-puberty as the cyclic pattern of the oral contraceptives also induces menstruation, although ovulation will not occur. The patients can carry a pregnancy as they have a uterus, but the lack of germ cells makes egg donation the only available option for pregnancy.

The dysgenetic gonads, residing in the abdomen in these 46,XY subjects are at high risk of malignant transformation into germ cell tumors. This is believed to be due to the presence of Y chromosome material, where specifically expression of the testis specific protein encoded by TSPY on Yp11.2 is an important marker for malignancy potential. Due to the risk of gonadoblastoma, prophylactic gonadectomy is strongly recommended [16-19].

Other 46,XY gonadal DSD

There are patients presenting with partial forms of 46,XY gonadal dysgenesis, classified as partial GD. Here the phenotype is dependent on the hormonal function of the differentiated testicular tissue, with varying degree of ambiguous external genital depending on androgen production levels. The fate of the internal ducts is dependent on extent of AMH secretion and androgen levels.

In gonadal regression, one or both testicles are absent upon external and internal examination despite a most often normal male phenotype. The unilateral absence of a testis is estimated to be as prevalent as 1 per 1250 men [20].
A rare form of 46,XY gonadal DSD is ovotesticular DSD where both ovarian and testicular tissue is present in one or both gonads. As in partial GD, the clinical presentation will vary between patients depending on androgen and AMH production. The treatment is based on individual consideration.

One patient has been described with 46,XY ovarian DSD. This was a girl presenting with a normal female phenotype at birth despite a prenatal 46,XY karyotype. Laparoscopy and histological examination at 4 years of age identified bilateral ovaries with normal ovarian tissue and the presence of primordial follicles. The ovarian function could not be determined as the girl was still prepubertal [21].

There are syndromic forms of 46,XY gonadal DSD, such as campomelic dysplasia (MIM 114290) and Denys-Drash syndrome (MIM 194080). It can also be noted that CYP11A1, or cytochrome P450 side-chain cleavage (P450scc) enzyme mutations, categorized as a disorder of androgen synthesis and action, can present with a variable phenotype depending on mutation, from complete 46,XY GD to mild hypospadias. In these patients the manifestation is always accompanied by adrenal failure, with a later onset in milder phenotypes [22].

A specific molecular diagnosis can be established in 20-50% of patients with 46,XY gonadal DSD, depending on group classifications [15].

### 46,XX gonadal DSD

Patients with a female chromosome complement can present with ovotesticular DSD and are managed clinically with individual consideration, as for patients with 46,XY ovotesticular DSD. The clinical presentation is variable with different degrees of ambiguous external genitalia, and the AMH and androgen levels influence the fate of the internal genitalia.

Patients with 46,XX testicular DSD present with a male phenotype, testes and a male hormonal profile. The patients are azoospermic, and often diagnosed later in life during an infertility investigation. The phenotype of external genitalia can vary, with presentation range from ambiguous external genitalia to mild hypospadias.

In 46,XX gonadal dysgenesis only streak gonads are present and patients present as normal girls, who do not reach menarche. However, as the ovarian follicle organization is dependent on functioning interaction between germ cells and granulosa cells, conditions with completely undifferentiated streak gonads cannot be distinguished from conditions with premature loss of germ cells. As the term gonadal dysgenesis makes a
mechanistic supposition, is not a suitable term to describe all 46,XX patients with the clinical presentation of primary amenorrhea. In addition, ovarian dysfunction can be seen as a continuous spectrum from a complete form of undifferentiated gonads to a partial defect with premature loss of ovarian function. Therefore, these patients will be discussed under the term primary ovarian insufficiency. The patients are further subdivided depending on the clinical presentation of primary or secondary amenorrhea.

As POI is a very heterogeneous condition with variable presentation, it is not suitable to define it as a form of gonadal DSD. Instead, POI will be discussed separately, and the term POI with primary amenorrhea will replace 46,XX GD in this thesis.

**PRIMARY OVARIAN INSUFFICIENCY**

Ovarian dysfunction that occurs before the age of 40 is called primary ovarian insufficiency (POI), defined as prolonged amenorrhea accompanied by repeated serum FSH levels in the menopausal range [23, 24]. Elevated FSH levels with concurrent low estrogen levels is a status termed hypergonadotropic hypogonadism. The estimated general population incidence of POI is 1% [25].

POI can present with either primary (PA) or secondary amenorrhea (SA), where the latter has previously been termed premature ovarian failure (POF). However, as some remaining ovarian function can exist despite SA, the terms POF or premature menopause are considered misnomers. Approximately 5-10% of these women have delivered a child after receiving the diagnosis [23]. For patients with PA, fecundity has not been reported, but egg donation is an available option for achieving pregnancy.

Hormonal substitution is necessary to avoid negative effects of lack of estrogens [26]. Often a combined oral contraceptive is used to mimic the normal cyclic hormonal pattern, but hormone replacement therapy regimes can also be followed. In the case of patients with PA who have not reached puberty, this can first be induced using low dose transdermal estrogen substitution. Substitution therapy is recommended up until the age of normal menopause, approximately 50 years of age. Clinical supervision of these patients is important to avoid negative effects of under-substitution such as osteoporosis, where repeated dual-energy X-ray absorptiometry exams are recommended.

POI pathogenesis can be divided into follicle dysfunction and follicle depletion. Follicle depletion can be caused either by a small initial germ cell count, or a rapid
Introduction

germ cell loss. External factors, such as surgery, irradiation, chemotherapy as well as autoimmune and metabolic disorders, can cause POI.

Follicle dysfunction can be caused by enzyme defects, such as mutations in CYP17A1 and CYP19A1. Also, autoimmunity with specific antibodies against the enzymes CYP11A1 (p450scc) CYP17A1 (17α-hydroxyase) and CYP21A2 (21-hydroxylase) is associated with POI, most often in conjunction with autoimmune Addison’s disease [27, 28]. Thyroid disease is linked to ovarian function in women, and the presence of autoantibodies against TPO (thyroid peroxidase) is more frequent in women with POI [29].

POI can be isolated, or associated to other symptoms, such as in Turner syndrome. Other examples include BPES (Blepharophimosis, ptosis and epicanthus inversus syndrome, MIM 110100) where women present with eyelid deformities and PA, and Perrault syndrome (MIM 233400) where female patients present with PA and progressive hearing loss. POI can also be a symptom in mitochondrial diseases, such as POLG mutations and PEOA1 (progressive external ophtalmoplegia with mitochondrial DNA deletions, MIM 157640).

The absolute majority (90%) of all cases of isolated POI remain idiopathic [23]. However, there is a significant genetic component in POI and familial clusters are not uncommon [23, 30, 31].

GENETICS

The genes known to cause gonadal disorders of development and function have been identified by studying patients with DSD and different animal models. The sex determination process is similar for many mammals, and findings in one species can often be relevant for another. For instance, mouse Sry and Fshr mutation models are human phenocopies [7, 32]. In animal model studies, the mouse is the most frequently used. Other species are also used, for instance cattle and sheep for investigation of ovarian disorders. For these, interspecies differences in ovulation patterns must be taken into account. For instance, Bmp15 mutations do not affect phenotype in the polyovulatory mouse, but cause superovulation in the monoovulatory sheep [7].

Genetic changes can cause gonadal DSD through different mechanisms. Point mutations of specific genes can cause loss of gene product function, loss of function mutations (LOF). If one allele is affected, a LOF mutation can cause gene haploinsufficiency, and if both alleles are affected, complete gene product loss. When
the phenotype shows autosomal dominant inheritance, haploinsufficiency is enough to cause the phenotype, whereas complete loss is necessary in autosomal recessive disease. A complete loss can occur when both alleles harbor the same mutation, where the patient is homozygous for the mutation. The patient can also carry two different mutations, called compound heterozygosity.

If one allele is deleted, the patient is hemizygous. In an autosomal dominant condition, gene deletion of one allele is sufficient to cause the phenotype. If a deletion of one allele occurs together with a LOF mutation in the remaining allele, this causes a complete loss of gene product that can cause autosomal recessive disease. Entire gene duplications can lead to increased gene expression. In dosage sensitive developmental steps, such as gonadal differentiation, that may have detrimental phenotypic effects. Duplications, that due to a positional effect cause functional gene loss, can affect the phenotype in the same manner as a LOF mutation.

Other chromosomal rearrangements such as translocations can cause haploinsufficiency by gene disruption at breakpoints. However, translocations can also render loss or gain of genetic material, causing a functional gain or loss of gene copies.

In gonadal DSD all the mentioned mechanisms of disease have been described. As a further complication, the phenotypic effect of a mutation in DSD is dependent on the chromosomal sex. A loss of a gene will only affect the development of the gonadal form it is necessary for, and a gain will only affect the development of the gonad type in which it is normally down regulated. By extension, mutations that cause gonadal DSD can be transmitted through healthy carriers of the opposite chromosomal sex. There are however also genes that are necessary for the development of both gonadal forms, where a phenotypic effect will be seen in both sexes. In addition, penetrance can vary between individuals.

A proposed schematic overview of known and candidate genes in sex determination and gonadal development is given in figure 2 at the end of the introduction section.

**Genes in 46,XY gonadal DSD**

The understanding of the genetic mechanisms of 46,XY gonadal DSD took a great step forward in 1990 with the identification of the **SRY** (Sex determining region Y) gene [33]. **SRY** mutations/deletions account for approximately 10-15% of all cases of 46,XY GD [34] (table 3a).

LOF mutations of **WT1** (Wilms tumor 1) cause different forms of renal and gonadal
### Table 3a. Genes in 46,XY gonadal DSD.

<table>
<thead>
<tr>
<th>Gene</th>
<th>OMIM no</th>
<th>Locus</th>
<th>Function</th>
<th>46,XY Inheritance</th>
<th>Gonadal phenotype</th>
<th>External genitalia</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARX</td>
<td>300382</td>
<td>Xp21.3</td>
<td>TF</td>
<td>X-linked</td>
<td>GD</td>
<td>Ambiguous</td>
<td>X-linked lissencephaly, epilepsy.</td>
</tr>
<tr>
<td>ATRX</td>
<td>300032</td>
<td>Xq21.1</td>
<td>Helicase</td>
<td>X-linked</td>
<td>GD</td>
<td>Female / ambiguous / male</td>
<td>Associated mental retardation and X-linked thalassemia.</td>
</tr>
<tr>
<td>CBX2</td>
<td>602770</td>
<td>17q25.3</td>
<td>TR</td>
<td>AR</td>
<td>Ovarian</td>
<td>Female</td>
<td>Trans-activates NR5A1.</td>
</tr>
<tr>
<td>DHH</td>
<td>605423</td>
<td>12q13.1</td>
<td>SF</td>
<td>AD/AR</td>
<td>GD</td>
<td>Ambiguous</td>
<td>Associated minifascicular neuropathy.</td>
</tr>
<tr>
<td>DMRT1</td>
<td>602424</td>
<td>9p24.3</td>
<td>TF</td>
<td>Del9p24.3 (reduced penetrance)</td>
<td>GD</td>
<td>Female/ambiguous</td>
<td>Associated mental retardation depending on extent of gene deletion.</td>
</tr>
<tr>
<td>GATA4</td>
<td>600576</td>
<td>8p.23.1-p22</td>
<td>TF</td>
<td>AD</td>
<td>GD / testis</td>
<td>Ambiguous</td>
<td>Interacts with ZFPM2, NR5A1 and the AMH promotor. ASD, VSD, atrial fibrillation.</td>
</tr>
<tr>
<td>MAP3K1</td>
<td>600982</td>
<td>5q11.2</td>
<td>SF</td>
<td>AD</td>
<td>GD partial or complete / testis</td>
<td>Female /ambiguous /hypospadia</td>
<td>-</td>
</tr>
<tr>
<td>NR0B1</td>
<td>300473</td>
<td>Xp21.3</td>
<td>NR TF?</td>
<td>DupXp21.2</td>
<td>GD</td>
<td>Female</td>
<td>LOF mutations causes adrenal defects in males with an X-linked inheritance.</td>
</tr>
<tr>
<td>NR5A1</td>
<td>184757</td>
<td>9q33</td>
<td>NR</td>
<td>AR/AD</td>
<td>GD partial or complete / anorchia</td>
<td>Female /ambiguous /hypospadia</td>
<td>Mutations also seen in isolated spermatogenic failure. Associated with adrenal failure.</td>
</tr>
<tr>
<td>SOX9</td>
<td>608160</td>
<td>17q24.3</td>
<td>TF</td>
<td>AD</td>
<td>GD / ovotestis</td>
<td>Female /ambiguous</td>
<td>Upregulated by SRY and NR5A1. (See table 3b)</td>
</tr>
<tr>
<td>SRY</td>
<td>480000</td>
<td>Yp11.3</td>
<td>TF</td>
<td>Y</td>
<td>GD / ovotestis</td>
<td>Female</td>
<td>10-15% of XY GD.</td>
</tr>
<tr>
<td>TSPY1L</td>
<td>604714</td>
<td>6q22.1</td>
<td>SF</td>
<td>AR/AD</td>
<td>GD / azoospermia</td>
<td>Female /male</td>
<td>SIDDIT</td>
</tr>
<tr>
<td>ZFPM2</td>
<td>603693</td>
<td>8q23.1</td>
<td>TF</td>
<td>46,XY,t(8;10)</td>
<td>Hypotrophic testes</td>
<td>Male</td>
<td>LOF mutations , heart defects interacts with GATA proteins.</td>
</tr>
<tr>
<td>WNT4</td>
<td>603490</td>
<td>1p36.23-p35</td>
<td>SF</td>
<td>Dup1p35</td>
<td>GD</td>
<td>Ambiguous</td>
<td>Associated mental retardation. (See table 3b)</td>
</tr>
<tr>
<td>WT-1</td>
<td>607102</td>
<td>11p13</td>
<td>TF</td>
<td>AD</td>
<td>GD</td>
<td>Female /ambiguous</td>
<td>Wilms tumor, renal defects, WAGR, Denys-Drash and Frasier syndrome.</td>
</tr>
</tbody>
</table>

TF, transcription factor; GD, gonadal dysgenesis; TR, transcriptional regulator; AR, autosomal recessive; SF, signaling factor; AD, autosomal dominant; Del, deletion; AMH, anti Müllerian hormone; NR, nuclear receptor; Dup, duplication; LOF, loss of function; WAGR, Wilms tumor, aniridia, genitourinary anomalies and mental retardation syndrome.
defects depending on mutation. Mutations can also cause the renal cancer form Wilms tumor. A wide range of gonadal defects in 46,XY is reported in heterozygous WT1 mutations, from complete GD to partial forms, and also ovotesticular DSD, caused by haploinsufficiency. DHH (Desert Hedgehog) is necessary for NR5A1 upregulation of Leydig cells. Several patients with complete or partial 46,XY GD due to mutations in this gene are reported, both with autosomal dominant and recessive inheritance. Homozygous mutations of TSPYL1 (TSPY-like 1) cause sudden infant death with dysgenesis of the testes syndrome (SIDDT, MIM 608800) and heterozygous mutations are described in both GD and isolated azoospermia [35].

Mutations in MAP3K1 (mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase) cause autosomal dominant complete or partial 46,XY GD. Mutations of the two interacting genes GATA4 (GATA-binding protein 4) and ZFPM2 (zinc finger protein, FOG family member 2), formerly known as FOG2 (friend of GATA2) cause autosomal dominant gonadal DSD in addition to several different types of heart anomalies.

Copy number changes affecting NR0B1 (Nuclear receptor subfamily 0, group B, member 1) and DMRT1 (Doublesex- and MAB3-related transcription factor 1) can cause GD, either in isolated form, or with associated features depending on the extent of the duplication/deletion. Also WNT4 duplication causes 46,XY gonadal DSD with ambiguous genitalia.

Mutations in several genes cause syndromic forms of 46,XY gonadal DSD. For instance, campomelic dysplasia (CD), a skeletal malformation, is caused by loss of SOX9 (SRY-box 9) due to LOF mutations, gene deletion or genetic rearrangements affecting regulatory regions. Patients with CD without GD are reported, and vice versa. Mutations in ARX (aristaless related homeobox, X-linked) and ATRX (alpha thalassemia/mental retardation syndrome X linked) cause X-linked 46,XY GD with associated symptoms.

**Genes in 46,XX gonadal DSD**

Ovarian development is dependent on RSPO1 (R-spondin family, member 1) and WNT4 (Wingless-type MMTV Integration site family, member 4) (table 3b). LOF mutations of RSPO1 cause autosomal recessive 46,XX testicular DSD with associated cutaneous symptoms. Also ovotesticular DSD with palmoplantar keratoderma has been found in one patient [36].
<table>
<thead>
<tr>
<th>Gene</th>
<th>OMIM no</th>
<th>Locus</th>
<th>Function</th>
<th>Inheritance</th>
<th>46,XX</th>
<th>Gonal phenotype</th>
<th>External genitali</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HSD17B4</strong> 601860</td>
<td></td>
<td>5q23.1</td>
<td>Enzyme</td>
<td>AR</td>
<td>46,XX</td>
<td>Ovarian dysgenesis</td>
<td>Female</td>
<td>Perrault syndrome with sensorineural deafness. 17-β-estradiol dehydrogenase IV</td>
</tr>
<tr>
<td><strong>MAMLID1</strong> 300120</td>
<td></td>
<td>Xq28</td>
<td>TF</td>
<td>AD</td>
<td>46,XX</td>
<td>Ovarian dysgenesis</td>
<td>Clitoromegaly</td>
<td>GOF mutation identified in one patient. Mutations cause hypospadia in 46,XY.</td>
</tr>
<tr>
<td><strong>RSPO1</strong> 609595</td>
<td></td>
<td>1p34.3</td>
<td>SF</td>
<td>AR</td>
<td>46,XX</td>
<td>Ovotestis</td>
<td>Ambiguous</td>
<td>Palmoplantar hyperkeratosis. Associated with congenital corneal opacities, onychodystrophy and hearing impairment.</td>
</tr>
<tr>
<td><strong>SOX3</strong> 313430</td>
<td></td>
<td>Xq27.1</td>
<td>TF</td>
<td>Del/DupXq27.1</td>
<td>46,XX</td>
<td>Testis</td>
<td>Male</td>
<td>Altered gene regulation by disruption of promotor region.</td>
</tr>
<tr>
<td><strong>SOX9</strong> 608160</td>
<td></td>
<td>17q24.3</td>
<td>TF</td>
<td>Dup17q24</td>
<td>46,XX</td>
<td>Testis, azoospermia</td>
<td>Male</td>
<td>(See table 3a)</td>
</tr>
<tr>
<td><strong>SOX10</strong> 602229</td>
<td></td>
<td>22q13.1</td>
<td>TF</td>
<td>Dup22q13</td>
<td>46,XX</td>
<td>Testis / ovotestis</td>
<td>Male, hypospadia / ambiguous</td>
<td>LOF mutation cases Waardenburg-Shah syndrome.</td>
</tr>
<tr>
<td><strong>SRY</strong> 430000</td>
<td></td>
<td>Yp11.3</td>
<td>TF</td>
<td>Translocation</td>
<td>46,XX</td>
<td>Testis / ovotestis</td>
<td>Male / ambiguous</td>
<td>(See table 3a)</td>
</tr>
<tr>
<td><strong>WNT4</strong> 603490</td>
<td></td>
<td>1p36.23-p35</td>
<td>SF</td>
<td>AD</td>
<td>46,XX</td>
<td>Ovary/ovotestis, ectopic ovary</td>
<td>Variable Müllerian duct abnormalities.</td>
<td>SERKAL, AR</td>
</tr>
</tbody>
</table>

AR, autosomal recessive; TF, transcription factor; AD, autosomal dominant; GOF, gain of function; SF, signaling factor; Del, deletion; Dup, duplication; LOF, loss of function, SERKAL, 46,XX sex reversal with dysgenesis of kidneys, adrenals and lungs

**WNT4** mutations cause variable forms of 46,XX gonadal DSD with variable Müllerian duct abnormalities due to excess AMH production. Homozygous mutations of **WNT4** have been described in the SERKAL syndrome (Female to male Sex Reversal and Kidney, Adrenal and Lung dysgenesis). **WNT4** and **RSPO1** are believed to initiate ovarian development by stabilization of β-catenin (**CTNNB1**), so far only seen in mouse models, but it is shown that patients with SERKAL exhibit downregulated **WNT4** dependent inhibition of β-catenin degradation [37]. Another syndromic form of 46,XX gonadal DSD is autosomal recessive Perrault syndrome, described earlier, caused by mutations in **HSD17B4** (17-beta-hydroxysteroid dehydrogenase IV).
One patient is described with a homozygous gain of function (GOF) mutation affecting the *MAML1* gene, otherwise known to cause DSD in 46,XY. The patient, 46,XX presented with clitoromegaly, primary amenorrhea and one streak gonad with surrounding persistent Wolffian ducts. The other gonad was absent. Uterus and bilateral fallopian tubes were present [38].

46,XX *SRY* positive patients, due to *SRY* gene translocations, can present with testicular or ovotesticular DSD. Also, copy number changes affecting whole genes, or gene regulatory elements of the *SRY*-box genes *SOX3*, *SOX9* and *SOX10* cause 46,XX testicular DSD and/or ovotesticular DSD.

**NR5A1**

The steroid production inducing transcription factor *NR5A1* (Nuclear receptor subfamily 5, group A, member 1), previously called *SF-1* (Steroidogenic factor 1), is important for both male and female sex development. *NR5A1* is expressed in the developing embryo in the urogenital ridge and bipotential gonad, before *SRY* onset in 46,XY. In male sex determination, *NR5A1* is important for *SOX9* up-regulation together with *SRY*. After the transient *SRY* expression has subsided, *SOX9* expression is maintained by *NR5A1* and a *SOX9* feed-back loop.

Post sex determination, *NR5A1* regulates several genes in steroid synthesis such as *CYP11A1* and *CYP21A2*, and is expressed in all primary steroidogenic tissues including adrenal cortex, Leydig cells, and ovarian theca and granulosa cells. *AMH* and *INHA* are regulated by *NR5A1*. *NR5A1* expression is regulated by a promotor region E-box and is activated by the MAPK (mitogen-activated protein kinase) pathway.

Due to its wide function, mutations of *NR5A1* can affect different stages of gonadal differentiation and function, as well as cause adrenal failure. The described spectrum of DSD reported varies from 46,XY complete GD to mild testicular dysgenesis or spermatogenic failure. External genitalia can present with hypospadias, ambiguous genitalia, or female with clitoromegaly. Both point mutations and partial or complete gene deletions have been described, with either autosomal dominant or recessive inheritance. In addition, *NR5A1* mutations have been reported in patients with POI, both PA and SA.

**CBX2**

The possible impact of *CBX2* (Chromobox homolog 2) in human gonadal development was highlighted by Biason-Lauber *et al* reporting a patient with 46,XY ovarian DSD.
compound heterozygous for two \textit{CBX2} mutations [21]. Mice models had previously shown male-to-female sex reversal in XY \textit{Cbx2}\textsuperscript{−/−} knock out (KO) mice, despite an intact \textit{Sry} gene [39]. Gonadal phenotype varied, ranging from ovaries with follicles, indistinct gonads or both testicular and ovarian tissue. Furthermore, XX \textit{Cbx2}\textsuperscript{−/−} mice exhibited gonadal defects, such as smaller ovaries and/or the absence of an ovary, as well as structural damage of oocytes, suggesting a role for \textit{CBX2} also in ovarian development and maintenance, in parallel to \textit{NR5A1} mutations, first only seen in 46,XY.

**Candidate genes in gonadal DSD**

Based on findings in mice models, there are several proposed candidate genes for gonadal DSD, summarized in table 4. Human phenotypic counterparts have not yet been described, and in some cases not identified despite larger studies. The investigated models are often complete KO mice, but in humans haploinsufficiency could be enough to cause a phenotype, as for \textit{NR5A1} mutations causing autosomal dominant gonadal DSD. \textit{Nr5a1} null mice exhibit gonadal and renal defects, while haploinsufficient mice present with isolated renal defects. The \textit{Emx2} deficient mice also display combined renal and genital defects, while the +/- mice are fertile.

Interestingly, the kinase \textit{Map3k4} causes XY sex reversal in both +/- and +/- mice. It is shown to be essential for \textit{Sry} expression, which makes it an interesting candidate gene for investigation in human 46,XY gonadal DSD. Also \textit{Fgf9}, \textit{Fgfr2} and \textit{Lhx9} are interesting candidate genes, but so far human mutations have not been found despite several studies. \textit{Fgf9} deficiency presents with multiple system defects in mice and could perhaps be more suitable for investigation in patients with syndromic DSD. \textit{ZGLP1} is an interesting candidate for both 46,XX gonadal DSD and POI as the mouse model exhibits early blocked germ cell development and that the gene is normally expressed in granulosa cells. The two genes \textit{BAX} and \textit{DDX4} (table 6) are interesting candidates genes for POI, but are mentioned here as XY KO mice also exhibit phenotypic changes. \textit{Bax} deficient mice present with male infertility and disturbed seminiferous structures, \textit{Ddx4} deficient males are infertile.

Most likely, there are unknown genes involved, as well as possible unknown regulatory or synergistic effects between known genes.
### Table 4. Candidate genes for gonadal DSD.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Locus (Human)</th>
<th>Function</th>
<th>Mouse model</th>
<th>Human phenotype</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EMX2</strong>&lt;br&gt;600035</td>
<td>10q26.11</td>
<td>TF</td>
<td>-/- mutants lack gonads, genital tract, kidneys and ureters.</td>
<td>Mutations found in endometrial carcinoma.</td>
<td>Heterozygous mutations in schizencephaly.</td>
</tr>
<tr>
<td><strong>FGF9</strong>&lt;br&gt;600921</td>
<td>13q11-q12</td>
<td>GF</td>
<td>Sex reversal / testicular hypoplasia. Lethal lung hypoplasia.</td>
<td>46,XY GD patients investigated, no mutation found.</td>
<td>Role in angiogenesis and tumorigenesis suggested. Represses Wnt4 stabilization of β-catenin in XY.</td>
</tr>
<tr>
<td><strong>LHX9</strong>&lt;br&gt;606066</td>
<td>1q31.3</td>
<td>TF</td>
<td>-/- mutants lack gonads, normal female phenotype. Nr5a1 expression minimal.</td>
<td>46,XY GD patients investigated, no mutation found.</td>
<td>Binds to Nr5a1 promoter. Additive effect together with Wt-1 in Nr5a1 activation.</td>
</tr>
<tr>
<td><strong>MAP3K4</strong>&lt;br&gt;602425</td>
<td>6q26</td>
<td>Kinase</td>
<td>-/- both +/- sex reversal.</td>
<td></td>
<td>Somatic mutations in endometrial tumors. Map3K4 deficiency disrupt Sry expression.</td>
</tr>
<tr>
<td><strong>PTGDS</strong>&lt;br&gt;176803</td>
<td>9q34.3</td>
<td>Enzyme</td>
<td>Paracrine signal for induction of Sertoli cell differentiation.</td>
<td>Expression in XX gonads in culture causes male development.</td>
<td>Upregulated by Sox9.</td>
</tr>
</tbody>
</table>
| **SOX8**  
| 605923 | **16p13.3** | TF | **-/−** gross normal phenotype, subfertile.  
| **-/−** mutants reproductively normal. | - | - | Associated alpha-thalassemia/mental retardation syndrome.  
| Reinforces Sox9.  
| Binds and activates Amh together with Nr5a1. |

| **ZGLP1**  
| 611639 | **19p13.2** | TF | **Defective sperm development.**  
| Expressed in Leydig cells.  
| Blocked germ cell development at E17.5.  
| Expressed in granulosa cells. | - | - | Transcriptional repressor of Gata6. |

TF, transcription factor; GF, growth factor; GD, gonadal dysgenesis; GFR, growth factor receptor; WD, Wolffian ducts; KO, Knock-out
POI

POI is a very heterogeneous condition and several genes as well as chromosomal aberrations have been shown to be associated to POI, discussed below. As for gonadal DSD, mutations and genomic imbalances can cause POI, both in an autosomal recessive, and autosomal dominant manner. An older study suggests a predominantly autosomal recessive pattern in POI [40].

The ovarian differentiation and follicle organization is a complex process with interaction between cell lines, and is dependent on transcription factors and paracrine signals. Genes causing POI can affect any of these steps, also in a synergistic manner. The first genes causing POI identified were by analysis of translocation breakpoints and terminal X chromosome deletions. A summary of known and candidate genes in POI is given in table 5 and 6, respectively.

Chromosomal aberrations

The most common genetic causes of POI are chromosomal aberrations. Turner syndrome is the most frequent, and is diagnosed by a karyotype performed on lymphocytes from peripheral blood samples. Turner syndrome can be either complete (45,X) or present in different mosaic forms, alternatively involve a structurally abnormal X-chromosome [41].

When an X chromosome is lost in one cell during embryogenesis, all daughter cells will be affected, creating a second cell population. Depending on when this occurs, the patient can be mosaic only in certain tissues. Therefore, it is important that when gonadal biopsies are taken, also tissue samples are analyzed for mosaicism, despite a normal karyotype in peripheral blood.

Other chromosomal aberrations, such as translocations and deletions involving the X chromosome, account for approximately 5-10% of all cases with POI [42].

FMR1 premutation

The FMR1 (fragile X mental retardation 1) gene contains a trinucleotide repeat (CGG) that when expanded (>200) causes Fragile X mental retardation. The normal interval is approximately 5-44 repeats, and expansions of between 55-200 repeats are defined as premutations. The premutation have a potential repeat instability upon transmission. Therefore, in oocytes from a woman carrying a premutation, there is a high risk for repeat number expansion to a full mutation, causing mental retardation in the offspring.
Introduction

The premutation carriers do not have symptom of mental retardation but are at risk to develop fragile X-associated tremor/ataxia syndrome (FXTAS) and/or POI [43].

Premutations of the \textit{FMR1} gene are considered responsible for 2-5\% of sporadic cases with POI presenting with SA [24, 44]. Conversely, premutation carriers have a high POI incidence of 15-20\% [43, 44].

It is not known through which mechanism the expanded \textit{FMR1} gene causes POI. It has been shown that mRNA levels are increased in premutations carriers, with a simultaneous lowered protein level, suggesting an inefficient translation with mRNA accumulation [45]. The \textit{FMR1} gene is expressed in both oocytes and granulosa cells [31], particularly during fetal life where accumulation of elongated mRNA could possibly damage the germ cells. Oocyte depletion has been seen in premutation carriers [46].

Diagnosed patients with POI and \textit{FMR1} expansion must be offered genetic counseling as they are at risk of having children with Fragile-X. In addition, close relatives should be offered investigation as there can be asymptomatic premutation carriers in the family at risk for POI or having children of their own with Fragile X.

\textit{Transcription factors}

The oocyte specific transcription factors \textit{FIGLA} (Folliculogenesis specific basic helix-loop-helix) and \textit{NOBOX} (Nobox oogenesis homeobox) are necessary for normal folliculogenesis (table 5). They are expressed in primordial and growing oocytes and were first described in mouse models. \textit{Nobox} null female mice exhibit atrophic ovaries and are infertile, with decreased expression of the oocyte specific genes \textit{Pou5fl} and \textit{Gdf9}.

\textit{Figla} is important in activation of genes in oogenesis, and also for repression of genes required for spermatogenesis in female mice. Male \textit{Figla} KO mice are in turn infertile, too. In humans, both \textit{FIGLA} and \textit{NOBOX} mutations have been identified in several patients with POI, with suggested autosomal dominant inheritance [47, 48].

Mutations in the granulosa cell expressed transcription factor \textit{FOXL2} (forkhead box L2) cause POI as an associated symptom in the BPES syndrome, mentioned earlier. Larger studies have investigated \textit{FOXL2} in isolated POI, but no causative mutation has yet been found [49].
Table 5. Genes in POI.

<table>
<thead>
<tr>
<th>Gene</th>
<th>OMIM no</th>
<th>Locus</th>
<th>Function</th>
<th>Inheritance</th>
<th>Phenotype</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMP15</td>
<td>300247</td>
<td>Xp11.2</td>
<td>GF</td>
<td>X-linked, dominant</td>
<td>POI with PA and SA.</td>
<td>Forms heterodimers with GDF9. Previously termed GDF9B. Expressed by oocytes in primary follicles.</td>
</tr>
<tr>
<td>FIGLA</td>
<td>608697</td>
<td>2p13.3</td>
<td>TF</td>
<td>AD</td>
<td>POI with SA.</td>
<td>Expressed in all follicular stages.</td>
</tr>
<tr>
<td>FMR1</td>
<td>309550</td>
<td>Xq27.3</td>
<td>Unknown</td>
<td>CGG repeat</td>
<td>POI with SA.</td>
<td>Premutation status (55-200 CGG repeats) associated with risk of POI.</td>
</tr>
<tr>
<td>FOXL2</td>
<td>605597</td>
<td>3q23</td>
<td>TF</td>
<td>AD</td>
<td>POI with SA.</td>
<td>BPES type 1. Upregulated in male Dmrt1 -/- mice.</td>
</tr>
<tr>
<td>FSHR</td>
<td>136435</td>
<td>2p21-p16</td>
<td>HR</td>
<td>AR</td>
<td>POI with PA and SA.</td>
<td>FSHR is not expressed in primordial follicles.</td>
</tr>
<tr>
<td>GDF9</td>
<td>601918</td>
<td>5q31.1</td>
<td>GF</td>
<td>AD</td>
<td>POI with SA, PCOS.</td>
<td>See BMP15. Implicated in dizygotic twinning.</td>
</tr>
<tr>
<td>NOBOX</td>
<td>610934</td>
<td>7q35</td>
<td>TF</td>
<td>AD</td>
<td>POI with SA.</td>
<td>Transactivates GDF9 promotor.</td>
</tr>
<tr>
<td>NR5A1</td>
<td>184757</td>
<td>9q33.3</td>
<td>NR</td>
<td>AR/AD</td>
<td>POI with PA and SA.</td>
<td>Also in gonadal DSD, see table 3.</td>
</tr>
<tr>
<td>PSMC3IP</td>
<td>608665</td>
<td>17q21.2</td>
<td>Several</td>
<td>AR</td>
<td>POI with PA.</td>
<td>Described in one consanguineous family with several affected members.</td>
</tr>
</tbody>
</table>

GF, Growth factor; POI, primary ovarian insufficiency; PA, Primary amenorrhea; SA, secondary amenorrhea; TF, transcription factor; AD, autosomal dominant; BPES, Blepharophimosis, ptosis, and epicanthus inversus syndrome; HR, hormone receptor; AR, autosomal recessive; PCOS, polycystic ovary syndrome ;NR, nuclear receptor; DSD, disorders of sex development

Oocyte secreted factors

Both BMP15 (Bone morphogenetic protein 15) and GDF9 (Growth differentiation factor 9) are members of the transforming growth factor beta super family (TGFB1). They are synthesized as prepropeptides and after processing become dimeric proteins. The expression of GDF9 and BMP15 is almost identical and exclusive to expression in follicles from primordial stage up until ovulation, and it is suggested that they form heterodimers. Heterozygous mutations in BMP15 and GDF9 have been described in patients with POI [50, 51].
Receptors

The FSHR (Follicle stimulating hormone receptor) gene is located on chromosome 2p16.3, close to the LH receptor gene. FSHR is expressed from the primary follicle stage with an increased expression with follicular development. Mutations in the FSHR gene has been described, first in Finnish families with POI, and later mutations have been described in other populations. Variants in the FSHR gene have also been found in patients with spontaneous ovarian hyperstimulation syndrome (OHSS) [52].

Mutations in the nuclear receptor NR5A1 are associated with POI in several reports, with variable reported frequency between studies. So far more than 40 human mutations have been described. Functional studies of several mutations show impaired transcriptional activation of target genes [53, 54], which highlights NR5A1’s importance in gonad development and function.

PSMC3IP

Recently the PSMC3IP (PSMC3 interacting protein) gene was proposed as a novel gene in POI with an autosomal recessive inheritance [55]. A consanguineous family with several affected members with PA was described, where a homozygous 3bp deletion (c.600_602del, p.Glu201del) segregated with the phenotype. The mutant PSMC3IP protein showed a significantly decreased function as an estrogen co-activator, possibly affecting prenatal development of the follicle pool. Estrogen is also important in the second stage of follicular development at puberty and PSMC3IP mutations could therefore possibly cause POI, with either PA or SA.

Candidate genes in POI

There are many proposed candidate genes in POI, an overview is provided in table 6. The majority have been identified in mouse models, some through genome wide association studies (GWAS).

Two counteracting genes, BAX, inducing apoptosis, and BCL2 opposing apoptosis induction, were recently suggested as candidate genes for POI. In mouse KO models, a decreased follicle number has been seen, but for BAX two contradictory studies are reported. One study describes atrophic ovaries, and the other an increased follicle number. In the latter study however the aged mice were infertile, suggesting a possible follicle depletion. Immunohistochemical analysis of human ovarian samples shows expression of both genes in developing follicles. Both BAX and BCL2 should be investigated further in POI.
### Table 6. Candidate genes for POI.

<table>
<thead>
<tr>
<th>Gene</th>
<th>OMIM no</th>
<th>Locus (Human)</th>
<th>Function</th>
<th>Female mouse model</th>
<th>Human phenotype investigated</th>
<th>Association</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BAX</strong></td>
<td>600040</td>
<td>19q13.33</td>
<td>Induces apoptosis</td>
<td>-/ - atretic follicles with excess granulosa cells. -/ - ovaries with 3-fold increase in follicle number but infertility in aged mice.</td>
<td>Human infant/prepubertal ovarian samples.</td>
<td>Expression in all stages of folliculogenesis.</td>
<td>Associated with cancer. Acts on mitochondria to induce apoptosis.</td>
</tr>
<tr>
<td><strong>BCL2</strong></td>
<td>151430</td>
<td>18q21.33</td>
<td>Resists apoptosis</td>
<td>-/ - decreased number of follicles. Deficiency lethal.</td>
<td>Human infant/prepubertal ovarian samples.</td>
<td>Expression in oogonia at gestational week 12-18 before primordial follicle stage.</td>
<td>Possible counteraction to <strong>BAX</strong>. Causes B-cell lymphoma.</td>
</tr>
<tr>
<td><strong>BRSK1</strong></td>
<td>609235</td>
<td>19q13.42</td>
<td>Kinase</td>
<td>-/ - mice normal. Double KO model including <strong>Brsk2</strong> lethal.</td>
<td>POI and natural menopause.</td>
<td>SNP association to POI and age at natural menopause.</td>
<td></td>
</tr>
<tr>
<td><strong>CTNNB1</strong></td>
<td>116806</td>
<td>3p22.1</td>
<td>Junctions</td>
<td>Stabilization of β-catenin in XY mice cause male-to-female sex reversal.</td>
<td>-</td>
<td>Associated with ovarian and endometrial cancers. Also called β-catenin. Regulated by <strong>Rspo1</strong> and <strong>Wnt4</strong>.</td>
<td></td>
</tr>
<tr>
<td><strong>DIAPH2</strong></td>
<td>300108</td>
<td>X21.33</td>
<td>Cytosolic</td>
<td>Cause sterility in D.melanogaster.</td>
<td>POI with SA, familial.</td>
<td>Translocations identified in POI.</td>
<td></td>
</tr>
<tr>
<td><strong>DLX5</strong> and <strong>DLX6</strong></td>
<td>600028 and 600030</td>
<td>7q21.3</td>
<td>TF</td>
<td>Double KO mice premature follicle maturation and reduction in follicles of all stages.</td>
<td>Mutation in <strong>DLX5</strong> associated with AR Split-hand/foot malformation and hearing loss.</td>
<td>Indication that Dlx5 inhibits <strong>Fox2</strong> expression, that in turn upregulates <strong>Fox2</strong> and <strong>Dlx5/6</strong> expression.</td>
<td></td>
</tr>
</tbody>
</table>
**Introduction**

| **ESR1** 133430 | 6q25.1 | TF | -/- mice infertile in single *Ers1a* and *Ers1b* KO. Double KO mice exhibit testicular phenotype in gonads. | POI and women after menopause. Patient with estrogen resistance. | SNP association to POI and age at natural menopause. LOF mutation identified. | Estrogen receptor alpha (ER-α). Infertile male mice, grossly normal reproductive tract. |
| **FOXO1A** 136533 | 13q14.11 | TF | *Foxo1a* regulator of G1/S transition in granulosa cells. | POI with SA. | Synonymous changes identified, not specific for patients. | Associated with longevity. |
| **FOXO3A** 602681 | 6q21 | TF | Early depletion of follicles and secondary infertility after global follicular activation. | POI with SA. | Heterozygous mutations identified, not POI specific. | Evolutionary conserved direct β-catenin interaction. Translocations associated with leukemia. |
| **FST** 136470 | 5q11.2 | GF | Upregulated in early follicle organization. | - | - | *Gdf9* suppresses and *Foxl2* upregulates expression. |
| **GATA4** 600576 | 8p23.1 | TF | Expressed in granulosa and theca cells in preantral and antral follicles. Regulates *StAR Cyp19a1* and *Inha*. | Expressed in granulosa and theca cells in primary to antral follicles, not in primordial. | Mutations seen in patients with congenital heart defects. | Regulated by *Zfpm2* and *Fsh*. Suggested to protect granulose cells from apoptosis in mice model. |
| **GPR3** 600241 | 1p36.11 | Receptor | Premature oocytes maturation, subfertility after superovulation. | POI with SA | No mutations found in two reports. |
| **HK3** 142570 | 5q35.2 | Kinase | - | POI and women after menopause. | SNP association. | Involved in glucose metabolism. |
| **INHA** 147380 | 2q35 | SF | -/- infertile, increased FSH levels. Gonadal and adrenal tumors. | POI with PA and SA. | SNP association. AR and AD POI reported, contradictory results. | Inhibin (INHA-INHB) inhibits pituitary FSH secretion. Activin (INHB-INHB) stimulates secretion. |
| **LAMC1** 150290 | 1q35 | Laminin | - | POI with SA. | SNP association study. | Small sample. |
Table 6. Candidate genes for POI, continued.

<table>
<thead>
<tr>
<th>Gene</th>
<th>OMIM no</th>
<th>Locus (Human)</th>
<th>Function</th>
<th>Female mouse model</th>
<th>Human phenotype investigated</th>
<th>Association</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHCGR</td>
<td>152790</td>
<td>2q16.3</td>
<td>Receptor</td>
<td>Indirect LH effect on follicle ovulation demonstrated.</td>
<td>POI with PA and SA.</td>
<td>Variable symptoms, oligoamenorrhea and infertility.</td>
<td>LH receptor. LOF Leydig cell hyperplasia in XY. GOF male-limited precocious puberty in XY.</td>
</tr>
<tr>
<td>LHX8</td>
<td>604425</td>
<td>1p31.1</td>
<td>TF</td>
<td>-/- lack germ cells. Identical to Sohlh1 KO mice.</td>
<td>POI with SA.</td>
<td>No association found.</td>
<td></td>
</tr>
<tr>
<td>NR5A2</td>
<td>604453</td>
<td>1q31.1</td>
<td>TF</td>
<td>-/- lethal. Binds promoters for Inha, Cyp19, Hsd3b2 and Star.</td>
<td>-</td>
<td>-</td>
<td>Expression is downregulated in TAF4B deficient mice.</td>
</tr>
<tr>
<td>NR6A1</td>
<td>602778</td>
<td>9q33.3</td>
<td>TF</td>
<td>-/- oocyte upregulated Bmp15 and Gdf9, dose dependent hypofertility.</td>
<td>-</td>
<td>-</td>
<td>Represses Pou5f1 in somatic cells. Expressed in primary to preovulatory follicles and preimplantation embryos.</td>
</tr>
<tr>
<td>PDPK1, RPS6</td>
<td>605213, 108460</td>
<td>16p13.3, 9p22.1</td>
<td>Kinase and ribosomal protein</td>
<td>-/- depletion of primordial follicles, POI in early adulthood.</td>
<td>-</td>
<td>-</td>
<td>Oocyte signaling in conjunction.</td>
</tr>
<tr>
<td>PGRMC1</td>
<td>300435</td>
<td>Xq24</td>
<td>Receptor</td>
<td>-</td>
<td>POI with SA.</td>
<td>Translocation and heterozygous mutation in POI.</td>
<td>Mediates progesterone anti-apoptotic effect in ovarian cells.</td>
</tr>
<tr>
<td>POU5F1 (OCT4)</td>
<td>164177</td>
<td>6p21.33</td>
<td>TF</td>
<td>-/- premature apoptosis of primordial follicles.</td>
<td>POI with SA.</td>
<td>No association found.</td>
<td>Expressed in germ cells in early fetal life. Important in stem cells.</td>
</tr>
<tr>
<td>Locus</td>
<td>Chromosome</td>
<td>Gene</td>
<td>Function and Effects</td>
<td>POI or SA</td>
<td>Link to Other Conditions</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>--------</td>
<td>----------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td><strong>PTEN</strong> 601728</td>
<td>10q23.31</td>
<td>Phosphatase</td>
<td>Premature activation and depletion of primordial follicle pool.</td>
<td>POI with SA</td>
<td>No association found.</td>
<td>Associated with cancer. Deficiency is lethal.</td>
<td></td>
</tr>
<tr>
<td><strong>SOHLH1</strong> 601224</td>
<td>9q34.3</td>
<td>TF</td>
<td>+/- mice infertile, lack germ cells, decreased <em>Figla, Nobox</em> and <em>Lkx8</em> expression. +/- mice normal.</td>
<td>Infertile men.</td>
<td>Mutations in men associated with nonobstructive azoospermia.</td>
<td>Abundant expression in ovary after E15.5. Specific expression in ovary and testis.</td>
<td></td>
</tr>
<tr>
<td><strong>TGFBR3</strong> 600742</td>
<td>1p33-p32</td>
<td>GFR</td>
<td>Binds inhibin and increases sensitivity to inhibin.</td>
<td>POI with SA</td>
<td>SNP association study, missense variants described.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gest, gestational; KO, Knock-out; POI, primary ovarian insufficiency; SA, secondary amenorrhea; TF, transcription factor; AR, autosomal recessive; SNP, single nucleotide polymorphism; LOF, loss of function; GF, Growth factor; FSH, follicle stimulating hormone; SF, signaling factor; AD, autosomal dominant; PA, primary amenorrhea; LH: luteinizing hormone; GOF, gain of function; GFR, growth factor receptor.
Novel candidate genes that have not yet been investigated in patient samples are the transcription factors ZGLP1 (table 4) SOHLH1, NR6A1 and the genes DAZL, FST, PDPK1 and RPS6, all with interesting mouse models.

Nr5a2 deficiency is lethal, and Nr5a2 activates expression of several important genes in follicular maturation. NR5A2 expression is possibly regulated by TAF4B as Taf4b KO mice have a decreased Nr5a2 expression. Taf4b encodes a subunit of the RNA polymerase II and KO mice have a disturbed folliculogenesis, and a decreased expression of several factors in follicle function. A GWAS study has recently shown a suggestive association for TAF4B to POI in humans and should it be evaluated further [56]. The same study also suggests ADAMTS19 as a candidate gene, but the association could not be replicated in an independent material and no mouse model exists for this gene.

Other genes identified in association studies are LAMC1 and inhibin alpha (INHA). An Inha KO mouse model exists with infertile mice with elevated FSH levels, but in humans, the association to POI is not yet proven and results are contradictory [57]. The receptor TGFBR3 binds inhibin and increases cell sensitivity to inhibin, suggesting a potential role in hormonal signaling. Four studies of TGFBR3 in POI are reported, three describing missense variants [58-60], but so far no functional study has been performed.

The luteinizing hormone (LHCGR) and estrogen receptor 1 (ESR1) are interesting candidates for POI given the role of the LH and estrogen in ovulation and follicle maturation. However, identified mutations in LHCGR cause variable menstrual irregularities with normal FSH levels [61]. A patient with estrogen insensitivity and a ESR1 LOF mutation is described, presenting with PA but with normal FSH levels [62]. One GWAS study however proposes an association between a single single nucleotide polymorphism (SNP) in ESR1 and POI, as well as between SNPs in the genes BRSK1 and HK3 and POI [63]. Unlike ERS1 and BRSK1, HK3 has no supporting mouse model and must be considered a less likely candidate despite a statistically significant association.

An early investigation of translocation breakpoints identified a disruption of the DIAPH2 gene in a mother and daughter with POI. However, mutations in this gene have not been found in other POI patients.
The *POF1B* gene was also described early, with autosomal recessive (AR) POI in a Lebanese mother and daughter. No other patients with *POF1B* mutations have been reported.

The transcription factors *LHX8, POU5F1, FOXO1A, FOXO3A* and the genes *PTEN* and *GPR3* have been investigated in patients with POI without identification of any patient specific sequence variant. Neither is any *CTNNB1/β-catenin* mutation or association reported in POI so far.

A proposed schematic overview for sex determination and signaling in gonad development, including candidate genes from mouse models, is given in figure 2.

**Figure 2.** Proposed model including candidate genes.

Names in italics symbolize genes. Names in smaller normal font represent protein, either enzyme or hormone. Names surrounded by black box represent transmembrane receptors. Cell type specific expression is indicated by circle. Positive interaction is indicated by arrow, negative by T-sign. Positive interaction involving hormones is indicated by dashed line.
Introduction

**Tilting the balance**

The mechanism how the expression of the sexually dimorphic genes control sex determination and maintain it is not completely understood. However, data suggest an important role for $\beta$-catenin in this process. Murine studies have shown that Rspo1 induces $\beta$-catenin signaling in the developing ovary, and that when stabilized maintains the tissue fate [8]. In fact, expression of a stable form of $\beta$-catenin will activate the WNT pathway in XY gonads [64]. The $\beta$-catenin expression is inhibited in the male pathway, possibly through MAP3K1. This as mutations of MAP3K1 have been shown to cause increased $\beta$-catenin expression in a human teratoma celline expressing testis determination genes. Simultaneously, SOX9 was down regulated [65]. The expression of $\beta$-catenin is however not responsible for this down-regulation of SOX9, although a loss of Wnt4 or Rspo1 causes partial sex reversal in XX mice. Instead, FOXL2 has been proposed to be the counteracting force of SOX9. Also, loss of Foxl2 expression will cause mature supporting cells of the ovary to develop into testicular somatic cell types [66]. This suggests that the sexually dimorphic expression patterns need to be maintained even after sex determination, and all trough adult life. In addition, the genes involved in sex determination are not expressed completely in an on/off fashion depending on gonadal fate, but can be important in both. For instance, Rspo1/WNT/ $\beta$-catenin expression is found in spermatogenesis [67] and Dhh signaling occurs in ovaries [68].
AIMS

The overall aim of this thesis was to identify and study candidate genes in gonadal DSD and POI. This was done in order to achieve a better understanding of the underlying molecular mechanisms, and to improve the available diagnostic tools and genetic counseling available for affected patients and family members.

The specific aims of the thesis were

- To identify new candidate genes necessary for normal testicular development by investigating submicroscopic genomic imbalances in patients with 46,XY gonadal dysgenesis.
- To screen for mutations in the CBX2 gene as a cause of 46,XY and 46,XX gonadal DSD and POI.
- To identify new candidate genes necessary for ovarian development and or maintenance, by studying submicroscopic genomic imbalances in patients with idiopathic POI.
- To sequence the PSMC3IP, FIGLA, NOBOX and GDF9 genes to evaluate the impact of mutations in these genes in a cohort of Swedish patients with POI.
- To develop new diagnostic tools and clinical genetic investigation guidelines based on findings from the project.
MATERIAL AND METHODS

PATIENTS

Patients were recruited among the patients referred to the Department of Clinical Genetics at Karolinska University Hospital, Stockholm, Sweden. The majority of patients are under clinical care at the Department of Obstetrics and Gynecology at the Karolinska University Hospital. All participants gave informed consent. In the case of minors, consent was obtained from legal guardians. The study was approved by the regional Ethics Committee at Karolinska Institutet, Stockholm, Sweden with approval numbers 2007/263-31/2 and 2011/276-32.

The patients participating in this study can be divided into the following groups based on the clinical diagnosis.

- Patients with 46,XY complete GD
  Women with a 46,XY karyotype, unambiguous female external genitalia and uterus were investigated to identify novel candidate genes in testicular development (paper I) and for mutations in the $CBX2$ gene (paper II).

- Patients with POI
  Women with PA, early SA and familial cases have been of particular focus. The patient cohort was investigated for novel candidate genes causing POI by array-CGH (paper III), and for specific gene mutations (paper II and IV).

- Patients with other forms of gonadal DSD
  Patients with 46,XY ovotesticular DSD, 46,XX testicular or ovotesticular DSD and one 46,XX patient with one streak gonad and one normal functioning ovary, were included in the investigation of the $CBX2$ gene (paper II).

The types of patient samples used in the studies were

- Genomic DNA
  extracted from peripheral blood lymphocytes, EBV-immortalized lymphocytes and mixed tissue from gonadal biopsies.

- RNA
  extracted from EBV-immortalized lymphocytes, and fibroblasts cultured from gonadal biopsies.
Clinical characterization
The clinical characterization of patients is based on phenotype, hormonal profile and karyotype. In addition, family history is investigated. For patients with POI, external causes such as surgery, radiation therapy or autoimmunity were excluded. The clinical characterization was performed by the patient’s gynaecologist, and the information was collected at the time of inclusion in the study.

The clinical diagnosis of 46,XY gonadal dysgenesis is based on a XY karyotype, female external genitalia, internal Müllerian structures (uterus) and hypergonadotropic hypogonadism (FSH >30 IU/L). There are two pairs of affected siblings in the patient group.

The diagnosis of POI is established by occurrence of PA or SA in a woman with female external genitalia as well as internal Müllerian structures and hypergonadotropic hypogonadism. In the studied group with SA, the overall median age at diagnosis was 22 years (12-37). Several of these patients presented with amenorrhea after oral contraceptive use, possibly delaying the diagnosis.

The clinical diagnosis of 46,XY and 46,XX ovotesticular DSD is based on histological examination of gonads, karyotype and phenotype of external and internal genitalia. The diagnosis of 46,XX testicular DSD is based on karyotype, phenotype of internal and external genitalia and hormonal profile.

The patient with a 46,XX karyotype, one streak gonad and one normal functioning ovary was discovered during abdominal surgery. The patient has normal external and internal female genitalia and normal female hormonal profile.

Histological examination of gonads is recommended for all patients if possible.

Clinical genetic analysis
Clinical genetic investigation is commenced with conventional karyotyping that can detect aneuploidy, structural chromosome aberrations and mosaicism. In the studied patients, low level sex chromosome mosaicism has been excluded using fluorescent in situ hybridization (FISH) on peripheral blood smears and when available on touch preparations from gonadal tissue. For patients with 46,XY gonadal DSD, sequencing of SRY, NR5A1, WT-1, SOX9 and DHH has been performed. In addition, MLPA analysis targeting the 9p24 region [69] and NR5A1, SRY, SOX9, WNT4, DHH, WT1, NR0B1 among other DSD genes [70] has been performed.
Material and methods

For patients with POI, sequencing has been performed for NR5A1, BMP15, and FSHR, and MLPA analysis for NR5A1. In patients with SA, FMR1 gene premutations have also been excluded.

**CONTROLS**

In paper II, DNA samples from anonymous blood donors at the Karolinska University Hospital, were used as normal controls in the investigation of novel sequence variants.

DNA from EBV-immortalized cell lines from healthy fertile men and women, were used as normal controls for probe validation experiments in MLPA set up (paper II and III).

For paper III, genomic DNA from 95 women without POI was collected as control samples. All women were above the age of 40 at blood sample collection and had given birth to at least one child. Exclusion criteria for participation were previous egg donation, *in-vitro* fertilization, fertility treatment or menopause before the age of 40.

**DNA SEQUENCING**

Novel candidate genes identified by array-CGH (C2ORF80, KRTAP2-3, KRTAP2-4, SUPT3H, TSPYL6, paper I and III), known gene for gonadal DSD (CBX2, paper II) and for POI (FIGLA, GDF9, NOBOX, PSMC3IP, paper III and IV) were sequenced to identify possible causative mutations. All exons, exon/intron boundaries and parts of 5’UTR and 3’ UTR were included. Primers for PCR amplification were designed using Primer3Input.

PCR conditions were optimized for each amplicon. PCR fragments were separated using agarose gel electrophoresis to ensure correct band size, and subsequently cleaned by Exonuclease I and Shrimp Alkaline Phosphatase enzyme reaction (both Thermo Scientific) to remove unconsumed deoxyribonucleotides (dNTP) and primers.

The sequencing reaction was performed using the BigDye terminator kit 3.1 (Applied Biosystems), based on Sanger methodology with fluorescent chain terminator dideoxynucleotides (ddNTP). Both DNA strands were sequenced using the PCR primers. Fragments were size separated and analyzed using an ABI 3730 XL capillary sequencer (Applied Biosystems). Electropherograms were analyzed, and compared to the corresponding reference sequence using the SeqScape version 2.5.0 software (Applied Biosystems).
REVERSE TRANSCRIPTASE-PCR

Reverse Transcriptase-PCR (RT-PCR) was used to investigate the expression of different mRNA isoforms for the *SUPT3H* and *CBX2* genes (Paper I and II) and to confirm and characterize a duplication within the *SUPT3H* gene (Paper I).

Total RNA was extracted from EBV-transformed lymphocytes and gonadal fibroblasts using the RNeasy kit (QIAGEN) including the DNAse digestion step. First strand synthesis, the conversion of mRNA to complementary DNA (cDNA), was performed using the First-Strand cDNA synthesis kit (Amersham Biosciences) and random hexamer primers. The DNA/RNA strand was subjected to PCR reaction using specifically designed primer pairs for the target of interest.

For *SUPT3H* duplication analysis, a specific primer pair was designed to amplify an mRNA containing an interstitial duplication where exon 12 was followed by exon 5. PCR amplified fragments were sequenced to verify that the right fragment was amplified and to determine the duplication junction.

For *CBX2* expression analysis, primer pairs were designed within isoform specific exons, and to include at least 2 exons for ensuring isoform and cDNA specific amplification.

ARRAY COMPARATIVE GENOMIC HYBRIDIZATION

Array comparative genomic hybridization (array-CGH) is a technique that allows the detection of copy number differences between two DNA samples, of which one can be a chosen reference control. The array consists of a glass slide with genomic fragments printed on one side. We have used a customized array with one million oligomarkers. The oligomarkers has a genome-wide distribution, with exclusion of repetitive DNA sequence regions, leading to an average theoretical resolution of 2.2kb. For 78 genes known to be involved in gonadal development, there are additional probes to the backbone coverage (3 probes per exon), increasing the resolution to such extent that partial gene deletion or duplications can be detected (paper I and III).

DNA quality is essential for optimal array-CGH quality control metrics (the probe to probe log ratio, or DLR spread). Therefore, all samples were controlled for quality, and purified using the QiAmp DNA minikit (QIAGEN) when appropriate to achieve acceptable quality values in analysis by Nanodrop 1000 spectrophotometer (260/280 1.8-2.0, 260/230 >2).
Material and methods

Procedure

Sample and control DNA are labeled using red (Cy5-dUTP) and green (Cy3-dUTP) fluorescent signals, respectively. The same amount of genomic DNA is used for both samples. Labeled samples are pooled together with hybridization buffer and Cot-1 DNA that blocks repetitive DNA sequences. The mixture is then pipetted onto a gasket slide that is pressed tightly against the array glass, and then placed in a rotating chamber inside a hybridization oven for 48h. A competitive hybridization of patient and control DNA to the oligomarkers on the array takes place. The array slide is then removed and washed to remove unhybridized material, and the fluorescent signal at each oligomarker is measured. Depending on the signal ratio between the red and green fluorescence, it is possible to determine the ratio of genomic material between the two samples. See figure 3 for an overview of the procedure.

Figure 3. Principles of Array-CGH.
Material and methods

In our studies we have used two commercial control samples of pooled DNA from 10 healthy males or females. For the study with 46,XY GD patients the male reference was used, while the female reference was used for the study with POI patients. The same reference DNA batch was used throughout each study.

Interpretation

Fluorescence data from laser scanning was analyzed using the Cytosure interpret software v3.4.3 (OGT, Oxford, Great Britain). After global normalization and exclusion of non-uniform outliers (10% outlier removal), circular binary segmentation analysis to detect copy number changes was performed with a minimum consecutive probe count of 5, a threshold for gains of 0.35 and a threshold for losses of 0.65.

Copy number variants

Deletions or duplications of genomic material are called copy number variants (CNV). The term CNV however is often used to denote the normally and frequently occurring copy number changes found in healthy controls, even though a CNV can represent a unique and causative change in one patient. International research data regarding CNVs in controls is collected and presented in the online Database of Genomic Variation (DGV). By comparing patient findings with this data, we have excluded common CNVs found in healthy control samples from further investigation. This exclusion was done with caution, as phenotypic effects of genetic imbalances affecting sexual development can be dependent on chromosomal sex. This information is not always available for the controls included in the DGV. Also, aberrations only partially overlapping with rare reported CNVs were not excluded. Small intronic variations and intergenic changes were excluded after verification that they were not located within or upstream/downstream a known gene causing DSD, as a positional effect could then be considered. All remaining novel copy number changes were confirmed by MLPA.

Limitations

The array-CGH method only distinguishes between ratio differences, not absolute copy numbers. If both control and sample carry the same number of alleles, regardless of number, no signal offset will be detected. The resolution of detected differences is dependent on type of array platform used, where number and distribution of probes vary between platforms. Copy number changes are also detected by SNP-array, and mixed arrays with both SNP and oligomarkers.
Positional information is not given by array-CGH. Therefore, balanced translocations cannot be detected as there is no change in copy numbers. In addition, detected duplications can theoretically be inserted anywhere in the genome. In paper I and II we have characterized two duplications and shown that they are placed in tandem.

**MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION**

Multiplex ligation-dependent probe amplification (MLPA) is a PCR based quantitative technique that allows for simultaneous detection of copy number changes of several target sequences in one reaction. Probes for target sequences can be placed to study several genes, or single exons within a gene, depending on design. The method was first described by Shouten et al [71] and has provided a valuable addition to the available cytogenetic techniques for detection of copy number changes in the size interval between detection by PCR and conventional cytogenetics.

In this thesis, MLPA has been used for verification of array-CGH findings (paper I and III), as well as for candidate gene studies. Two MLPA probes sets using synthetic probes have been developed for the detection of entire and/or partial gene copy alterations of the CBX2 and GDF9 genes (paper II and III).

**MLPA probes**

MLPA probes are synthetic oligonucleotides designed as pairs, with a left probe oligo containing the forward primer sequence tag and the left hybridizing sequence, and a right probe oligo containing the right hybridizing sequence and the reverse primer sequence tag (Figure 4a). The right probe oligo has a 5’ phosphor group, essential for ligation. The primer sequence tags are identical in all probe oligos and allow for simultaneous PCR amplification of all ligated probes with only one primer pair. MLPA probe oligos used together in one MLPA reaction are designed to create, after ligation, fragments of different size. In our experiments, a minimum of 3bp in size difference has been used to allow for good separation in the capillary electrophoresis apparatus (ABI 3100 genetic analyzer, Applied Biosystems). The probes were designed according to the recommendations by Stern et al [72] and SNPs where avoided at least near the ligation site. Target specificity was controlled using the BLAT function at the UCSC genome browser. Probe characteristics such as GC content and melting temperature (Tm) was obtained using the RawProbe software.

The probe oligo used in the MLPA probe sets for CBX2 and GDF9 analysis were validated for consistency by analysis of 10 healthy controls of both sexes, with a
standard deviation below 0.1 for each probe deemed acceptable. The control and pilot probe used in experiments have been described earlier [73].

**Procedure**

The synthetic oligonucleotide probe set is mixed with genomic DNA and denatured. Then hybridization of left and right probe oligos to single stranded DNA is allowed for 16h at 60°C (Figure 4b).

Left and right probe oligos adjacently hybridized to the DNA are ligated together at 54 °C to form the complete MLPA probe. Unhybridized, and unpaired probe oligos will not be ligated. Ligated MLPA probes are then amplified by a multiplex PCR using a universal primer pair. The forward primer is fluorescently labeled, allowing quantification during capillary electrophoresis separation of the different PCR fragments. The fluorescence allows for quantification of PCR products by measuring the area or height of each peak corresponding to the targeted sequence.

**Figure 4.** Principles of Multiplex Ligation-dependent Probe Amplification (MLPA).
Material and methods

Controls
For comparative quantification of copy number changes, the MLPA method is dependent on the use of internal reference probes. In our MLPA experiments, a minimum of at least three control probe pairs with stable amplification has been included in each probe set. In a process of internal control probe normalization, each sample’s peak height and area is normalized to the average peak height and area of the reference probes. The normalized values are then compared to the average of normalized peak height or area of control samples included in the run. A minimum of two appropriate control samples has been used in MLPA experiments.

Interpretation
The comparative quantification after normalization gives results in ratio values between sample and control. A ratio value of 1 represents equal DNA amount in the sample and the control for the specific target sequence of the MLPA probe. For a sample run to be deemed acceptable in our experiments the ratio values for the reference probes has to be between 0.8 and 1.2. For the target of interest a threshold value for deletion and duplication detection should be set. The analysis of MLPA results was initially made in an Excel spreadsheet and later using the GeneMarker v1.90 (Soft Genetics) software. Threshold values for deletions and duplications were set at 0.75 and 1.25 in Excel, and 0.75 and 1.3 in GeneMarker, respectively. The values differ as the GeneMarker software uses a more complex statistical computation of the results changing the duplication cut-off value.

All detected deletions and duplications were verified by at least two independent MLPA runs and if available, a separate DNA sample from the patient was also investigated for confirmation.

IN SILICO ANALYSIS, ONLINE RESOURCES AND TOOLS
The methods used in this thesis, and the interpretation of results were dependent on several programs, tools and databases. For an overview including URL or manufacturer, see appendix II.

As general reference source the website for the National Center for Biotechnology Information (NCBI) was used, including its constituent databases. In addition to PubMed, the database for single nucleotide Polymorphisms (dbSNP)[74] at NCBI should be mentioned.
The UCSC (University of California Santa Cruz) genome browser [75] with the alignment tool BLAT [76] were frequently used for all studies. Also Ensembl [77] and GeneCards were used frequently for general gene information. Software for PCR primer and MLPA probe design was used or downloaded online, respectively [78, 79], whereas commercial software for sequence, MLPA and array-CGH analysis was used.

For interpretation of array-CGH findings the Database of Genomic Variants (DGV) v10 released Nov 2010 [80, 81] and DECIPHER (Database of chromosomal imbalances and phenotype in humans using Ensembl resources) [82] were used.

Data regarding expression patterns of genes and proteins was extracted from the general sources mentioned above, as well as specific online databases such as Human Protein Atlas [83], Database of Gene expression profiles during sex determination [84] and the LifeMap Database of Embryonic development, Stem Cell Research and Regenerative Medicine [85].

Protein and nucleotide alignment was performed using ClustalW2 and ClustalWomega. Prediction of amino acid substitution on protein function was performed using the commercial software AlaMut and online tools Polyphen2 [86], SNP&GO and SIFT [87]. The prediction of nucleotide substitution on splicing efficiency was performed with AlaMut, as well as online tools NNSPLICEv0.9 [88], NatGene2v2.4 [89] and FSPLICEv1.0.

**STATISTICAL ANALYSIS**

In paper IV, binomial probability calculation is performed for estimation of statistical probability of negative findings. The binomial parameters were calculated using StatXact 4 (CYTEL Software Corporation, Cambridge, USA). 95% confidence interval was calculated according to Clopper-Pearson.
RESULTS AND DISCUSSION

CLINICAL ANALYSES (UNPUBLISHED DATA)

Patients referred to the Department of Clinical Genetics at Karolinska University Hospital, Stockholm, Sweden, for genetic investigation of gonadal DSD and POI have been considered for this thesis. For these patients, the initial investigation of choice is a conventional karyotype. If an aberration is identified, this is often diagnostic. Further investigations that can be offered include:

- **46,XY gonadal DSD**: Sequence analysis of SRY, NR5A1, WT-1, SOX9, DHH and MLPA analysis detecting aberrations in NR0B1, SOX9, SRY, WNT4 and NR5A1.

- **46,XX gonadal DSD**: Sequence analysis of SRY and NR5A1. MLPA analysis can be considered depending on clinical presentation.

- **POI**: Sequence analysis of BMP15, FSHR and NR5A1. FMRI analysis in SA.

Some patients have received a molecular diagnosis by routine clinical genetic investigation or in previous research studies. An overview is given in table 8. The remaining patients have been included in the studies in this thesis.

**XY gonadal DSD**

Three patients with 46,XY GD have been diagnosed with SRY point mutations, and one patient has a diagnostic SRY deletion. One patient has a missense mutation in the WT-1 gene considered causative. Also different forms of chromosomal mosaicism have been found in three patients. For the remaining 14 patients with GD, clinical genetic investigations have not revealed any causative mutation. Among the three patients with other gonadal DSD forms, one patient with mixed GD has a causative dicentric Y-chromosome rearrangement.

Previous research studies in the group using BAC array and MLPA analysis, have identified a pair of affected siblings carrying NR0B1 duplications [90], detected before the development of the MLPA set now used at the Department of Clinical Genetics. Also one patient with a causative 9p deletion has been identified [69].

The overall diagnostic rate in 46,XY GD can be estimated to 44% in our material. Previous reports vary between 20-50% [15].
Table 8. Results of clinical genetic investigations.

<table>
<thead>
<tr>
<th>Clinical diagnosis</th>
<th>Patients studied (n)</th>
<th>Total (n)</th>
<th>Genetic diagnosis (n)</th>
<th>Diagnosis established (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>46,XY complete GD</td>
<td>14</td>
<td>25</td>
<td>SRY mutation/deletion (4)</td>
<td>44%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WTI mutation (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45,X/46,XY (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45,X/47,XY (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NR0B1 duplication (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Del9p (1)</td>
<td></td>
</tr>
<tr>
<td>46,XY gonadal DSD (other forms)</td>
<td>2</td>
<td>3</td>
<td>45,X/46,X,idic(Y) (1)</td>
<td>33%</td>
</tr>
<tr>
<td>46,XX gonadal DSD</td>
<td>3</td>
<td>4</td>
<td>SRY positive (1)</td>
<td>25%</td>
</tr>
<tr>
<td>POI with PA</td>
<td>23</td>
<td>24</td>
<td>46,XX, t(X;14)</td>
<td>4%</td>
</tr>
<tr>
<td>POI with SA</td>
<td>31</td>
<td>35</td>
<td>FMR1 premutation (1)</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46,XX,del(X) (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46,XX, t(X;4) (1)</td>
<td></td>
</tr>
</tbody>
</table>

GD, gonadal dysgenesis; Del, deletion; DSD, disorders of sex development; POI, primary ovarian insufficiency; PA, primary amenorrhea; SA, secondary amenorrhea.

**XX gonadal DSD**

One patient with 46,XX testicular DSD has been diagnosed as SRY positive, explaining the phenotype. The remaining patient with 46,XX testicular DSD, 46,XX ovotesticular DSD and the patient with one streak gonad and one functioning ovary, have not received a molecular diagnosis.

**POI**

Four patients with POI had diagnostic aberrant karyotype with either X-chromosome deletion or translocations involving the X-chromosome. One had PA and three had SA. In three patients, balanced translocations involving autosomes have been identified, but not considered causative. One of these patients has participated in the studies included in this thesis. The remaining two patients have not been available for inclusion.

One patient with SA has been diagnosed with a causative FMR1 premutation. This corresponds to a detection rate of 1/54=2.8% which is expected for unrelated cases (2-5% in previous reports) [24, 44].

In total, 54 patients with POI were included in the study, 23 with PA and 31 with SA. All patients with POI participating in the studies of this thesis have undergone sequence
Results and discussion

analysis of the BMP15, FSHR and NR5A1 genes. No causative mutation has been identified, although three novel changes have been found.

One patient is a heterozygous carrier of a novel silent change in the BMP15 gene (c.393T>C), inherited from the mother. A novel silent change was found in the FSHR gene in another patient in heterozygous form (c.1653C>T), no parental samples were available. The silent changes were not predicted to affect splicing.

In NR5A1 a missense change was identified in one patient with SA, c.771C>G, p.Asp257Glu. The patient is heterozygous, and the mother does not carry the change. Paternal samples were not available. The amino acid substitution occurs in a nuclear hormone receptor ligand binding domain, but the substitution is, based on prediction tools and the similarity between exchanged amino acids, considered benign.

These patients are included in study II, III and IV of this thesis.

Conclusion

The diagnostic rate is higher for the group with 46,XY gonadal DSD than for 46,XX gonadal DSD and POI. This can be due to differences in mutation frequency of causative genes in 46,XY and 46,XX respectively, reflect that POI is a more heterogeneous condition, or illustrate that the group with 46,XX gonadal DSD and POI is far less studied than 46,XY.

Through previous research in the group studying 46,XY GD, three patients could receive a molecular diagnosis. This was the rationale to extend the study to also include 46,XX patients and to further analyze the undiagnosed 46,XY patients in the current project.

A karyotype is always necessary for the analysis of patients with gonadal DSD and POI as it is the basis of classification of DSD forms, and can also be diagnostic alone, such as in the case of mosaicism or translocations involving the X chromosome. For patients with SA, FMR1 analysis must be recommended, especially in familial cases. It is a simple and affordable analysis, and there can be undetected younger carriers in the family at risk for POI. In addition, premutation carriers have an increased risk of having children with Fragile-X mental retardation, and must be offered genetic counseling.

Today, there are several known genes in gonadal DSD and POI to investigate in patients. A suitable genetic investigation must be recommended for giving affected patients a molecular diagnosis. Even if not influencing therapy recommendations, a molecular diagnosis will provide the patient with an explanation, which can be some
consolation in a difficult condition. A summary of the clinically relevant genes as of today is given in the concluding remarks.

**GENOME WIDE APPROACH TO IDENTIFY NOVEL CANDIDATE REGIONS IN 46,XY GONADAL DYSGENESIS (PAPER I)**

Attempting to identify new potential candidate regions involved in testicular development, we have investigated patients with complete 46,XY GD for submicroscopic genetic imbalances using a genome wide approach. The process of sex determination is sensitive to gene dosage, both to haploinsufficiency and gene duplication, and genes affecting gonadal development are located not only on the sex chromosomes but also on autosomes.

DNA samples from nine unrelated patients with 46,XY GD, of which one also has an affected 46,XY sister, were analyzed by array-CGH. A 1 M array-CGH platform with 2.2 kb probe spacing for whole genome coverage and enrichment targeting 78 genes involved in gonadal development was used.

Four of the nine analyzed patients had novel copy number alterations, with a total of five rearrangements. Of these, one is a previously described benign duplication of the **SRD5A2** gene, paternally inherited [73]. Four additional patients with 46,XY GD and one with ovotesticular DSD were included in follow-up studies of novel regions. The four most interesting candidate genes identified in the study are discussed below.

**SUPT3H on 6p21.1**

A duplication within the **SUPT3H** gene (suppressor of Ty 3 homologue *S. cerevisiae*) and a deletion affecting the **C2ORF80** gene, discussed in the next paragraph, were detected in the affected pair of siblings. Both changes are inherited from the healthy mother. It is possible that one of these changes is a benign variant, or that the causative mechanism is by an interacting or additive effect. However, at the moment no common pathways are known between SUPT3H and C2ORF80.

The duplication within the **SUPT3H** gene was further characterized by MLPA and RT-PCR, and shown to be an interstitial tandem duplication of exon 5 through 12, placing exons in the order 1-12 followed by 5-13 (figure 5). The duplication affects both major isoforms of **SUPT3H** and causes a likely C terminal amino acid change. The SUPT3H protein is well conserved in mammals and it is difficult to predict if a terminal change affects protein function or stability. SUPT3H expression is high in testicular cells.
Figure 5. SUPT3H duplication.

A, SUPT3H gene, exon organization. Green exons are not present in mRNA variant 1. Exon 5 and exon 12 have been colored in blue and yellow respectively, to help to understand the duplication structure.; B, SUPT3H mRNA variants. The exons in gray are not coding. ATG indicates the initiation of translation.; C, Duplication structure.

(Leydig, interstitial and germ cells), and seminiferous ducts. It is a transcriptional co-activator [91] and the human homolog of the yeast transcription factor Spt3 in which mutations cause defects in mating and sporulation [92].

C2ORF80 on 2q34

The deletion of the C2ORF80 gene (chromosome 2 open reading frame 80) spans the first eight of a total of nine exons, thus removing one allele. The gene is poorly characterized. C2ORF80 is moderately expressed in testicular cell types (Leydig cells and cells in seminiferous ducts) and ovarian stroma cells, among others. The protein does not contain any known functional domain, and does not present homology with any known protein. It is highly conserved in other species down to frog and fish, with a 47% sequence homology with D.rerio.

We investigated all available patients with 46,XY gonadal DSD by sequence analysis, to detect inactivating mutations of SUPT3H and C2ORF80, although none were found. Due to the small group size this negative results must be interpreted with caution, and
we consider both *SUPT3H* and *C2ORF80* as interesting candidate genes for 46,XY GD.

**PIPK1B and PRKACG on 9q21.11**

An unrelated patient carried a 454 kb duplication within 9q21.11, possibly affecting two interesting candidate genes, *PIPK1B* and *PRKACG*.

The duplication spans the last four exons of the *PIPK1B* (Phosphatidylinositol-4-Phosphate 5-kinase, Type1, Beta) gene. Positional data cannot be inferred by array-CGH and it is possible the patient has two normal alleles. The duplication can also disrupt one allele.

The *PIPK1B* gene is highly conserved in the mammalian clade, and has a 67% sequence homology with *D.rerio*. The *PIPK1B* enzyme function is not completely understood, but it is an active signaling molecule involved in cytoskeleton reorganization, cell survival, epithelial cell morphogenesis and more [93].

Mouse *Pip5Ka* (corresponding to human *PIPK1B*) has a very high expression in testis and in germ, Sertoli, and Leydig cells [93]. *Pip5Ka* and *Pip5Kb* (corresponding to human *PIPK1A*) double deficient mice are completely sterile due to lack of germ cells. *Pip5Ka KO* mice are subfertile due to reduced motility and abnormal morphology of sperm. Based on these findings, the potential role for *PIPK1B* in human gonadal development must be further evaluated.

The *PRKACG* gene (homo sapiens protein kinase, cAMP-dependent, catalytic, gamma) encodes the gamma form of the catalytic subunit of cAMP-dependent kinase, a retrotransposon derived from the alpha form. In contrast to the widely expressed alpha and beta forms, the gamma form is highly tissue-specific to human testis, and germ cells in particular [94]. We consider *PRKACG* an interesting candidate gene for XY GD.

**Conclusion**

By array-CGH we have identified novel candidate genes in 46,XY GD. Proving causality for identified variants is challenging and experimental functional characterizations should ideally be undertaken. The identification of recurrent or related genetic variants in other groups of patients with well documented phenotypes could also provide support for the potential role of these genes in gonadal DSD.

All studied patients had undergone MLPA analysis for detection of dosage imbalances in several known gonadal DSD genes [69, 73, 95]. Therefore it was not surprising that
Results and discussion

no such rearrangements were detected here. However, array-CGH is a powerful technique with high potential resolution, capable of detecting diagnostic genomic imbalances, as well as novel candidate genes, and we recommend that all patients with 46,XY GD without a molecular diagnosis should undergo copy number change analysis.

CBX2 CANDIDATE GENE STUDY (PAPER II)

The purpose of this study was to investigate the presence of point mutations and gene copy number changes in the CBX2 gene in a group of patients with gonadal DSD. One reported patient, compound heterozygous for CBX2 mutations and with 46,XY ovarian DSD highlighted the potential role for CBX2 in human gonadal development. In murine models, Cbx2 KO mice exhibit XY and XX gonadal disturbances, including germ cell instability in XX gonads. Therefore we proposed CBX2 as a candidate gene for 46,XY and 46,XX gonadal DSD, as well as for POI. A cohort of 47 patients with different gonadal DSD forms or POI were included in the study. We used DNA sequencing and an in-house developed MLPA probe set for specific identification of deletions/duplications affecting CBX2.

Sequence alterations

10 changes, nine single nucleotide variations and one insertion were identified. Nine are reported SNPs. A novel silent G to A transition (c.1356G>A) was identified in heterozygous form in a patient with POI and PA, maternally inherited. Both alleles were expressed in gonadal fibroblast cultured from the patient. No effect on splicing could be predicted by three different splice site prediction tools. The nucleotide change was not identified in 156 control alleles, however we believe it is a rare normal variant. Its maternal inheritance in the patient argues further against a pathogenic role in POI when in heterozygous form.

All patients were found to carry the G allele for SNP rs71368052 (c.-52C>G) in the CBX2 promotor. This is likely an annotation error as the available data from two population studies in dbSNP reveals only G allele carriers.

A non synonymous SNP, p.Pro471Ala was identified in a patient with PA, inherited from the healthy father. Both alleles were expressed in gonadal fibroblasts from the patient, excluding a second mutation affecting gene expression. Pro471 is conserved in eutherian mammals (figure 6). It is not in one of the highly conserved domains of CBX
proteins [96, 97], but it is the first amino acid after the conserved CBX2-specific motif Cx2.2 [97], making it difficult to predict a specific functional role. It is reported as a

**Figure 6. CBX2 amino acid conservation.**

Partial alignment of the CBX2 protein. Pro471 and corresponding residues in other species is highlighted in yellow. Blue and green show part of the functionally important Cx2.2 and Pc box, respectively, in human CBX2. Consensus symbols are indicated considering only mammalian (nine top) or all species proteins. An * (asterisk) indicates positions with a fully conserved residue; a : (colon) indicates conservation with residues of strongly similar properties; a . (period) indicates conservation with residues of weakly similar properties.

very rare SNP and was not detected in 156 control alleles in our study. Causative mutations for rare recessive diseases are expected to be listed as very rare SNPs as there will be healthy heterozygous carriers in the population. It cannot be excluded that the nucleotide variant constitutes a potentially causative mutation. However, in the reported patient, CBX2 mutations caused autosomal recessive disease, affecting a 46,XY subject. Therefore it is less likely that a heterozygous change inherited from a 46,XY father would be causative alone in the 46,XX patient, in whom we have also confirmed expression of both two alleles.

**Copy number changes**

No deletion or duplication affecting the CBX2 gene was detected in the patient cohort by MLPA. These results are surprising to us as four copy number variants involving CBX2 are registered in the CNV database. Three variants are reported encompassing the entire gene detected at a reported frequency of one loss in 39 controls, one loss in 269 and two gains and one loss in 30. In addition, a change spanning exon 1 to 3 is also described with one loss and 22 gains out of 50 controls by de Smith et al. Due to the reported high frequency of CNVs, we expected to find some patients with a heterozygous deletion, or partial gene duplication. The reported high frequency by de Smith et al in a French population not replicated in our experiments could constitute a
population difference, or it could also represent a false positive array-CGH result in the previous study.

Expression
The CBX2 gene is reported to be transcribed in two isoforms, but no expression pattern has been reported. Therefore, we investigated the expression of the two CBX2 isoforms in two cell lines and have successfully shown that both RNA isoforms are expressed in gonadal fibroblasts as well as in EBV-immortalized lymphocytes. It still remains to be determined if isoform 2 is translated into a protein.

Conclusion
Despite the potential importance of CBX2 in human sex development our study on a group of 46,XX and 46,XY gonadal DSD cases did not reveal any causative point mutations in the CBX2 gene. Neither was any gene deletion or duplication detected by MLPA. Consequently, our study does not support CBX2 gene mutations as a common cause of gonadal DSD.

GENOME WIDE APPROACH TO IDENTIFY NOVEL CANDIDATE REGIONS IN PRIMARY OVARIAN INSUFFICIENCY (PAPER III)

The aim of this study was to identify novel candidate genes for ovarian development and function by investigating submicroscopic genetic imbalances in patients with POI. By a genome wide approach, novel candidate genes could be identified both on the X chromosome and on autosomes. DNA samples from 26 patients with POI were analyzed by array-CGH using the same platform as described for Paper I.

11 unique copy number changes were identified in a total of 13 patients. One aberration affects the GDF9 gene, already associated to POI, the other ten changes unravel novel candidate regions. All aberrations were confirmed by MLPA. When possible, inheritance pattern was investigated. Maternally inherited changes have not automatically been excluded as a variable phenotype expression due to polygenic or environmental factors, or an autosomal recessive inheritance cannot be excluded.
An MLPA probe set for identification of copy number changes in the novel candidate regions identified by array-CGH was developed. This was used for investigation of 95 controls, and an additional 28 patients with POI. The extended cohort and control material were also used in the follow-up studies of candidate genes. Below, the
identified duplication affecting the $GDF9$ gene, and six candidate genes for POI are discussed.

**GDF9 on 5q31.1**

A partial duplication within the $GDF9$ (Growth differentiation Factor 9) gene was identified in a patient who had SA at 15 years of age. Using PCR, we amplified the duplication junction (figure 7) and could by sequencing confirm a tandem, head to tail duplication, of 475bp containing the first part of the 5’ UTR and a short upstream DNA sequence.

**Figure 7. GDF9 duplication.**

A  **GDF9**

![Schematic representation of the GDF9 gene](image1)

B  **Duplication**

![Schematic representation of the duplication](image2)

C  **UCSC representation**

![Representation from the UCSC genome browser](image3)

A. Schematic representation of the $GDF9$ gene. ATG indicates the initiation of translation. Red dots represent $NOBOX$ binding elements (NBE). Purple dot represents the E-box. Yellow arrow indicates the duplicated segment. Blue arrows represent PCR primers for duplication breakpoint amplification. R, reverse primer; F, forward primer. B. Schematic representation of the duplication. Blue arrows represent PCR primers. Green line represents PCR product containing duplication junction. C. Representation from the UCSC genome browser, GRCh37/hg19 assembly. Horizontal yellow arrow line indicates the duplicated segment. Location of NBE (TAATTA, TAATTG) and E-box (CAGCTG) are shown as vertical red and purple lines, respectively. Vertical black boxes represent MLPA probes.
Results and discussion

*GDF9* is an oocyte secreted factor, expressed from the primary follicle stage, and is necessary for normal folliculogenesis and fertility. Mutations of *GDF9* are associated with POI, although at a low frequency [51, 98-100].

Within the duplicated region there are three NOBOX binding elements (NBE), and an E-box sequence. In mouse, NBEs regulate *Gdf9* expression and the E-box sequence is necessary for the ovarian specific expression of *Gdf9*. Both the NBEs and the E-box are conserved between mouse and human, as well as several other species and are likely important for regulation of human *GDF9* expression. The identified duplication is likely to have caused an altered *GDF9* expression in the patient’s ovary, leading to POI. This is the first mutation identified affecting the regulatory region of *GDF9*.

As a further investigation, the entire patient cohort (total 54 patients) was investigated by an in-house designed MLPA probe set, for the identification of deletions/duplications within or encompassing *GDF9*. No additional case was found.

**DNAH6 on 2p11.2**

A 171kb deletion affecting the *DNAH6* (Dynein axonemal heavy chain 6) was identified in one patient with SA at 22 years of age. Another array-CGH study has described a smaller, partially overlapping deletion detected in a patient with SA (figure 8) [101]. An additional study has also identified the related *DNAH5* as a candidate gene for POI [102]. Mutations in *DNAH5* are otherwise known to cause primary ciliary dyskinesia (PCD).

*DNAH6* is a heavy axonemal dynein chain. Dyneins are microtubule-associated motor protein complexes, which are important for ciliary and flagellar motility. Also, dyneins have been suggested to have a role in ovarian development. In adult tissues *DNAH6* can be found in fallopian tubes and testis, among other tissue types. We believe that *DNAH6* is a very interesting candidate gene for POI.

**Figure 8. DNAH6 deletion.**

Representation from the UCSC genome browser, GRCh37/hg19 assembly. The green line represents the deletion detected in Patient 21. The orange line show the deletion reported by Ledig *et al*. Vertical black lines represent MLPA probes.
**Results and discussion**

**TSPYL6 on 2p16.2**

A deletion affecting the *TSPYL6* gene (TSPY-Like 6) was identified in a patient with PA. The *TSPYL6* gene function is unknown, but it has a C-terminal NAP (nucleosome assembly protein) domain, where NAP proteins are known to affect gene transcription by histone moving and nucleosome assembly. Homozygous mutations of *TSPYL1* have been found to cause 46,XY gonadal dysgenesis and sudden infant death syndrome (SIDDT). We consider *TSPYL6* an interesting candidate gene for POI. As a further investigation, we sequenced the *TSPYL6* gene in the entire patient cohort. No inactivating mutations were identified.

**SMARCC1 and CSPG5 on 3p21.31**

A 135kb duplication on chromosome 3p21.31 was found in a patient with PA. The duplication spans the *SMARCC1* gene (SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily c member1) and at least the first two exons of the *CSPG5* gene (chondroitin sulfate proteoglycan 5). SMARCC1 regulates transcription by chromatin remodeling and is highly expressed in both follicle and ovarian stroma cells, among other cell types in various tissues. CSPG5 is a proteoglycan that may function as a differentiation factor. In mice, there is a differential expression profile of *Cspg5* in *Sf1*+ somatic cells in developing gonads at embryonic day 12.5 (E12.5) with a higher expression in male gonads, and with an increasing expression in female gonads on E13.5.

We consider both *SMARCC1* and *CSPG5* possible candidate genes for POI.

**SH3GL3 on 15q25.2**

A duplication of almost the entire sequence of the first intron of the *SH3GL3* gene (SH3-domain GRB2-like 3) was identified in a patient with PA. The *SH3GL3* gene contains a SRC Homology-3 (SH3) domain involved in signal transduction for transcriptional regulation. *SH3GL3* expression is abundant in testis where it is believed important for spermatogenesis. It is moderately expressed in follicle cells.

By array-CGH no positional data can be obtained. An interstitial duplication could affect splicing or have no effect. An association between *SH3GL3* and POI should be further investigated.
Del 19p13.3

A 128kb deletion on chromosome 19 was identified in a patient with SA affecting 6 genes: TJP3, APBA3, MRPL54, RAX2, MATK, and ZFR2. Of these, the ZFR2 gene (zinc finger RNA binding protein 2) is an interesting candidate gene for POI. Its function is not fully understood, but it can bind double stranded RNA or RNA/DNA hybrids. ZFR2 is expressed in many adult cell types, including ovarian follicle and stroma cells. In addition, ZFR2 is expressed in female gametocytes and germ cells during embryonic development. Ovarian biopsies from the patient shows bilateral complete lack of oocytes and follicles.

KRTAP2-3 and KRTAP2-4 on 17q21.2

Three patients with SA were found to carry a 9kb deletion encompassing the KRTAP2-3 (Keratin associated protein 2-3) and KRTAP2-4 (Keratin associated protein 2-4) genes, two heterozygous and one homozygous for the deletion. We hypothesized an autosomal recessive causative mechanism and by sequencing we detected a missense substitution (p.Cys73Tyr) in the remaining allele of the two heterozygous patients. The variant is reported (rs113397060), but with no allele frequency. As the prediction program Polyphen2 predicted the change to be probably damaging we investigated all patients and the control material for this change. One patient was found to be homozygous for the substitution, but also one control. Therefore we do not consider KRTAP2-3 as a candidate gene for POI. It is possible that this deletion of KRTAP2-3 and KRTAP2-4 is a normal variant in the Swedish population, although not reported in DGV, as two controls were also found to be heterozygous for the deletion.

Conclusion

Using array-CGH, we have identified the first mutation affecting the regulatory region of GDF9 in a patient with SA, most likely pathogenic. Importantly, the small, 475bp duplication, would not have been detected without the customized probe enrichment, making clear that platform resolution and probe targeting is crucial. We have also identified a second DNAH6 deletion in a patient with SA, corroborating DNAH6 as a candidate gene for POI. In addition, we have identified several novel candidate genes in POI. Even though descriptive in its form, studies reporting recurrent and related genetic variants in phenotypically well characterized groups of patients can contribute to the identification of novel candidate genes in gonadal development and maintenance and supporting the initiation of functional studies.
**PSMC3IP GENE STUDY (PAPER IV)**

The *PSMC3IP* (*PSMC3 interacting protein*) gene was implicated in POI in 2011. In a consanguineous family with several members with PA, all affected patients shared a terminal homozygous 3bp deletion (c.600_602del, p.Glu201del). The mutant protein showed a significantly decreased function as an estrogen co-activator. Estrogen is important for prenatal development of the follicle pool and follicular development at puberty. A decreased estrogen-dependent transcription could thereby possibly cause POI presenting with PA or SA. No other study has investigated the *PSMC3IP* gene in a group of unrelated patients with POI. It is of interest to investigate the possible frequency of pathogenic mutations in new candidate genes for POI. This is both for diagnostic purposes, and for strengthening the understanding of the gene function in ovarian development and maintenance. Therefore, we investigated a cohort of 50 patients with POI for possible *PSMC3IP* mutations.

**Sequencing results**

Samples from 50 patients were successfully sequenced. No mutations or sequence variations in the coding region of the *PSMC3IP* gene were detected. In intron 4, a novel heterozygous change was identified in one patient with SA at 37 years of age. The intronic transition c.337+33A>G is not previously reported. *In silico* analysis of splicing efficiency using three different prediction tools did not support an effect of the variant on splicing. Although we could not detect this change in any of the 190 control alleles investigated we believe it constitutes a rare normal variant.

**Conclusion**

The *PSMC3IP* gene has so far only been investigated in a single consanguineous family with hereditary POI. The present study is based on a group of 50 patients with POI where no mutation of the *PSMC3IP* gene was detected. Using binomial probability calculation we can note that the probability of detecting mutations in zero out of fifty patients with POI with an assumed POI population prevalence of possible causative *PSMC3IP* mutations of 5.8%, is only 5% (exact confidence interval 0-7.11%). This is less than reported for mutations in the autosomal gene *NOBOX* in Caucasian patients with POI and some reported studies of *FMR1* premutations. We can therefore conclude that *PSMC3IP* gene mutations are not a common cause of POI in this Swedish cohort. A population difference cannot be excluded, our patients are mostly Swedish (Caucasian), whereas the family investigated by Zangen *et al* was of Palestinian
Results and discussion

descent. It remains possible that the described mutation can be found in a higher prevalence in patients of this ethnic background.

For clinical genetic investigation of POI, a single gene approach with consecutive screening by conventional sequencing of candidate genes has so far been quite unsuccessful and expensive, with a low detection rate per gene investigated. With the development of next generation sequencing techniques, a multiple gene approach where a panel of known candidate genes is analyzed simultaneously, could be more successful.

POI GENE SEQUENCING (UNPUBLISHED DATA)

In addition to the clinical genetic investigation of *BMP15*, *FSHR* and *NR5A1*, all included patients with POI, in total 54 patients without a molecular diagnosis, have been analyzed for inactivating mutations in the *GDF9*, *FIGLA* and *NOBOX* genes. No definitive causative mutation was found, although some novel or rare changes were identified. The results are summarized in table 9.

**FIGLA**

Seven reported changes were identified in the cohort (table 9a). Of these, one is a very rare missense variant p.Pro93Thr, detected in 5 out of 4093 European American controls, 0 out of 1842 African American in whole-exome sequencing reports. The change is predicted damaging Polyphen2, Mutation taster, SIFT and Align GCVD. The patient carries this change in heterozygous form. We investigated our control material of 95 healthy women without POI and did not detect any additional carrier. Previously reported *FIGLA* mutations with a possible causative effect have been described in heterozygous form [47]. Unfortunately, no parental samples were available for this patient. If we could show that the change was either paternally inherited or *de novo*, its pathogenicity would be strengthened. It is however an interesting candidate for further evaluation with functional studies.

The other detected changes are either common, likely benign or do not predict any potential change in splicing.
**GDF9**

In *GDF9*, five reported SNPs were identified in the cohort (table 9b), four common changes and one with no reported minor allele frequency (MAF). These are either silent changes, or outside the coding regions. Splice site prediction tools do not predict any change in splicing for any change.

**NOBOX**

Sequencing of the *NOBOX* gene revealed one novel change and twelve reported SNPs (table 9c). The novel change is silent (c.1548C>T) and juxtaposed a common polymorphism. It was not detected in any control allele, but is maternally inherited. The change is not predicted to introduce any splice site and is likely a rare polymorphism.

In one patient the missense change p.Arg117Trp was detected. It has previously been described as potentially damaging [103], but it has a reported MAF of 1.7%, which is too common for a causative change in the rare condition POI. The MAF however is based on a sample of only 38 controls, meaning it was detected in heterozygous form in one control, which does not have to represent the distribution in the entire population. However another study reported the minor allele in 13 of 118 alleles in a control material, which speaks more strongly against a causative role. We did not have access to parental samples, and consider the variant interesting for further evaluation. The other changes are either common, maternally inherited, predicted benign or not influencing splicing.
### Table 9a. *FIGLA* sequencing results (unpublished data).

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Reported MAF</th>
<th>Nucleotide change</th>
<th>Location</th>
<th>Amino acid change</th>
<th>POI (54)</th>
<th>Study pop MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12713717</td>
<td>G=0.398/869</td>
<td>c.231+24T&gt;C</td>
<td>Intron 1</td>
<td>-</td>
<td>T/T:26</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/C:23</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/C:5</td>
<td></td>
</tr>
<tr>
<td>rs200665269</td>
<td>-</td>
<td>c.277C&gt;A</td>
<td>Exon 2</td>
<td><strong>p.Pro93Thr</strong></td>
<td>C/C:53</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/A:1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/A:0</td>
<td></td>
</tr>
<tr>
<td>rs7566476</td>
<td>C=0.249/543</td>
<td>c.422G&gt;C</td>
<td>Exon 3</td>
<td><strong>p.Ser141Thr</strong></td>
<td>G/G:9</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/C:27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/C:17</td>
<td></td>
</tr>
<tr>
<td>rs7566541</td>
<td>G=0.216/472</td>
<td>c.552C&gt;T</td>
<td>Exon 3</td>
<td><strong>p.His184=</strong></td>
<td>C/C:13</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/T:24</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/T:20</td>
<td></td>
</tr>
<tr>
<td>rs144711401</td>
<td>~0.016/35</td>
<td>c.644+109_644+112delGGAA</td>
<td>Intron 4</td>
<td>-</td>
<td>GGAA/GGAA:51</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GGAA/-:3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-/-:0</td>
<td>0.03</td>
</tr>
<tr>
<td>rs56135050</td>
<td>T=0.317*691</td>
<td>c.*5T&gt;A</td>
<td>3'UTR</td>
<td>-</td>
<td>T/T:23</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/A:28</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/A:3</td>
<td></td>
</tr>
<tr>
<td>rs56316086</td>
<td>G=0.314/686</td>
<td>c.7* A&gt;G</td>
<td>3'UTR</td>
<td>-</td>
<td>A/A:23</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/G:28</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/G:3</td>
<td></td>
</tr>
</tbody>
</table>

NM_001004311.3 is used as reference sequence.

SNP ID, single nucleotide identification number; MAF, minor allele frequency; POI, primary ovarian insufficiency; pop, population; UTR, untranslated region.
### Table 9b. GDF9 sequencing results (unpublished data)

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Reported MAF</th>
<th>Nucleotide change</th>
<th>Location</th>
<th>Amino acid change</th>
<th>POI (54)</th>
<th>Study pop MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs17166282</td>
<td>C=0.044/95</td>
<td>NM_005260.3:c.-32A&gt;G</td>
<td>5'UTR</td>
<td>-</td>
<td>A/A:53</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/G:0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/G:1</td>
<td></td>
</tr>
<tr>
<td>rs254285</td>
<td>C=0.185/402</td>
<td>NM_005260.3:c.398-39G&gt;C</td>
<td>Intron 1</td>
<td>-</td>
<td>G/G:1</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/C:6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/C:47</td>
<td></td>
</tr>
<tr>
<td>rs254286</td>
<td>A=0.454/993</td>
<td>c.447C&gt;T</td>
<td>Exon 2</td>
<td>p.Thr149=</td>
<td>C/C:13</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/T:18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/T:23</td>
<td></td>
</tr>
<tr>
<td>rs10491279</td>
<td>T=0.162/353</td>
<td>c.546G&gt;A</td>
<td>Exon 2</td>
<td>p.Glu182=</td>
<td>G/G:43</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/A:10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/A:1</td>
<td></td>
</tr>
<tr>
<td>rs376173165</td>
<td>-</td>
<td>c.1182T&gt;C</td>
<td>Exon 2</td>
<td>p.Tyr394=</td>
<td>T/T:53</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/C:1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/C:0</td>
<td></td>
</tr>
</tbody>
</table>

SNP ID, single nucleotide identification number; MAF, minor allele frequency; POI, primary ovarian insufficiency; pop, population.
Table 9c. *NOBOX* sequencing results (unpublished data).

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>MAF reported</th>
<th>Nucleotide change</th>
<th>Location</th>
<th>Amino acid change</th>
<th>POI (54)</th>
<th>MAF Study pop</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1208179</td>
<td>G=0.079/171</td>
<td>c.42T&gt;C</td>
<td>Exon 1</td>
<td>p.Gly14=</td>
<td>T/T:33</td>
<td>0.20</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/C:18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/C:2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs727714</td>
<td>G=0.336/732</td>
<td>c.262C&gt;T</td>
<td>Exon 3</td>
<td>p.Leu88=</td>
<td>C/C:9</td>
<td>0.56</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/T:27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/T:17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7800847</td>
<td>T=0.017/38</td>
<td>c.349C&gt;T</td>
<td>Exon 4</td>
<td><strong>p.Arg117Trp</strong></td>
<td>C/C:53</td>
<td>0.01</td>
<td>Potentially causative.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/T:1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/T:0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs201806397</td>
<td>-</td>
<td>c.454G&gt;A</td>
<td>Exon 4</td>
<td>p.Gly152Arg</td>
<td>G/G:51</td>
<td>0.03</td>
<td>Predicted benign.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/A:3</td>
<td></td>
<td>Maternally inherited (2).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/A:0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs373867875</td>
<td>-</td>
<td>c.715C&gt;T</td>
<td>Exon 4</td>
<td><strong>p.Arg239Trp</strong></td>
<td>C/C:53</td>
<td>0.01</td>
<td>Predicted benign.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C/T:1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/T:0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs370846754</td>
<td>-</td>
<td>c.795G&gt;A</td>
<td>Exon 4</td>
<td>p.265Pro=</td>
<td>G/G:53</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/A:1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A/A:0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP ID</td>
<td>A/MAF</td>
<td>Change</td>
<td>Location</td>
<td>Type</td>
<td>Genotype Frequencies</td>
<td>p-Value</td>
<td>Annotation</td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td>----------------</td>
<td>----------</td>
<td>------</td>
<td>----------------------</td>
<td>---------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>rs757388</td>
<td>0.337/733</td>
<td>c.1154+1T&gt;C</td>
<td>Intron 6</td>
<td>-</td>
<td>T/T:32</td>
<td>0.40</td>
<td>-</td>
</tr>
<tr>
<td>rs11769847</td>
<td>0.336/732</td>
<td>c.1155-23G&gt;A</td>
<td>Intron 6</td>
<td>-</td>
<td>G/G:11</td>
<td>0.56</td>
<td>-</td>
</tr>
<tr>
<td>rs2525702</td>
<td>0.118/257</td>
<td>c.1444G&gt;A</td>
<td>Exon 8</td>
<td>p.Gly482Ser</td>
<td>G/G:34</td>
<td>0.20</td>
<td>Predicted benign.</td>
</tr>
<tr>
<td>Novel change</td>
<td>-</td>
<td>c.1548C&gt;T</td>
<td>Exon 9</td>
<td>p.Pro516=</td>
<td>C/C:53</td>
<td>0.01</td>
<td>Maternally inherited (1). Not found in 190 control alleles.</td>
</tr>
<tr>
<td>rs2699503</td>
<td>0.350/762</td>
<td>c.1549T&gt;C</td>
<td>Exon 9</td>
<td>p.Phe517Leu</td>
<td>T/T:20</td>
<td>0.56</td>
<td>Common, likely tolerated.</td>
</tr>
<tr>
<td>rs1208216</td>
<td>0.082/179</td>
<td>c.1796C&gt;A</td>
<td>Exon 10</td>
<td>p.Pro599His</td>
<td>C/C:34</td>
<td>0.20</td>
<td>Predicted benign.</td>
</tr>
<tr>
<td>rs77802098</td>
<td>0.024/52</td>
<td>c.1991A&gt;G</td>
<td>Exon 10</td>
<td>p.Lys664Arg</td>
<td>A/A:52</td>
<td>0.20</td>
<td>Predicted benign. Paternally inherited (1).</td>
</tr>
</tbody>
</table>

NM_001080413.3 is used as reference sequence. SNP ID, single nucleotide identification number; MAF, minor allele frequency; POI, primary ovarian insufficiency; pop, population; UTR, untranslated region.
CONCLUDING REMARKS - FUTURE PERSPECTIVE

Gonadal disorders of sex development (DSD) are rare conditions affecting patients with 46,XY or 46,XX karyotype. The gonadal phenotype can vary, and depending on the hormonal profile, the phenotype of internal and external genitalia can differ. Primary ovarian insufficiency (POI) can present as a continuous spectrum of ovarian dysfunction, from primary amenorrhea (PA) to secondary amenorrhea (SA). The overall aim of this thesis was to identify and study candidate genes in gonadal DSD and POI. A molecular diagnosis in these conditions can offer the patient a causative explanation for their condition and provide for a more specific genetic counseling of family members. In addition, by identifying novel genes we can also achieve a better understanding of the underlying molecular mechanisms of gonadal development and maintenance.

By genome wide array-CGH analysis of patients with 46,XY gonadal dysgenesis we have identified new candidate genes for testicular development. Copy number changes affecting the SUPT3H and C2ORF80 genes were identified and characterized in two affected sisters with 46,XY GD. We also implicate the PIP5K1B and PRKACG genes as potential candidate genes.

By a candidate gene approach, we have investigated the CBX2 gene for sequence alterations and copy number changes in a group of patients with 46,XY or 46,XY gonadal DSD, or POI. Our negative findings do not support CBX2 gene mutations as a common cause for these conditions.

Using array-CGH we have identified the first mutation affecting the regulatory region of GDF9 in a patient with SA, most likely pathogenic. We have also identified a second patient with SA and DNAH6 deletion, corroborating the role for DNAH6 in POI and making it worth proceeding with functional characterization. In addition, we have also identified TSPYL6, SMARCC1, CSPG5, SH3GL3 and ZFR2 as candidate genes in POI.

We have investigated the PSMC3IP, FIGLA, NOBOX and GDF9 genes for sequence alterations in a cohort of patients with POI. No causative mutations were found, but two changes to follow up were identified. In addition, we have investigated the cohort for copy number changes affecting GDF9 where no additional case to the patient identified by array-CGH was found.

MLPA has been used to develop diagnostic tools for the identification of copy number changes of GDF9 and CBX2.
The genetic investigation of patients with gonadal DSD should always be centralized to a specialized DSD unit. Also, the investigation and treatment of patients with POI is suitable for centralization to subspecialists within reproductive endocrinology. In addition to the clinical investigation and substitution therapy requiring special interest from the physician, results from the clinical genetic investigation can have impact on the entire family. Genetic counseling must be offered these patients and the family members. A suggested clinical investigation procedure for POI is given in Appendix 1.

Based on the current knowledge, I can recommend the following genetic investigations today:

- For patients with isolated 46,XY GD: investigation of sequence alterations of SRY, NR5A1, WT-1, SOX9, DHH, MAP3K1, TSPYL1 and GATA4 and copy number change analysis of NR0B1, SOX9, SRY, WNT4, NR5A1, DMRT1. Only one patient has been described with mutations in CBX2 and our study has not identified any other mutation. Before any new supporting data is provided CBX2 analysis cannot be recommended as a clinical investigation.

- For patients with POI: investigation of sequence alterations of BMP15, FIGLA, FSHR, GDF9, NOBOX, and NR5A1. Also FMR1 analysis is recommended in patients with SA. FOXL2 mutations have so far only been described in patients with BPES, but it is possible that mutations of this gene could present with isolated POI, as in acampomelic SOX9 mutations. Therefore it should be considered for diagnostic investigation. Mutations in PSMC3IP have so far only been described in one family and our study has not revealed any additional case. Until another report supports its importance it cannot be recommended as a clinical analysis. Copy number change analysis of the aforementioned genes must also be considered in patients with POI. Investigation of promotor regions in sequence and copy number change analyses should be considered.

The knowledge of implicated genes in gonadal development is rapidly increasing and it is important that the genetic investigations conducted are constantly evolving, and also retrospectively offered patients diagnosed earlier.

For genetic investigation of patients with gonadal DSD or POI, the simultaneous detection of gene dosage aberrations and small DNA sequence aberrations would be preferred, and ideally conducted in a validated, fast and cost-efficient manner. In the future, once next generation sequencing will provide reliable data on copy numbers, whole-genome sequencing could be considered for diagnostic purposes.
Concluding remarks – future perspective

It would be interesting to apply next generation sequencing to identify novel candidate genes in gonadal DSD and POI by studying families with several affected members. We have now samples from another pair of sisters with 46,XY GD where we would like to collect DNA samples from parents and the unaffected sibling. In addition, we are trying to include family members of a patient with POI with PA, where several sisters are also affected with POI with a variable phenotype. Both these families would be very interesting to study further as the familial cluster strongly suggests a common genetic mechanism.

However, the number of sequence alterations identified in each investigated person by next generation sequencing is extensive and both laborious and difficult to interpret. Using a family based approach, we could exclude changes that do not segregate with phenotype, and also perform analysis depending on an assumed dominant or recessive inheritance pattern. This would greatly decrease the amount of data noise in search of the true signal.

Novel candidate genes identified by such an approach, as well as the rare missense variants identified in the FIGLA and NOBOX genes, would be interesting to investigate by functional studies. Also DNAH6 is a candidate gene for functional investigation. Animal models could be considered, but in vitro techniques and cell-based systems could be designed to evaluate the expression and function of mutant protein compared to its wild type counterparts.

For POI, a multicentre approach with a large numbers of patients and controls could form the basis of association studies evaluating SNP frequency in patients compared to controls, in attempts to localize genes and genomic regions important for gonadal function.
"Hur många barn får jag när jag blir stor?"


Det finns många bakomliggande orsaker till infertilitet, både manliga och kvinnliga faktorer. Denna avhandling rör de ovanliga tillstånd där äggstockar eller testiklar, gemensamt kallade gonader, inte har utvecklats eller har förlorat sin funktion i förtid. Dessa patienter är normala flickor, förutom att de antingen inte kommer i puberteten eller förlorar sin mens i mycket unga år. Den drabbade kan ha en manlig kromosomuppsättning, 46,XY, eller kvinnlig, 46,XX, och motsvarande kliniska diagnos kallas 46,XY gonad dysgenesi (GD) eller primär ovariell insufficiens (POI). POI kan orsakas av externa faktorer såsom kirurgi, strålning, cytostatikabehandling, och även av autoimmunitet, men genetiska faktorer spelar en stor roll. Vi vet också att styrningen under fosterlivet från ett neutralt gonadanlag till utvecklingen av testikel eller äggstockar regleras av genetiska mekanismer.

Genom att studera patienter med 46,XY GD och POI hoppas vi kunna få en bättre förståelse för de underliggande mekanismer som styr den normala gonadutvecklingen. Därmed hoppas vi kunna ge fler patienter en förklaring till varför just de drabbats, vilket kan vara en liten lättnad i ett tillstånd som det annars kan vara svårt att förlika sig med. En diagnos medför även att den genetiska rådgivning vi kan erbjuda patienten och familjen blir säkrare och bättre.

Denna avhandling består av fyra delprojekt:

- Identifiering av nya kandidatgener för testikelutvecklingen. Vi har använt en specialdesignad högupplöst platform för detektion av mycket små förluster eller dubblering av avrsmassa. Metoden kallas array-CGH. Med denna teknik har vi identifierat fyra gener som vi tror kan ha betydelse för den normala utvecklingen av testikeln. Generna heter SUPT3H, C2ORF80, PIP5K1B och PRKACG. Alla uttrycks normalt i testikeln, och förlust av Pipk51b ger sterilitet hos manliga möss.
vilket ytterligare stöder vår hypotes. Dessa gener behöver studeras vidare för att bättre kunna förstå deras funktion. (Artikel I)

- Studier av kandidatgenen CBX2. Genen Cbx2 har beskrivits påverka gonaduvecklingen hos möss, och en patient har beskrivits med mutationer i denna gen och normal äggstocksuveckling trots en manlig karyotyp. Vi undersökte huruvida mutationer i denna gen förekom hos våra patienter med 46,XY GD och POI. Inga mutationer påträffades i vår undersökta grupp och vi kan därmed sluta oss till att CBX2 mutationer inte är en vanlig orsak till dessa tillstånd. (Artikel II)

- Array-CGH studier av patienter med POI. Vi har identifierat den första mutationen som påverkar den reglerande regionen för genen GDF9 som är essentiell för äggstocksuvecklingen. Dessutom har vi påvisat en andra förlust av genen DNAH6 hos en patient med POI, vilket stärker dennes troliga roll i normal äggstocksuveckling och funktion. Vi beskriver även nya kandidatgener för POI, TSPYL6, SMARCC1, CSPG5, SH3GL3 och ZFR2. (Artikel III)

- Studier av kandidatgenerna PSMC3IP, FIGLA och NOBOX. En mutation i genen PSMC3IP har beskrivits orsaka POI i en familj med flera drabbade. I vårt material kunde vi inte påvisa någon mutation, och därmed vet vi att mutationer i PSMC3IP inte är en vanlig orsak till POI i Sverige. Vi har hittat två möjligt sjukdomsorskande förändringar i generna FIGLA respektive NOBOX som behöver studeras vidare. (Artikel IV och opublicerade data)

Sammanfattningsvis kan vi se att 46,XYGD och POI är genetiskt heterogena tillstånd. Kunskapen kring de mekanismer som styr utvecklingen av och funktionen i testikel och äggstock ökar snabbt. Den genetiska utredning som görs bör följa med i utvecklingen och innefatta analys av både punktmutationer av arvsmassan såväl som analys av förlust eller dubblering av genetiskt material.

De drabbade patienterna bör skötas av specialister, och för POI innebär detta av gynekolog med subspecialisering inom reproduktiv endokrinologi. Centralisering är viktigt för utredning, rådgivning och behandling då patienter utan egen könshormonsproduktion måste få individuellt anpassad ersättning och uppföljning.

Patienter med en manlig kromosomuppsättning löper risk att drabbas av tumörer i den kvarvarande testikelvävnaden, vilket kräver specialistbdömning och handläggning.

Oavsett kromosomuppsättning kan alla patienter med 46,XYGD och POI bli mammor med hjälp av äggdonation, och flera av våra patienter har redan fått barn på detta sätt.
APPENDIX 1. SUGGESTED CLINICAL INVESTIGATION OF WOMEN WITH POI

CLINICAL PRESENTATION:

Primary amenorrhea with absent or delayed pubertal development. Infantile external female genitalia. (Tanner stage can vary depending on sex steroid production from adrenal glands)

OR

Secondary amenorrhea before 40 years of age in a previously menstruating woman for at least 4 months accompanied by hypergonadotropic hypogonadism

BASIC INVESTIGATION

Clinical / gynecological examination.

Ultrasonography or magnetic resonance imaging for investigation of presence of internal female genitalia.

Hormonal profile: FSH, LH, estradiol, testosterone, TSH, T4, prolactin

Genetic investigation: Karyotype

Gonadal biopsies should be taken if laparoscopy is performed, and correct sample handling must be ensured before surgery.

Patients with a 46,XY karyotype should be referred to a DSD specialist.

Patients with a normal 46,XX karyotype should be referred to a gynecological specialist in reproductive endocrinology.

Patients with 46,XY or 46,XX chromosomal DSD should be referred for clinical genetic counseling. The clinical care is recommended to be handled by a DSD specialist.
SPECIALIST INVESTIGATION

The following investigations are recommended to be handled by, or performed in collaboration with, a subspecialist within reproductive endocrinology.

Immunological profile: autoantibodies against cytochrome P450scc, 17α hydroxylase, 21 hydroxylase and TPO. In the case of positive findings, an extended analysis of autoantibodies against transglutaminase, parietal cells, GAD, IA2 and TRAK, is suggested.

Genetic investigation: FMR1 premutation analysis in secondary amenorrhea. Gene sequencing and copy number analysis in collaboration with clinical genetic laboratory.

GONADAL BIOPSY INVESTIGATION

Biopsy material from gonads from patients with POI should be investigated by a pathologist for histological examination. In addition, a biopsy should be sent for FISH (fluorescent in situ hybridization) analysis at a clinical genetic laboratory to exclude gonadal chromosomal mosaicism. It is also suggested that fibroblasts be cultured and saved from this material for possible future analysis.

A flowchart for the referral of patients is given below.

```
POI

Karyotype 46,XY 46,XX

<table>
<thead>
<tr>
<th>46,XY</th>
<th>DSD specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td>46,XX</td>
<td>Specialist in reproductive endocrinology</td>
</tr>
<tr>
<td></td>
<td>Clinical geneticist</td>
</tr>
</tbody>
</table>

Chromosomal DSD
```
## APPENDIX II. *IN SILICO* ANALYSIS, PROGRAMS, RESOURCES AND TOOLS.

<table>
<thead>
<tr>
<th>Name</th>
<th>URL / Manufacturer</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCSC</td>
<td><a href="http://genome.ucsc.edu/">http://genome.ucsc.edu/</a></td>
<td>Genome browser. Use to retrieve sequences and gene information. BLAT tools used for primer and probe specificity control.</td>
</tr>
<tr>
<td>Ensembl</td>
<td><a href="http://www.ensembl.org/">http://www.ensembl.org/</a></td>
<td>Genome browser.</td>
</tr>
<tr>
<td>Primer3Input</td>
<td><a href="http://primer3.ut.ee/">http://primer3.ut.ee/</a></td>
<td>Primer design.</td>
</tr>
<tr>
<td>RaW-Oligov0.15β</td>
<td><a href="http://www.mlpa.com/WebForms/WebFormMain.aspx?Tag=zejCZBtdOUyAt3KF3EwRZhAPz9QEm7akikAm7AOEGw1vtZvffaZPOiSig8uqel7Yd">http://www.mlpa.com/WebForms/WebFormMain.aspx?Tag=zejCZBtdOUyAt3KF3EwRZhAPz9QEm7akikAm7AOEGw1vtZvffaZPOiSig8uqel7Yd</a></td>
<td>MLPA probe design, available for download at the MRC Holland website.</td>
</tr>
<tr>
<td>SeqScapev2.5</td>
<td>Applied Biosystems</td>
<td>Sequence analysis.</td>
</tr>
<tr>
<td>Gene Mapper</td>
<td>Applied Biosystems</td>
<td>MLPA analysis.</td>
</tr>
<tr>
<td>GeneMarkerv1.90</td>
<td>Soft genetics</td>
<td>MLPA analysis.</td>
</tr>
<tr>
<td>Cytosure Interpret Software v3.4.3</td>
<td>Oxford Gene Technology</td>
<td>Array-CGH analysis.</td>
</tr>
<tr>
<td>DGV</td>
<td><a href="http://dgv.tcag.ca/dgv/app/home">http://dgv.tcag.ca/dgv/app/home</a></td>
<td>CNV data.</td>
</tr>
<tr>
<td>Tool</td>
<td>Website</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DECIPHER</td>
<td><a href="http://decipher.sanger.ac.uk/">http://decipher.sanger.ac.uk/</a></td>
<td>CNV and phenotype data.</td>
</tr>
<tr>
<td>The Human Protein Atlas</td>
<td><a href="http://www.proteinatlas.org/">http://www.proteinatlas.org/</a></td>
<td>Human protein expression data.</td>
</tr>
<tr>
<td>during sex determination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LifeMap</td>
<td><a href="http://discovery.lifemapsc.com">http://discovery.lifemapsc.com</a></td>
<td>Expression data.</td>
</tr>
<tr>
<td>AlaMutv2.3</td>
<td>Interactive Biosoftware</td>
<td>Mutation analysis by combining Align GVGC, SIFT, MutationTaster and PolyPhen2. Splicing analysis by combining SpliceSiteFinder-like, MaxEntScan, NNSplice, GeneSplicer and Human Splicing finder.</td>
</tr>
<tr>
<td>Polyphen2</td>
<td><a href="http://genetics.bwh.harvard.edu/pph2/">http://genetics.bwh.harvard.edu/pph2/</a></td>
<td>Missense mutation analysis (protein level).</td>
</tr>
<tr>
<td>SIFT</td>
<td><a href="http://sift.bii.a-star.edu.sg/">http://sift.bii.a-star.edu.sg/</a></td>
<td>Missense mutation analysis (protein level).</td>
</tr>
<tr>
<td>NetGene2v. 2.4</td>
<td><a href="http://www.cbs.dtu.dk/services/NetGene2/">http://www.cbs.dtu.dk/services/NetGene2/</a></td>
<td>Splice site analysis.</td>
</tr>
<tr>
<td>NNSPLICEv0.9</td>
<td><a href="http://www.fruitfly.org/seq_tools/splice.html">http://www.fruitfly.org/seq_tools/splice.html</a></td>
<td>Splice site analysis.</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who have helped and encouraged me during my work with this thesis. In particular I would like to thank:

My first main supervisor Anna Wedell, you are a true inspiration. Always ready with scientific input and encouragement, and your passion for science is really contagious. I am very grateful for your support in this project.

My main supervisor Michela Barbaro, my Italian supervisor. With passion and fire you have taught me how to do things properly, and not just the ‘doctor’ way. You have made me realize what I am capable of, and to trust my judgment. And you have with quite some work on your part, turned me in to a scientist. I deeply thank you for this, and all the patience I now understand you must have had.

My co-supervisor Erik Iwarsson, thank you for your calm and encouragement, always ready to discuss clinical questions and making me understand genetics better. And also and for making sure that me, my time and my life doesn’t get stuck in the middle during this project.

My co-supervisor Angelica Hirschberg, I am so happy you’ve been a part of this project, without you so much would not have been possible. Thank you for your help, encouragement, and future belief in me.

I would also like to thank my absolute first supervisor Anna Karlsson at IMPI, who gave me that first PhD project at a high school graduation lunch so long ago. I apologize for never finishing, but I’m very glad for time we had in the lab, and thank you for your support to work at Institut Pasteur. And I’m very happy that you, and Magnus Johansson, taught me so much, which made this project possible.

Co-author Bengt Persson, thank you for your help and input, and for your invaluable knowledge regarding gene and protein conservation and evolution.

Co-author Kenny Rodriguez-Wallberg, thank you for your help and input, and for the clinical treatment perspective.

Collaborators and contributing clinicians Anne-Grethe Myhre, Annika Reims, Johan Svensson, Marie Hofling, Eva Horezemuzova, Lena Arvidsson, Elisabeth Sinnerstad, and Sten Ivarsson, thank you for your work and help, without your work this study would not have been possible.
Acknowledgements

My mentor, Amelia Adamo, thank you for encouraging discussions about intrepreneurship and life.

Present and former members of the AW group at CMM, Svetlana Lajic, always ready with information about diamond quality in between scientific discussions. Tatja Hirvikoski, you are one of the most level-headed people I’ve met, and thank you for all the time we shared writing and talking. Leif Karlsson, your work has been invaluable for this project, I wish you all the best in your scientific future.

Magnus Nordenskjöld, for letting me take advantage of all the knowledge and technical apparatus at the Department of Clinical Genetics for my work, and also for the opportunity to spend part of my residency at the clinic. Thank you Margareta Lagerberg for teaching me array-CGH during all those months, and Isabel Neira Sandoval for making me feel like a star, I love it and thank you for all your help. Ulla Grandell and Monika Carlsson, thank you for all the help with clinical analyses.

Agne Lieden, Kristina Lagerstedt and Helena Malmgren, thank you for always being so happy and helpful, despite my constant interruptions. Your input has been invaluable. Thank you so much, all you wonderful ladies at the department, letting me run my MLPA samples last minute, and scan arrays during your booked time, and helping with EBV transformations on Friday afternoons.

Thank you Lotta Blomberg, Berit Lagerstam and Siv Rödin Andersson, for the invaluable help with control samples collection, and with all the practicalities at Kvinnohälsan.

Thank you Christina Nyström for your help and company in the lab, it’s been great fun, and I think I owe you at least 10 kitchen duty weeks. Thank you Anna-Lena Kastman, Nina Jäntti, Fariba Foroogh, Oliviera Werngren and Karin Husman, for your help with my lab problems, no matter how strange, and for keeping everything and everyone in order.

Ann Britt Wikström, thank you for helping out with “dagens Ameli” so many many days. Thank you Daniel Uvehag, Rudolf Matousek and Dagmar Vejsicka for always fixing my computer problems, ASAP.

To fellow CMM-ites, past and present, Liv Eidsmo, Alexandra Löfstedt, Vasilios Zachariadis, Ellika Sahlin, Bianca Tesi, Johanna Winberg, Ingegerd Överholm, Josephine Wincent, Marie Meeths, Ekatarina Chernogubova, Tobias Laurell, Anna Lindstrand, Emma Tham, Peter Gustavsson, Wolfgang Hofmeister, Aram Rasul, Lousie Lönnndahl, Clara Chamarro, Anna Svenningsson, Anna Skarin, Jia
Cao, Giedre Grigelioniene, Ellen Markljung, Susanna Von Holst and Izabella Körberg, thank you for all nice chats in the lab, corridor and at lunch. It would have been a lonely project without you guys. Extra thanks to all fellow journal club members for addressing studies on facebook use as seriously as next generation sequencing.

Jag vill också tacka följande personer, viktiga i livet i och utanför forskningen:

Min första ST chef Lisskulla Sylvén som anställde mig 2006 trots medföljande spädbarn på intervjun, och stort tack till min nuvarande chef Sophie Graner som stått upp för mig och min forskning i turbulenta tider.

Tack till min kliniska ST handledare Lena Marions som klarar av att handleda en extremt bångstyrig ST läkare, och tack till min första handledare Eva Eneroth, ett kliniskt föredöme som gjorde stort intryck på en nyanställd underläkare.

Stort tack till mina härliga kollegor på Kvinnekliniken, jag trivs så bra bland alla er kloka, bestämda och orginella människor. Tack också till alla barnmorskor och sköterskor som gjort både dagarna på kliniken och nattjourerna roliga i alla möjliga och omöjliga situationer.

Ett stort tack till kollegorna på Klinisk Genetik som gör kliniken så trevlig, och där jag alltid kommer sakna att dela skrivbord med Eva Rudd.

Tack till Erik Sundström för den ovärderliga forskarskolan i molekylär medicin, och till vänner och forskarskolekollegor Sebastian Gidlöf, Tina Villard, Caroline Staff, Mia Karlberg, Malin Kvarnung, Carl-Henrik Shah, Kristina Stenstedt, Louise Ekholm och Frida Lédel. Ett extra stort tack till vennen Josefín Lysell, som med skärpa och engagemang kan hjälpa till att lösa de mest komplicerade av frågor.

Tack till Katarina Remeaus för både samarbete och vänskap, din åsikt och dina tankar är alltid värda att lyssna till.

En stor kram till mina VOV-ladies, Anna Sandström, Marija Simic, Nathalie Roos och Susanne Sjöström, ni är mina förebilder, och jag är redo att resa vart som helst, bara jag får hänga med er.

Stort tack till mina goda vänner, gamla och nya, utspridda och i Enskede, som håller humöret uppe på mig i medgång och motgång. Extra tack till Katarina Hjerotonsson, alltid redo för ett glas vin och trädgårdshäng en tråkig vardagskväll i förorten.

Tack till min familj; mamma, pappa, Kristofer och Jacob, för all hjälp och stöd i livet, stor kram!
Till mina älskade barn, Henry och Svante, ni är mitt hjärta och jag är så tacksam över att ni finns.

Detta hade inte gått utan min Magnus, min GW, min Roy, min Samarkanda. Det är inte alltid lätt, men du gör både resan och målet värda besväret. Tack.

This study was supported by grants from Stiftelsen Frimurare Barnhuset and Karolinska Institutet doctoral funds.
REFERENCES


References


References


References

Wijmenga, B.C. Fauser, P.O.F.C. Dutch, Genome-wide association study in premature ovarian failure patients suggests ADAMTS19 as a possible candidate gene, Hum Reprod, 24 (2009) 2372-2378.


LITERATURE RECOMMENDATIONS

During the work on PhD project I have enjoyed several completely unrelated books, of which I would very much like to recommend the following:

Flowers for Algernon
Gone Girl; A novel
How to be a woman
Lean In: Women, Work, and the will to lead
Let’s discuss diabetes with owls
Middlesex
Special topics in calamity physics
The immortal life of Henrietta Lacks
The fault in our stars
The signal and the noise: why so many predictions fail—but some don’t”
The other hand
Thinking, fast and slow
The Help

In Swedish the following books are a must:

Varat och varan: prostitution, surrogatmödraskap och den delade människan
Flickan och skulden
Bitterfittan
Under det rosa täcket
Det kallas kärlek

Daniel Keyes
Gillian Flynn
Caitlin Moran
Sheryl Sandberg
David Sedaris
Jeffery Eugenides
Marisha Pessl
Rebecca Skloot
John Greene
Nate silver
Chris Cleave
Daniel Kahneman
Kathryn Stockett
Kajsa Ekis Ekman
Katarina Wennstam
Maria Sveland
Nina Björk
Carin Holmberg