eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Department of Genetics, Physical Anthropology and Animal Physiology

Doctoral thesis

Clinical and molecular characterization of DSD patients: Impact of Next Generation Sequencing in diagnosis

Idoia Martinez de LaPiscina Martin
2020

ABBREVIATIONS

17OHP	17-hydroxypreogesterone
17β-HSD3	17β-hydroxysteroid dehydrogenase
21OH	21-hydroxylase
3βHSD1	3b-hydroxysteroid dehydrogenase type 1
3βHSD2	3b-hydroxysteroid dehydrogenase type 2
A	Adenine
ABS	Antley–Bixler Syndrome
aCGH	Array-based Comparative Genomic Hybridization
ACMG	American College of Medical Genetics
ACTH	Adrenocorticotropic hormone
AHC	Adrenal Hypoplasia Congenita
AIS	Androgen Insensitivity Syndrome
AMH	Anti-Müllerian Hormone
AMHR2	Anti-Müllerian Hormone Receptor Type 2
AR	Androgen Receptor
aro	Aromatase
AS	Aldosterone synthase
ATPase	Adenosine triphosphatase
ATRX	a-Thalassemia/Mental Retardation X-linked syndrome
AZF	Azoospermia Factor
BLAST	Basic Local Alignment Search Tool
BMI	Body mass index
BMP15	Bone Morphogenetic Protein 15
BMP2	Bone Morphogenetic Protein 2
bp	Base pair
BPES	Blepharophimosis Ptosis Epicanthus inversus Syndrome
BWA-MEM	Burrows-Wheeler Aligner Memory
C	Cytosine
C10orf2	Chromosome 10 open reading frame 2
CAH	Congenital Adrenal Hyperplasia
CAIS	Complete Androgen Insensitivity Syndrome
cAMP	Cyclic adenosine monophosphate
CBX2	Chromobox 2
CDH1	Cadherin 1
CDS	Coding sequence
CE	Coelomic epithelium
CHD	Congenital heart defects
CIS	Carcinoma *in situ*
CLPP	Caseinolytic Mitochondrial Matrix Peptidase Proteolytic Subunit
CMV	Cytomegalovirus
CNV	Copy number variations
CRE	cAMP response element
CREB	cAMP response element binding protein
CRR	Cysteine-rich region
CT	Computed tomography

CTNNB1	Catenin B1
CYB5A	Cytochrome B5 Type A
CYP	Cytochrome P450
CYP11A1	Cytochrome P450 Family 11 Subfamily A Member 1
CYP11A1-FDX-FDXR	P450 side-chain cleavage/Adrenodoxin/Adrenodoxin reductase system
CYP11B1	Cytochrome P450 Family 11 Subfamily B Member 1
CYP17A1	Cytochrome P450 Family 17 Subfamily A member 1
CYP19A1	Cytochrome P450 Family 19 Subfamily A Member 1
CYP21A1P	Cytochrome P450 Family 21 Subfamily A Member pseudogene
CYP21A2	Cytochrome P450 Family 21 Subfamily A Member 2
DAX1	DSS-AHC critical region on the X chromosome protein 1
DBD	DNA-binding domain
DBP	D-bifunctional protein
dbVar	Database of human genomic structural variation
ddNTP	Dideoxynucleotides triphosphates
DDS	Denys-Drash Syndrome
DDT	Dichlorodiphenyltrichloroethane
DECIPHER	Database of genomic variation and Phenotype in Humans using Ensembl Resources
DGV	Database of Genomic Variants
DHEA	Dehydroepiandrosterone
DHEA-S	Dehydroepiandrosterone-sulfate
DHH	Desert Hedgehog
DHT	Dihydrotestosterone
DM	DNA-binding motif
DMC1	DNA Meiotic Recombinase 1
DMEM	Dulbecco's modified Eagle medium
DMRT1	Doublesex and Mab-3 Related Transcription Factor 1
DMRT2	Doublesex and Mab-3 Related Transcription Factor 2
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphates
DOC	11-deoxycorticosterone
Dpc	Days post-conception
DSD	Disorders of Sex Development
DSS-AHC	Dosage-sensitive sex reversal, adrenal hypoplasia congenital
EDC	Endocrine-disrupting chemical
EDTA	Ethylenediaminetetraacetic acid
EmPCR	Emulsion PCR
EMX2	Empty Spiracles Homeobox 2
ES	Enrichment system
ESHG	European Society of Human Genetics
ESP	Exome Sequencing Project
ESR1	Oestrogen Receptor 1
ESR2	Oestrogen Receptor 2
ExAC	Exome Aggregation Consortium
FGF9	Fibroblast Growth Factor 9
FGFR2	Fibroblast Growth Factor Receptor 2

FGFR3	Fibroblast Growth Factor Receptor 3
FIGLA	Folliculogenesis specific bHLH ranscription factor
FISH	Fluorescence *In Situ* Hybridization
FOG2	Friend Of GATA Protein 2
FOXL2	Forkhead Box L2
FOXO3	Forkhead Box O3
FS	Frasier syndrome
FSH	Follicle Stimulating Hormone
FSHR	Follicle Stimulating Hormone Receptor
Fst	Follistatin
FZD1	Frizzled Class Receptor 1
G	Guanine
G bp	Giga bp
GADD45G	Growth Arrest and DNA Damage-Inducible protein
GATA4	GATA Binding Protein 4
GATK	Genome Analysis ToolKit
GB	Gonadoblastoma
GDF9	Growth Differentiation Factor 9
gDNA	Genomic DNA
GH	Growth hormone
GLI	Glioma-associated oncogene
GnomAD	Genome Aggregation Database
GnRH	Gonadotropin Releasing hormone
GNRHR	Gonadotropin Releasing Hormone Receptor
GO	Gene Ontology
GPCR	G-protein coupled receptors
GR	Genital ridge
GTF2A1L	General Transcription Factor IIA Subunit 1 Like
HA	Human influenza hemagglutinin
HARS	Histidyl-TRNA Synthetase 2, Mitochondrial
hCG	Human chorionic gonadotropin
HGMD	Human Gene Mutation Database
HH	Hypogonadotropic Hypogonadism
HHIP	Hedgehog Interacting Protein 1
HLA	Human leukocyte antigen
HMG	High-mobility group
HNF1	Hepatocyte Nuclear Factor 1
HOX	Homeobox
HPG	Hypothalamo-Pituitary-Gonadal
HSD17B3	Hydroxysteroid 17-Beta Dehydrogenase 3
HSD17B4	Hydroxysteroid 17-Beta Dehydrogenase 4
HSD3B1	Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta and Steroid Delta-Isomerase 1
HSD3B2	Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta and Steroid Delta-Isomerase 2
HSDs/KSR	Hydroxysteroid dehydrogenases/Ketosteroid reductase
HTS	High-Throughput Screening
IGF	Insulin-like Growth Factor
IGF1R	Insulin-like Growth Factor type 1 Receptor

IGV	Integrative Genomics Viewer
IHH	Idiopathic hypogonadotropic hypogonadism
Indel	Insertion or/and deletion
INHA	Inhibin Subunit Alpha
INHBA	Inhibin Subunit Beta A
INHBB	Inhibin Subunit Beta B
INSL3	Insulin-Like Factor 3
INSR	Insulin Receptor
INSRR	Insulin Receptor-Related Protein
ISCA	International Standards for Cytogenomic Arrays
ISP	Ion sphere particle
Kb	Kilobase
KISS1	Kisspeptin
KISS1R	KISS1 Receptor
KITL	Ligand for the receptor-type protein-tyrosine kinase
KO	Knock-out
KS	Kallmann Syndrome
LARS2	Leucyl-tRNA synthetase 2, mitochondrial
LB	Luria-Bertani
LBD	Ligand-binding domain
LCAH	Lipoic Congenital Adrenal Hyperplasia
LCH	Leydig Cell Hypoplasia
LH	Luteinizing Hormone
LHCGR	Luteinizing Hormone/Choriogonadotropin Receptor
LHX1	LIM Homeobox 1
LHX9	LIM Homeobox 9
Luc	Luciferase
M bp	Mega bp
MAF	Minor Allele Frequency
MAIS	Mild Androgen Insensitivity Syndrome
MAMLD1	Mastermind-like Domain-containing 1
MAP3K1	Mitogen-Activated Protein Kinase Kinase Kinase 1
MAP3K4	Mitogen-Activated Protein Kinase Kinase Kinase 4
MAPD	Median Absolute Pairwise Difference
MC2R	Melanocortin 2 Receptor
MEM	Minimum essential medium
MEN2	Multiple endocrine neoplasia type II
MGD	Mixed Gonadal Dysgenesis
MLPA	Multiplex Ligation-dependent Probe Amplification
MNV	Multiple nucleotide variant
MRAP	Melanocortin 2 Receptor Accessory Protein
MRI	Magnetic resonance imaging
MRKH	Mayer-Rokitansky-Küster-Hauser
mRNA	Messenger RNA
MTM1	X-linked Myotubular Myopathy
MutPred	Mutation Prediction
MYCBP2	MYC Binding Protein 2

NAD+	Nicotinamide adenine dinucleotide+
NADPH	Nicotinamide adenine dinucleotide phosphate hydrogen
NAV3	Neuron Navigator 3
NCBI NR	National Centre for Biotechnology Information Non-Redundant
NC-LCAH	Non-Classic Lipoic Congenital Adrenal Hyperplasia
NGS	Next Generation Sequencing
nIHH	Normosmic idiopathic hypogonadotropic hypogonadism
NNT	Nicotinamide Nucleotide Transhydrogenase
NOBOX	Newborn Ovary Homeobox-Encoding
NPPB	Natriuretic Peptide B
NR0B1	Nuclear Receptor Subfamily 0 Group B Member 1
NR5A1	Nuclear Receptor Subfamily 5 Group A Member 1
nsSNP	Non-synonymous single-nucleotide polymorphism
NTD	Amino-terminal domain
OCT3/4	Octamer-Binding Protein 4
OMIM	Online Mendelian Inheritance in Man
OT DSD	Ovotesticular DSD
OT2	Ion OneTouch 2
P450scc	P450 cholesterol side-chain cleavage
PAIS	Partial Androgen Insensitivity Syndrome
PANTHER	Protein Analysis Through Evolutionary Relationships
PBX1	Pre-B-Cell Leukemia Transcription Factor 1
PCOS	Polycystic Ovary Syndrome
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
PGM	Personal Genome Machine
PKA	Protein Kinase A
PKC	Protein Kinase C
PMDS	Persistent Müllerian duct Syndrome
POF	Primary Ovarian Failure
POI	Primary Ovarian Insufficiency
PolyPhen-2	Polymorphism Phenotyping v2
POR	P450 Oxidoreductase
PROVEAN	Protein Variation Effect Analyzer
PSI-BLAST	Position-Specific Iterative Basic Local Alignment Search Tool
PSMC3IP	Proteasome 26S Subunit, ATPase, 3-Interacting Protein
PTCH1	Patched 1
PTCH2	Patched 2
PTGDS	Prostaglandin D synthase
PTPN3	Protein Tyrosine Phosphatase Non-Receptor Type 3
QMPSF	Quantitative Multiplex Polymerase Chain Reaction of Short Fluorescent Fragments
qPCR	Quantitative PCR
RA	Retinoic acid
RAD51	DNA repair protein RAD51 homolog 1
RBM8A	RNA Binding Motif Protein 8A
rDNA	Ribosomal DNA
RET	REarranged during Transfection

Rluc	Renilla luciferase
RNA	Ribonucleic acid
RSPO1	R-Spondin 1
Runx1	Runt-related transcription factor 1
RXFP2	Relaxin-Family Peptide Receptor 2
SAP	Shrimp alkaline phosphatase
Scc	Side-chain cleavage
SDR	Short chain oxidoreductase
SEEP	*Sociedad Española de Endocrinología Pediátrica,* Spanish Society for Paediatric Endocrinology
SERKAL	SEx Reversion, Kidneys, Adrenal and Lung dysgenesis
SF1	Steroidogenic Factor 1
SNP	Single-nucleotide polymorphism
SHOX	Short stature Homeobox
SIDDT	Sudden Infant Death with Dysgenesis of the Testes
SIFT	Sorting Intolerant From Tolerant
SNV	Single-nucleotide variants
SOX	SRY-Box
SOX10	SRY-Box 10
SOX3	SRY-Box 3
SOX8	SRY-Box 8
SOX9	SRY-Box 9
SRD5A1	Steroid 5 Alpha-Reductase 1
SRD5A2	Steroid 5 Alpha-Reductase 2
SRY	Sex determining Region of Y chromosome
STAR	Steroidogenic Acute Regulatory protein
SVM	Support vector machine
SW1/SNF	SWItch/Sucrose Non-Fermentable
T	Thymine
T DSD	Testicular DSD
TACR3	Tachykinin Receptor 3
TALE	Three amino acid loop extension
TBX6	T-Box Transcription Factor 6
TES	Testis-specific enhancer
TESCO	Testis-specific enhancer core element
TGFb	Transforming Growth Factor-beta
TIMP1	Tissue Inhibitor of Matrix MetalloProteinase 1
TSPY	Testis Specific Protein Y
TSPY/TSPYL/SET/NAP-1	Testis-specific protein Y-encoded, Testis-specific protein Y-encoded like, Drosophila proteins Su(var)3-9, Enhancer-of-zeste and Trithorax, Nucleosome assembly protein 1
TSPYL1	Testis-Specific Y-Encoded-Like Protein 1
TSS	Transformation storage solution
UCSC	University of California Santa Cruz Genomics Institute
UGT	Undifferentiated gonadal tissue
UniProtKB	UniProt Knowledgebase
US	Ultrasound
UTR	Untranslated region

UV	Ultraviolet
VNN	Vanin 1
VUS	Variants of unknown significance
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
WNT4	Wnt Family Member 4 gene
wt	Wild type
WT1	Wilms Tumour 1 Transcription Factor
WWOX	Tryptophan, Tryptophan Domain-Containing Oxidoreductase
ZFPM2	Zinc Finger Protein FOG Family Member 2

INDEX

INTRODUCTION

As an introduction to this work I describe the genes and pathways that are involved in the development of the gonad, from the undifferentiated gonad to the completely developed testicles and ovaries. Implicated genes in the formation of internal and external genitalia that lead to sexual differentiation are also explained. Then, the widely varied disorders of sex development caused by impaired alterations in those previously mentioned genes are explained. Finally, I describe current clinical management of the sex anomalies and the role of next generation sequencing in the molecular diagnosis of this pathology.

1. SEX DEVELOPMENT AND GENETIC REGULATION

In humans and other mammals, sex development is a sequential process that involves a large number of genes and pathways acting to completely acquire functional gonads and consecutive differentiation of internal and external genitalia to finish with secondary sexual characterization at puberty (Figure 1A and 1B). The first step, sex determination refers to the decision of the bipotential gonad to continue through the testicular or ovarian pathway. In contrast to other organs, the early gonads have the potential to differentiate into two functionally different organs, testes or ovaries. Later, sex differentiation comprises the formation of the internal and external genitalia due to the sex-specific hormones secreted by the developing gonads and the secondary sexual characterization at puberty.

In males and females, sex determination and the correct development of the gonad and genitalia depend on a tightly controlled network of transcription factors and signalling pathways. Moreover, sex steroid hormone production is dependent on consecutive enzymatic steps and functional hormone receptors. Any changes in this delicate genetic system result in atypical sex development leading to disorders of sex development (DSD) in humans.

1.1. FORMATION OF THE UNDIFFERENTIATED GONAD

This process starts during the fourth week of embryonic life with the formation of the undifferentiated urogenital ridge, comprising the pronephros, mesonephros and metanephros, which gives rise to the adrenal cortex, gonadal and urinary systems, respectively (1).

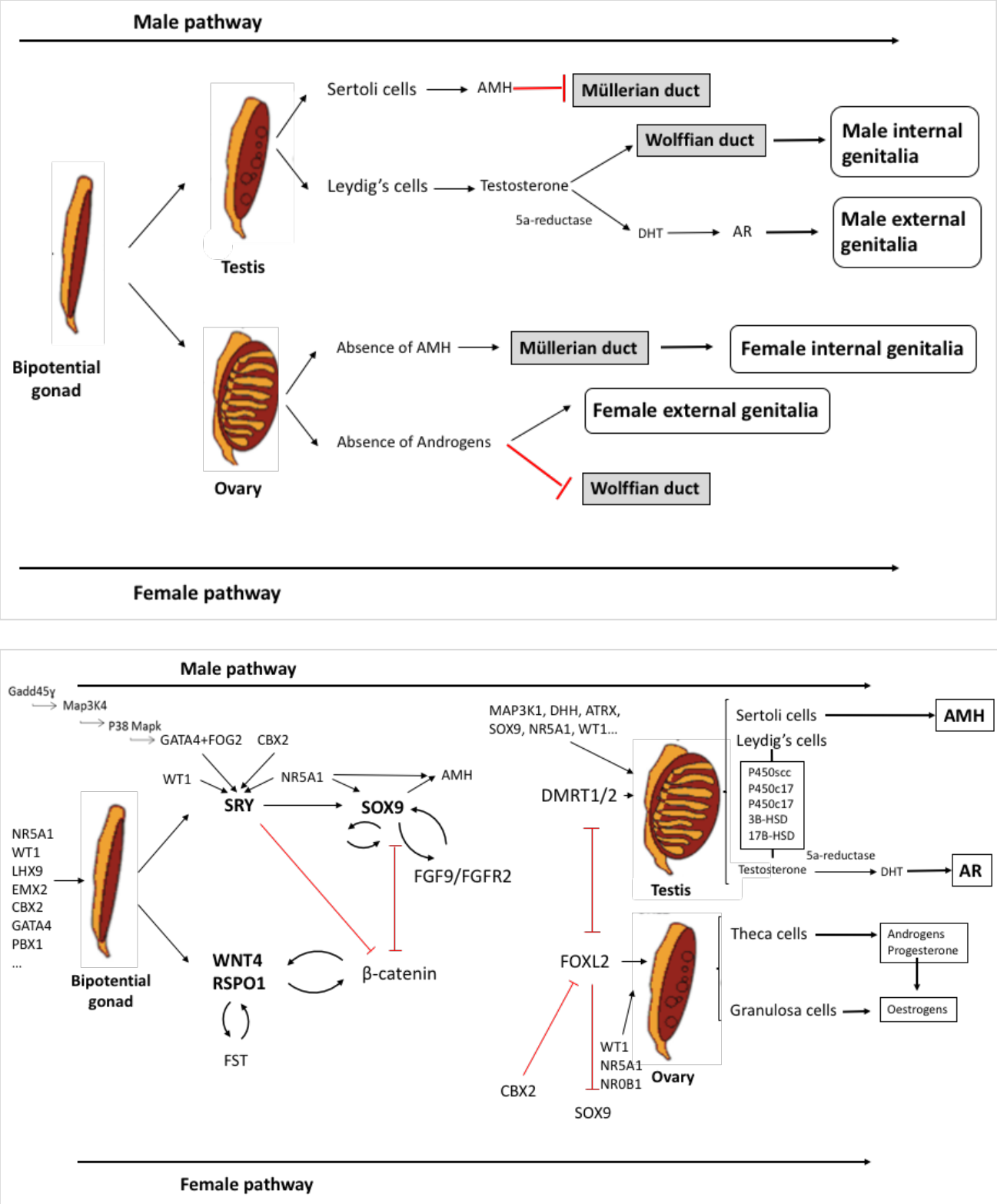

Figure 1. A, Overview of the development of male and female genitalia. B, Overview of the genes and pathways involved in the sex development.Black arrows indicate activation of a downstream target. Red lines ending in bars show repression of a downstream target. Modified from Ohnesorg T., 2014 (2).

In humans, the testis and the ovary arise from a sexually undifferentiated precursor, the genital ridge (GR) which appears as a thickening of the coelomic epithelium (CE) on the ventromedial surface of the mesonephros, creating a thick

epithelial cell layer. Together with the proliferation of the CE, the epithelial cells from the underlying membrane move towards the mesonephros and as mesenchymal cells colonize the space between the coelomic epithelium and the mesonephros. The mesenchymal cells of the GR are gonadal precursors that differentiate into interstitial and somatic supporting cells of the early differentiated gonad. For the testicular development, supplementary cells are recruited from the mesonephros into the genital ridge to increase the endothelial cell population that establish the male vascular system (1). Prior to gonadal determination, primordial germ cells, precursors of sperm and oocytes, migrate into the GR. In case of male sex determination, the mesonephros also gives rise to the mesonephric or Wolffian duct and to a primordial urogenital tissue that develops epididymis, vas deferens and seminal vesicles. In females, the paramesonephric or Müllerian duct is also comprised in the mesonephros (3).

At 5 weeks, the early gonad is bipotential and has the ability to differentiate into a testis or an ovary, depending on the presence or absence of the *SRY* (Sex determining region of Y chromosome) gene on the Y chromosome. However, at this point no morphological differences are observed between the XX and XY early gonads. Furthermore, genes that are later associated with the testicular fate such as *Sox9* (SRY-Box 9) or ovarian fate, as *Wnt4* (Wnt Family Member 4), are expressed equally in XX and XY gonads, until the beginning of the differentiation (4).

In mice at least seven genes, including *Nr5a1* (Nuclear Receptor Subfamily 5 Group A Member 1), *Wt1* (WT1 Transcription Factor), *Lhx9* (LIM Homeobox 9), *Emx2* (Empty Spiracles Homeobox 2), *Cbx2* (Chromobox 2), *Pbx1* (Pre-B-Cell Leukemia Transcription Factor 1) and *Gata4* (GATA Binding Protein 4) have been demonstrated to be necessary for the initial thickening of the coelomic epithelium, progression and maintenance of the genital ridge (5). Knock-out (KO) mouse studies have revealed that mutations or absence of these gene products might lead to a failure in the gonadal progress or to the presence of streak gonads (3, 6), emphasising the critical role of some genetic factors in both males and females.

Despite their early role in gonadal development, the loss-of-function of these genes in humans, result in a range of phenotypes that vary from complete gonadal dysgenesis to adult infertility (3).

1.1.1. Nuclear Receptor Subfamily 5 Group A Member 1 (*NR5A1*)

NR5A1, encoding steroidogenic factor 1 (SF1) is a member of the orphan nuclear receptor located on 9q33. It plays a central role in regulating adrenal development, gonadal determination and differentiation and in the hypothalamic-pituitary axis. In the mouse, *Nr5a1* is expressed in the GR from 9 dpc (days post-

conception) and thereafter, in the Sertoli and Leydig lineages, as well as in the adrenal gland. *Nr5a1*-null models showed gonadal agenesis and lack of adrenal glands (7).

Human SF1 is a 461 amino acid protein that contains a DNA (Deoxyribonucleic Acid) binding domain (DBD) with two zinc fingers, an accessory "A"-box that extends the DBD and mediates specific DNA binding, and a hinge region followed by a ligand-binding domain (LBD) that forms two activation function domains (AF-1 and AF-2) (Figure 2). The expression of *NR5A1* has been demonstrated during the formation of the bipotential gonad and adrenal glands (32-33 dpc), but also in gonad development and steroidogenesis (8).

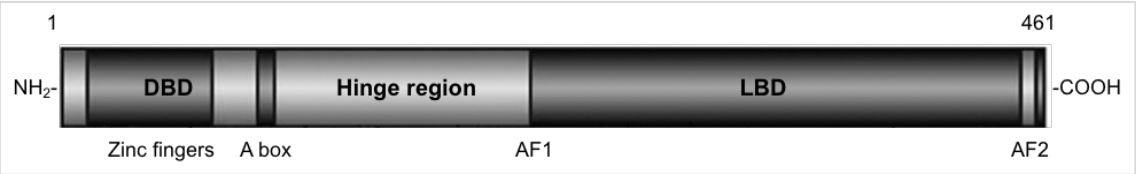

Figure 2. Diagram of the SF1 protein. The critical functional domains found in SF1, a DNA-binding domain (DBD) at the amino terminal (From amino acid 13 to 112), the flexible hinge region (amino acid 112-225) and a ligand-binding domain (LBD) (amino acid 225-458). Modified from Tantawy S., 2014 (9).

SF1 upregulates a number of testis-determining genes, such as *SOX9* in the somatic cells of early testis, anti-Müllerian hormone (*AMH*) in Sertoli cells, and the enzymes included in the steroidogenic system in Leydig cells (8, 10). While the expression of *Nr5a1* continues during testicular development in mice, it is downregulated during ovarian development. However, SF1 has been detected in Granulosa and theca cells of an adult ovary in humans but its role is unclear (8). It has been suggested that SF1 may promote the expression of *FOXL2* (Forkhead Box L2), which ultimately suppresses *SOX9* (8) and exemplifies the outcome of antagonistic pathways that compete to manage the differentiation and repression of supporting cell precursors to develop the bipotential gonad into a testis or an ovary.

In 1999, Achermann *et al* identified the first *NR5A1* heterozygous mutation in a patient with 46,XY gonadal dysgenesis, Müllerian structures and primary adrenal failure (11). Now, human variations in *NR5A1* show variable phenotypic expressivity and incomplete penetrance and have been widely characterized in 46,XY DSD, primary ovarian insufficiency (POI) and 46,XX testicular (T DSD) and ovotesticular DSD (OT DSD) (12-14).

1.1.2. Wilms' tumour Suppressor 1 (*WT1*)

The *WT1* gene, located on 11p13, codifies a zinc finger transcription factor and is known to act as a repressor or activator of numerous genes involved in the establishment of the bipotential gonad and in testicular development. In adults, *WT1* is

expressed in Sertoli and Granulosa cells. Among the multiple protein isoforms that are generated in mammals through complex regulatory mechanisms, two main WT1 isoforms are important in the gonadal development and are defined by the presence or absence of three amino acids, lysine, threonine and serine (KTS) between zinc finger three and four (15).

Wt1-null mice of both sexes fail to develop gonads and kidneys, suggesting an early role in the formation of the bipotential gonad (16). Moreover, the expression of *Wt1* begins with that of *Nr5a1* and the DNA-binding form WT1 (-KTS) acts together with Sf1 to ensure proper formation of the urogenital system (2). The other WT1 isoform (+KTS), is involved in the early testis development. Mice lacking the gene, presented male-to-female sex reversal through the failure in the Sry expression regulation, which influences cell proliferation and Sertoli cell differentiation (3). Other genes regulated by WT1 (-KTS) are *Sry, Amh, Sox9, Wnt4* and *Nr0b1* (Nuclear Receptor Subfamily 0 Group B Member 1) (10).

Mutations in the *WT1* gene have been classically identified in anomalies of testicular development leading to 46,XY DSD. Also, *WT1* gene deletions in the Wilms' tumour (OMIM 194070) and missense variants in the zinc finger domains or in intron 9 in Denys-Drash (DDS, OMIM 194080) and Frasier syndromes (FS, OMIM 136680), respectively, have been reported (2, 17). However, the application of NGS to the molecular diagnosis of DSD patients and associated syndromes has enlarged the phenotypic spectrum, like the *WT1* frameshift activating variant in a zinc-finger domain found in a 46,XX T DSD girl (18) or the six intronic and two missense changes identified in a cohort of women presenting with POF (Premature ovarian failure, OMIM 311360). Further *in vitro* analyses showed that missense variants downregulated *AMH* and *CDH1* (Cadherin 1), while increased expression of *FSHR* (Follicle Stimulating Hormone Receptor) and *CYP19A1* (Cytochrome P450 Family 19 Subfamily A Member 1) (19).

1.1.3. GATA binding protein 4 (*GATA4*)

The human *GATA4* gene on chromosome 8p23.1 encodes an essential transcription factor for the developing gonad and heart (20). GATA proteins have two zinc fingers (ZNI and ZNII), which are highly conserved and are necessary for protein–protein interactions with other transcription factors (Figure 3). The sequences of the carboxyl-terminal and amino-terminal domains are required for the DNA recognition and binding, and contribute to stability (21).

Gata4-null mice die due to severe abnormalities in heart tube formation and ventral morphogenesis. *Gata4* expression and function seems not only essential for normal testicular and genital development (22), but also to initiate formation of genital ridge. Furthermore, mouse embryos conditionally lacking *Gata4*, show no signs of GR

formation and failure of LHX9 and SF1 gonadal markers, suggesting an earlier role in gonadogenesis than *Sf1, Wt1, Lhx9* and *Emx2* (3). In humans, GATA4 interacts with several proteins, including SF1, WT1, and FOG2 (Friend Of GATA Protein 2) to regulate the expression of sex determining genes *SRY, SOX9*, and *AMH*, as well as genes involved in sex differentiation such as *STAR* (Steroidogenic acute regulatory protein)*, CYP17A1* (Cytochrome P450 family 17 subfamily A member 1)*, CYP19A1, INHA* (Inhibin subunit alpha) and *HSD3B2 (*Hydroxy-delta-5-steroid dehydrogenase, 3 Beta and steroid delta-Isomerase 2) (22-24).

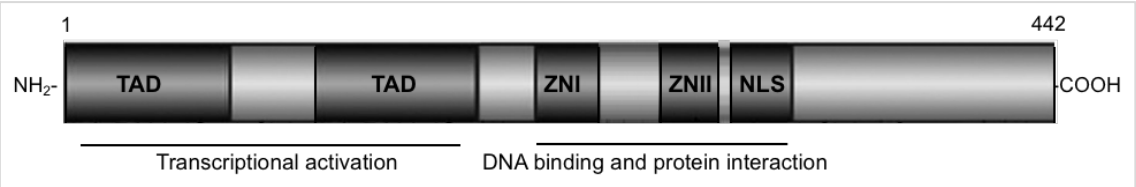

Figure 3. Scheme of the structure of the GATA4 protein. GATA4 contains two distinct zinc finger domains (ZNI and ZNII) and a carboxi terminal nuclear localization sequence (NLS) which consists of a DNA-binding domain and a protein-protein interaction domain. Transcriptional activation domains (TAD) are located in the N-terminus. Modified from Martinez de LaPiscina I., 2018 (25).

GATA4 haploinsufficiency has been described in patients with different forms of congenital heart defects (CHD) since 1999, while only few studies have reported mutations related to a 46,XY DSD phenotype (12, 25-29).

1.1.4. Chromobox Homolog 2 (*CBX2*)

This gene, located on 17q25.3, codifies for a component of the polycomb group complex of proteins responsible for epigenetic regulation through the recognition of methylated histones. Among other adrenal and spleen defects, the targeted deletion of *Cbx2* caused male-to-female sex reversal in XY gonads. Posterior forced expression of *Sry* and *Sox9* in this murine model resulted in smaller testes and indicated its role in testis differentiation through regulation of *Sry*. Cbx2 is also involved in the upregulation of other genes, such as *Nr5a1* and *Wt1,* and downstream targets such as *Sox9* (30). Variants in this gene are associated with gonadal dysgenesis in humans (31, 32).

1.1.5. Pre-B-Cell Leukemia Transcription Factor 1 (*PBX1)*

PBX1 encodes a three amino acid loop extension (TALE) class homeodomain protein that plays an important role in the adrenal and urogenital development. $Pbx1^{-/-}$ mice decreased cell proliferation in the genital ridge and thus, testicular development was impaired. In the female reproductive system, the inactivation of the gene led to the absence of uterus and vagina (33). More recently, a missense mutation in the TALE domain was found in a child with 46,XY gonadal dysgenesis. This study suggested that

the loss-of-function results in an altered interaction with CBX2 and EMX2, driving to irregular downstream events that interrupted the gonadal development in the patient (34).

1.1.6. Empty-Spiracles Homeobox 2 (*EMX2*)

EMX2, encoding a homeobox-containing transcription factor is expressed in the urogenital system and in the early stages of the cerebral cortex and olfactory system. The role of this protein in the early gonadal development was supported when KO mice (*Emx2*$^{-/-}$) showed complete gonadal and genitourinary system agenesis (35). As mentioned, it interacts with other partners to bind DNA regulatory targets and activate transcription (34). In males, few deletions that encompass the gene in 10q26.11 have been reported (35).

1.1.7. Lim Homeobox 9 (*LHX9*)

LHX9, encodes a protein of the LIM gene family and contains a homeodomain and two cysteine domains involved in protein interactions. During mice embryonic development, Lhx9 protein transcripts are present in urogenital ridge. In *Lhx9*-deficient mice, somatic cells of the genital ridge don't proliferate and gonad fails to form. In the absence of testosterone and AMH the genetically XY mice develop a female phenotype with reduced expression of *Nr5a1* (36). *In vitro* studies demonstrated the regulation of *Nr5a1* gene through the direct binding of Lhx9 and Wt1 to the promoter (22, 37). Recently, two heterozygous variants in *LHX9* have been reported in two 46,XY DSD patients with variable degrees of undervirilization (38).

1.2. MOLECULAR GENETICS IN MALE GONADAL DETERMINATION

In humans, gonadal differentiation occurs from 6 to 10 gestational weeks and in mice from embryo day 11 to 12. During sex development, the gonadal precursors differentiate into two somatic cell lineages. The supporting cell precursors contribute to Sertoli and Granulosa cells in the XY and XX gonads respectively, whereas steroidogenic progenitors give rise to Leydig cells in the male gonads or theca cells in the ovary. The differentiation of the germ cell depends on the surrounding niche, then the sex role chosen by the cell lineage begins a differentiation cascade that finishes with a complete functional gonad (39).

Differentiation of the bipotential gonad into testis or ovary is the result of the antagonistic male and female pathways that take part to control the differentiation of supporting cell precursors, via regulation of *SOX9* and others. The commitment of the Sertoli or Granulosa cell fate supposes the activation of one program and the

repression of the alternative pathway of development as well as the control of the gonad field and the fate of other somatic lineages and germ cells. The *SRY* gene works as a switch that guides the initial gonadal determination to a testicular fate, while *RSPO1 (R-Spondin 1)* and *FOXL2,* among others, direct gonadal development to an ovarian destiny.

SRY gene is necessary and sufficient to initiate testicular development. At around week 6 of gestation in humans and embryo day 11 in mice, pre-Sertoli cells express the *SRY* in the XY gonad and together with *NR5A1*, trigger the testis development program through activation and maintenance of the *SOX9* signalling pathway. Other important target genes and components of the early testis development are *NR0B1, AMH* and *GATA4* and its co-factor *FOG2* (40). In the absence of *SRY*, female-specific pathway is initiated which includes genes such as *RSPO1, WNT4* and *NR0B1* to promote the organogenesis of the ovaries (22).

1.2.1. Sex determining region Y *(SRY)*

This single-exon gene encodes a member of the high-mobility group (HMG) box family of transcription factors. It is characterized by the presence of the highly conserved HMG box DNA-binding domain and acts as the testes determining factor. Approximately 80% of the 46,XX T DSD and 10% of OT DSD had *SRY* material translocated into one of their X chromosomes. Moreover, nearly 20% of the 46,XY DSD females with gonadal dysgenesis had a loss-of-function variant in the *SRY* gene. Female mice transgenic for *Sry* developed testes, which finally demonstrated that this gene triggers male sexual determination and testis formation (2).

In mice, the expression of *Sry* begins in pre-Sertoli cells at 10 days post-conception (dpc), stimulating the development of testicular cords and Leydig cell formation. Its expression reaches the highest levels at 11.5 dpc and diminishes at around 12.5 dpc, thus *Sry* is not required for the maintenance or function of the testis. In contrast, in humans *SRY* is expressed in male somatic cells at 6 weeks and continues to be expressed in the adult testis (2). It seems that SRY acts upregulating *Sox9* and stimulates Sertoli cell formation and thus, testicular differentiation. In the absence of Sry, *Sox9* is suppressed and female gonadal development begins (41).

Positive regulators of the *Sry* gene expression include GATA4, FOG2, WT1 and NR5A1. As mentioned before, Cbx2 regulates the expression of *Sry* itself (30).

1.2.2. SOX family genes

The SOX (SRY-Box) genes encode transcription factors involved in the development of several tissues and cell lineages. These transcription factors bend the

DNA after binding to the minor groove, which allows the interaction of activating or repressing cofactors to the promoter or regulatory regions of target genes. Thus, Sox proteins modulate the expression of downstream genes to correctly differentiate and preserve the distinct tissues and cell types (42). Besides *SRY*, three other members of the SOX family are expressed in the XY gonad but *SOX9* was the first to be related to testis development.

In mice, *Sox9* is slightly expressed in the genital ridge at embryo day 10.5. Immediately after the onset of *Sry* expression in XY pre-Sertoli cells, expression levels of *Sox9* increase. At this time Sox9 becomes sexually dimorphic as XX gonads decrease its expression levels. Up to date *Sox9* is the only known target of SRY (2).

Sox9 is also essential in the chondrogenic lineage specification and differentiation. Then, *SOX9* loss-of-function alterations cause campomelic dysplasia (OMIM 114290), a condition characterized by skeletal defects and typical facies. Approximately 75% of 46,XY individuals present with campomelic dysplasia combined with complete or partial gonadal dysgenesis (2, 43). Furthermore, genomic rearrangements upstream the regulatory region of the *SOX9* gene have been identified in a wide spectrum of DSD, including 46,XX T DSD and 46,XY OT DSD (44, 45). Studies in Oddsex mice showed sex reversal phenotypes when a 150Kb deletion occurred upstream *Sox9* (46) and a testis-specific enhancer (TES) was found as a gene expression regulator. Sf1 and Sry bind the testis-specific enhancer core element (TESCO), located 13Kb upstream *Sox9*, to upregulate its expression. Once Sox9 is increased, it displaces Sry and synergizes with Nr5a1 to regulate its own expression by a positive feedback loop (47). This positive regulatory loop occurs via Fgf9 (Fibroblast growth factor 9), which in turn activates the FGF receptor 2 (*Fgfr2*) and Prostaglandin D synthase (*Ptgds*), resulting in the secretion of Prostaglandin D2, Amh or Vnn (Vanin 1) (48, 49). Other genes have been identified downstream Sry and Sox9 and might act as potential targets of these, but evidence is still insufficient. On the other side, Sox9 represses the WNT4/FOXL2 pathways, involved in the ovarian formation. *Fgf9/Wnt4* double mutant XY mice developed male gonads which indicated that the principal role of Fgf9 is to repress *Wnt4* (50).

Studies in murine models have suggested that *Sox9* could effectively replace *Sry* as the sex-determining gene. Transgenic *Sox9* XX mice developed testes while KO XY embryos led to the formation of ovaries. Moreover, specific annulation of the gene in Sertoli cells resulted in normal testicular development and fertile mice (46, 51). This phenomenon may be due to the functional redundancy amongst the Sox gene family members, which means that functionality of the Sertoli cells might have been rescued by either *Sox8* or *Sox10* (42).

SOX8 and *SOX10* are also expressed in the XY gonad. In mouse model, *Sox8* is expressed in the testis cords at embryo day 13.5 and *in vitro* it upregulates the expression of anti-Müllerian hormone (Amh). *Sox8*-deficient male mice have abnormal Sertoli cells and infertility, but normal male internal structures (22, 52). Moreover, *SOX8* variants have been recently described in a range of reproductive anomalies (53). In contrast, *Sox8* and *Sox9* double KO mice lack testis cord, which supports the functional redundancy between these two proteins in testicular differentiation, either by the mediation of the HMG-box or to the inadequate expression of HMG-box proteins (42). The transgenic expression of *Sox10* in XX gonads has resulted in the development of testes and male phenotype, suggesting that Sox10 is able to activate transcriptional targets of Sox9 (2). On the other hand, *SOX3* a single-exon gene located on the X chromosome is expressed in brain and gonads. Studies in animal models demonstrated that *Sox3*-null female mice were able to develop defective ovaries while male mice (-/Y) developed testes with loss of germ cells and disruption of seminiferous tubules. They conclude that *Sox3* is important for testicular differentiation and gametogenesis (54). In humans, *SOX3* gene variants have been identified in X-linked mental retardation (OMIM 309530) and in a 46,XX DSD patient with bilateral ovotestes (55).

1.2.3. Fibroblast growth factor 9 (*FGF9*)

The *FGF9* gene encodes a member of growth factors critical in cell proliferation, survival, migration and cell differentiation (22). It is expressed after *Sry* in the bipotential gonad. Studies with *Fgf9*-deficient mice have shown different results, either a male to female sex-reversed phenotype or mutant mice that died with gonadal abnormalities after birth (56). Some data support that Sry and Nr5A1 begin the upregulation of *Sox9*, increasing the expression of *Fgf9* which in turn further increases *Sox9* expression (48). Therefore, in the absence of Fgf9, the expression of *Sox9* is not maintained, and Sertoli cells don't differentiate. Then, the balance between *Fgf9* and *Wnt4* favours to female development

The molecular function of FGF9 is mediated through its receptor FGFR2, which is embedded in the membrane of progenitor Sertoli cells. Although an important crucial role of both genes is suspected in human sex development, a single variant in *FGFR2* has been identified in a patient with 46,XY sex reversal and craniosynostosis (57) and a *FGF9* gain of copy number in a 46,XX male with elevated FSH (Follicle Stimulating Hormone) level, low testosterone and ambiguous genitalia (58).

1.2.4. Nuclear Receptor Subfamily 0 Group B Member 1 (*NR0B1*)

The *NR0B1* gene encodes DAX1 (DSS-AHC, dosage-sensitive sex reversal, adrenal hypoplasia congenital critical region on the X chromosome protein 1), on

chromosome Xp21.2, which is an orphan nuclear hormone receptor. The carboxil-terminal region of the protein, shows similarities to ligand-binding domain of other nuclear receptors and contains twelve helices, whereas the N-terminal represent a domain involving 3.5 repeats of a 65 to 67 amino acid motif with two putative zinc fingers, which possibly act as nucleic acid-binding domain (41). DAX1 acts throught the retinoic acid receptor as an inhibitory protein of the SF1 nuclear receptor, estrogen receptor and androgen receptor (59).

In humans, loss-of-function mutations or deletions of *NR0B1* results in patients with X-linked congenital adrenal hypoplasia (AHC, OMIM 300200) and hypogonadotropic hypogonadism (HH). Patients with AHC have poorly differentiated adrenal gland after fetal stage and present with steroidogenesis disturbances and a disorganized testis cord (59). On the other side, duplications of *NR0B1* have been shown to cause 46,XY gonadal dysgenesis and a female phenotype (60).

It has been proposed that DAX1 functions as a dosage-based mechanism. In 46,XY males, *NR0B1* is maintained inactive by SRY and the testicular differentiation pathway continues. When duplications occurs, SRY is not enough to block the amount of DAX1 and the female pathway is followed after DAX1 has repressed testis-specific enhancer (TES), the enhancer of SOX9, and the testicular formation is stopped. Lack of testosterone and AMH blocks development of external genitalia and the regression of Müllerian ducts (59).

As mentioned*, NR0B1* also promotes the ovarian pathway and loss-of-function heterozygous and homozygous nonsense mutations in XX individuals have been associated with delayed puberty and hypogonadotropic hypogonadism (61, 62), which disagrees with mouse modelling studies. Due to the dimorphic expression between ovary and testis *NR0B1* was considered an “anti-testes” gene. It has been indicated that DAX1 inhibits the interaction of SF1 and WT1 and thus, reduces the activation of SOX9, which blocks steroidogenesis and the production of anti-Müllerian hormone (59).

This was validated when $Dax1^{-}/Y^{pos}$ murine models, testes were not differentiated due to Sox9 low levels and absence of Amh (63), suggesting that Sry and Dax1 may act to upregulate *Sox9*. These genetic males had female external and internal genitalia. SRY and DAX1 interact in early development of the gonadal ridges and are expressed in both testicular and ovarian tissue. In the absence of *Nr0b1* male mice showed complete gonadal sex reversal, suggesting its role in testicular determination (64). After sex-determining period, Dax1 is downregulated in the testes to be upregulated again at 17.5 days post-conception. This break is necessary for the male correct differentiation, since Dax1 inhibits the expression of *Amh* and *Hsd3B2*,

both important in the testicular development (65). In differentiating ovaries, Dax1 is upregulated downstream Wnt4, which indicates that Dax1 support Sry (66).

1.2.5. GATA binding protein 4 (*GATA4*) and Friend of GATA 2 (*FOG2*)

The role of *GATA4* in male sex determination has been explained before (See 1.1.3, page 6).

FOG2 is a zinc finger cofactor that binds the N-terminal zinc finger of the GATA4 protein and acts as a transcriptional coactivator or corepressor. *Fog2 (or ZFPM2,* Zinc Finger Protein, FOG Family Member 2*)* is co-expressed with *Gata4* in XY gonad during Sertoli cell formation and although the exact mechanism is unknown, the physical interaction between Gata4 and fog2 proteins is needed for testicular development (67). Indeed, an impaired interaction between GATA4 and FOG2 was reported in a 46,XY DSD patient presenting with a missense variant in the zinc finger of *GATA4* (27). In humans, mutations in the *FOG2* gene have demonstrated to cause 46,XY DSD due to modulation of the expression of target genes involved in testes development (67). It was found that *Fog2*-null mice died before embryo day 15.5 with a complex cardiac defect, failure of testis differentiation and reduced gene expression, such as those required for Sertoli cell differentiation and Leydig cell function. Moreover, haploinsufficiency of either *Gata4* or *Fog2* is sufficient to induce gonadal sex-reversal in murine model.

1.2.6. Desert Hedgehog (*DHH*)

DHH, located on 12q12 and encoding a 396-amino acid protein, is a member of the hedgehog family of signalling molecules. Among them, *Dhh* is expressed at 11.5 days post-conception in Sertoli cell precursors, after the expression of *Sry* had begun. This protein is also expressed in Schwann cells, vascular endothelium, endocardium and seminiferous epithelium of the mouse embryo (39) and no expression has been detected in female gonads. $Dhh^{-/-}$ mutant mice were sterile and showed abnormal peripheral nerves (68). *Dhh* signalling is also required for the upregulation of *Nr5A1* and receptor Patched 1 (*Ptch1)* expression in Leydig cells (69). Several mutations have been associated with complete or partial 46,XY gonadal dysgenesis (22, 69).

1.2.7. Doublesex and Mab-3 Related Transcription Factor 1 (*DMRT1*) and 2 (*DMRT2*)

DMRT1 protein is a transcription factor with a DNA-binding domain similar to a zinc finger, named DM (DNA-binding motif) domain. In mice, *Dmrt1* is expressed in the primordial gonads of both sexes at 10.5 days post-conception (dpc). Later it is only increased and maintained in Sertoli and premeiotic germ cells until adulthood,

whereas its expression is decreased in the ovary. In humans, *DMRT1* is expressed in undifferentiated XY gonadal primordium at 6 weeks of gestation and in Sertoli cells between gestation weeks 8 to 40. During childhood and postpuberty it is also abundant in spermatogonia. At about gestation week 20, it has been detected in oogonia and oocytes but is downregulated when germ cells enter at meiosis (70).

Studies in mice revealed that the gene important in the maintenance and growth of the testis in the postnatal and adult period, whereas in humans is required for the development of the testis during foetal period (71). *DMRT1* is involved in sex determination and gonadal development across a broad range of species. In humans, the role of *DMRT1* in male sex determination was confirmed when subtelomeric deletions causing sex reversal and gonadal dysgenesis in 46,XY individuals, were mapped to gene location 9p24.3.

In mice, Dmrt1 activated testis-promoting genes including *Sox9, Sox8* and *Ptgds*, and repressed ovary-promoting genes such as *Foxl2, Wnt4* and *Rspo1*, as well as retinoic acid (RA) signalling and its feminizing effects (70, 72). Moreover, it has been shown that the loss of *Dmrt1* in mouse Sertoli cells activates *Foxl2*, and thus, Sertoli cells differentiate to Granulosa and theca cells, with its consequent oestrogen production (73). *Dmrt1*$^{-/-}$ mice showed postnatal defects in testis maturation and germ cells died at 7 dpc. This gene may not have a critical role in the early stages of gonadogenesis in mice, however the default in germ cell migration and survival is related to the expression pattern in both somatic and germ cell types.

Further analysis of the 9p *locus* showed that *DMRT2* also maps to 9p24.3. The two DMRT1 and DMRT2 are 80% identical in the core region of the DM domain and are expressed in testis, suggesting that it might potentially be involved in 9p sex reversal. Although it was not believed to be a common cause of 46,XY sex reversal, it has been probed that deletions in 9p include the removal of both *DMRT1* and *DMRT2* genes (71). Individuals with distal monosomy 9p presented with both normal or ambiguous external genitalia and varying degrees of mixed gonadal dysgenesis, including fibrous streak gonads, hypoplastic testicles, or ovotestes (74, 75).

1.2.8. α-Thalassemia/Mental retardation X-linked gene (*ATRX*)

The ATRX protein is a member of the SW1/SNF (SWItch/Sucrose Non-Fermentable) DNA family, involved in chromatin remodelling. It is a nuclear protein found at heterochromatic structures, such as pericentromeric heterochromatin and ribosomal DNA (rDNA) repeats. These repeat elements are essential for the transcriptional regulation. ATRX modulates gene expression by binding to G-rich tandem repeat sequences. Variations in the size of the tandem repeats associated to

diverse gene expression levels could explain the phenotypic variability observed within families with the same mutation (39).

The amino-terminal of the ATRX protein is an ATRX-DNMT3-DNMT3L (ADD) domain with a plant homeodomain-like zinc finger. In the middle of the protein lies the helicase/adenosine triphosphatase (ATPase) domain that displays nucleosome-remodelling activity. Most of the mutations described so far, lie in the helicase and ADD domain. Mutations in this last region are associated with more severe psychomotor impairments. The C-terminus encodes the conserved P-box domain, involved in transcriptional regulation and a Q-box responsible for protein interaction. This region seems to be in charge of the urogenital formation. Patients lacking this C-terminal region have severe urogenital anomalies such as female external genitalia with streak gonads or ambiguous genitalia (76).

Mutations in this gene cause alpha-thalassemia/mental retardation syndrome X-linked (ATR-X, OMIM 301040), a sex-linked condition characterized by alpha-thalassemia, psychomotor retardation, distinct dysmorphic facies, skeletal, renal and cardiac anomalies, as well as genital abnormalities (76). Testicular irregularities are present in 80% of the XY patients, ranging from complete gonadal dysgenesis to mild hypospadias or micropenis. Histopathological analysis of dysgenetic testis in two *ATRX* positive patients confirmed its function after sex determination, possibly during neonatal period (77).

1.2.9. Anti-Müllerian hormone (*AMH*) and its receptor *AMHR2*

AMH is a member of the transforming growth factor-β family and is released by immature Sertoli cells at sex differentiation in foetal life. *AMH* expression begins in the pre-Sertoli cells, declines in the perinatal period and is highly maintained until Sertoli cells are mature in puberty. In females, its expression begins during the perinatal period and remains low in Granulosa cells until menopause (39). AMH is coded by a 5-exon gene located on 19p13.3 (78) and is translated as a dimeric precursor comprising 2 polypeptide chains, containing a large N-terminal fragment important for the correct folding of the protein, and smaller C-terminal mature domain.

In contrast, *AMHR2* (Anti-Müllerian Hormone Receptor Type 2) gene, located on 12q13 contains 11 exons and codifies for a protein with an intracellular domain with serine/threonine kinase activity, a transmembrane domain and an amino-terminal extracellular domain that binds AMH.

Anti-Müllerian hormone is responsible for regression of the female Müllerian ducts during male sex differentiation and mutations in *AMH* or its receptor *AMHR2* cause persistent Müllerian duct syndrome (PMDS, OMIM 261550). Transcription

factors NR5A1, SOX9, GATA4 and WT1 regulate the transcription of *AMH* through a critical DNA-binding region on its minimal promoter. In mice, SOX9 demonstrated to be essential to start *Amh* expression, whereas NR5A1 interacts with either *SOX9* or *WT1* through its DNA-binding domain to upregulate Amh transcripts. GATA4 alone or together with NR5A1 upregulates the expression of *AMH* too. On the other side, NR5A1 can cooperate physically with DAX1 to repress the expression of *AMH*, but with lower affinity compared to NR5A1-WT1. In women, AMH has an important role as an ovarian reserve marker (79).

1.2.10. Mitogen-activated protein kinase pathway (*MAP3K1*)

The mitogen-activated protein kinase (MAPK) pathway was linked to the early stages of male gonadal development when the *Map3k4* (Mitogen-Activated Protein Kinase Kinase Kinase 4) gene responsible for a male to female sex reversal in XY mice was discovered by Bogani *et al*. Moreover, expression analysis in *Map3k4* KO models showed a decrease of Sox9 and Sry production which lead to the absence of Sertoli cells and the development of ovaries (80).

Later, sequencing analyses revealed gain-of-function mutations in the *MAP3K1* (Mitogen-Activated Protein Kinase Kinase Kinase 1) gene in patients with streak gonads and female genitalia (12, 81, 82). In mice, *Map3k1* is expressed during testicular development at embryo day 11.5 but KO mice do not have testicular abnormalities. The molecular mechanism behind the gain-of-function mutations in humans differs from mouse models. Mutations in *MAP3K1* gene stimulate the ovarian-determining pathway, resulting in augmented β-catenin, *WNT4* and *FOXL2* expression and decreased *SRY* and *SOX9* (82).

1.2.11. Insulin receptor tyrosine kinases and insulin-like growth factors

Insulin and insulin-like growth factors (IGFs) are also involved in cell regulation during embryonic and postnatal development and mediate their functions by membrane-associated tyrosine kinase receptors. Genes codifying for *Insr* (Insulin receptor), *Igf1r* (Insulin-like growth factor type 1 receptor) and *Insrr* (Insulin receptor-related protein) are necessary for testicular determination pathway. Triple (*Insr/Igf1r/Insrr*) KO XY mice showed complete male-to-female sex reversal and reduced expression levels of *Sry* and *Sox9* (83). Later, it was demonstrated that only *Insr* and *Igfr1* were involved in genital ridge (GR) development and its absence is enough to cause the XY sex reversal phenotype (84). These mice embryos showed delayed *Sry* expression at embryo day 11.5, downregulation of several genes and delayed ovarian differentiation until embryo day 16.5. Genes which expression was decreased were *Wt1, Lhx9, Nr5a1, Gadd45g* and *Wnt4*, indicating the influence of insulin/IGF signalling in the GR development.

In humans, sequence variants in these genes have not been described in DSD.

1.2.12. Testis-Specific Y-Encoded-Like Protein 1 (*TSPYL1*)

The TSPYL1 protein is a member of the TSPY/TSPYL/SET/NAP-1 (Testis-specific protein Y-encoded, Testis-specific protein Y-encoded like, Drosophila proteins Su(var)3-9, Enhancer-of-zeste and Trithorax, Nucleosome assembly protein 1) family of chromatin modifiers. Its expression is limited to male germ cells and contains a NAP and SET domain that interact with cyclin B to regulate the cell cycle. Actually, NAP domain containing proteins act as histone chaperones to control chromatin and nucleosome assembly and control gene expression of a variety of proteins, like HMG box containing proteins SRY, SOX8 and SOX9 key elements. This fact may propose the mechanism whereby TSPYL1 causes defects in testicular development (85).

Variants in *TSPYL1* were recognized in an autosomal recessive syndrome called Sudden infant death with dysgenesis of the testes (SIDDT, OMIM 608800) (86). Affected males presented with either 46,XY partial or complete gonadal dysgenesis at birth and developed later viscero-autonomic dysfunction followed by death. Sexual development in 46,XX individuals was normal (86). Further studies associated *TSPYL1* gene variants with isolated testicular dysgenesis (85).

1.2.13. Mastermind-like domain-containing 1 (*MAMLD1*)

MAMLD1 was firstly identified as the responsible gene for X-linked myotubular myopathy (MTM1, OMIM 310400) and abnormal genitalia in two patients with a deletion in Xq28. MAMLD1 acts as a transcription factor by transactivating the promoter of the non-canonical Notch targeted Hes promoter, and increases testosterone through the regulation of SF1. The possible role in sex differentiation was probed when studies in mouse models showed increasing *Mamld1* expression in fetal Sertoli and Leydig cells at embryo day 12.5 to E14.5. Moreover, *Mamld1*-null mice showed reduced testicular expression of *Cyp11a1, Cyp17a1, Hsd3b1, Star* and *Insl3*, all Leydig-cell specific genes, although external genitalia and reproduction capacity was normal (13)(87).

In humans, *MAMLD1* sequence variations have been detected in 46,XY DSD in patients with a broad range of mild to severe phenotypes (87). A role in ovarian development was also suggested after a gain-of-function mutation was found in a 46,XX DSD woman with virilization (88). However, further evidence has questioned whether *MAMLD1* gene variants are able to explain the DSD phenotype. Besides the *MAMLD1* variants found in normal population (89) and the aforementioned KO mice with normal genital phenotype, promoter activation assays show similar results in

MAMLD1 mutations compared to Wt (87, 89-91). In fact, an oligogenic mechanism explaining the phenotype in *MAMLD1* positive patients has been recently published (91).

1.2.14. WW Domain-Containing Oxidoreductase (*WWOX*)

WWOX (Tryptophan, Tryptophan Domain-Containing Oxidoreductase) gene is located at 16q23.1-q23.2 and is a tumour suppressor gene with causative rearrangements described in several cancers. This gene encodes a protein with two WW domains at the amino-terminal and a short chain oxidoreductase (SDR) domain in the middle, thought to have a role in steroid metabolism. The WW domains are involved in protein-protein interactions. KO mouse models showed defects of Leydig cell function. This is possibly explained because WWOX inhibits the Wnt/β-catenin pathway in a dose-dependent manner, even though this inhibition is reduced when the SDR domain is eliminated. In humans, *WWOX* is expressed in pituitary, testis and ovary and has been suggested to have a role in gonadotrophin or sex-steroid biosynthesis (92). Missense variants as well as duplications and exon deletions have been identified in 46,XY DSD patients with ambiguous genitalia and in a 46,XX DSD patient that was referred with primary amenorrhea and hypergonadotropic hypogonadism (92, 93).

1.3. PATHWAYS INVOLVED IN THE DETERMINATION OF THE FEMALE GONAD

Although the noteworthy increase in the identification of the essential regulatory genes associated with ovarian development, little is known compared to testicular formation.

Female development has been considered the passive pathway in mammals, since the removal of male embryonic gonads before sex determination or the loss-of-function mutations in early gonad genes led to a female phenotype (2). In the last years, genetic factors involved in ovarian development have been identified and it has become apparent that both male and female pathways actively suppress the alternate state to maintain the somatic sex identity of the gonad as either male or female.

In the absence of *SRY* gene, the undifferentiated gonad activates a signalling pathway responsible for commitment to an ovarian fate. Ovarian sex determination requires Rspo1/Wnt4/β-catenin and Foxl2 signalling pathways.

1.3.1. Rspo1/Wnt4/β-catenin signalling pathway

WNT4 (Wingless-type MMTV integration site family member 4) and RSPO1 (R-spondin family member 1) ensure appropriate levels of stable β-catenin signalling,

essential for starting the ovarian development (94). In mouse models, the activation of β-catenin in male gonads promoted ovarian development, leading to XY sex reversal, while conditional ablation of *Ctnnb1* (Catenin B 1 gene) in XX somatic cells did not interrupt ovarian differentiation, suggesting that the loss of the *Ctnnb1* causes sex reversal when happening in both somatic and germ cells (95).

Rspo1 and *Wnt4* expression is increased in the ovary at embryo day 12.5 and although both are secreted by the somatic cells, Rspo1 has been detected in the membrane of germ cells too. In mice, loss-of-function mutations in either *Rspo1* or *Wnt4* cause complete or partial XX sex reversal and ovotestes (96). In *Wnt4*-deficient mice the expression of Rspo1 remains undamaged, whereas in *Rspo1*-null mice Wnt4 expression is absent. Thus, Rspo1 might collaborate with Wnt4 in a linear pathway to stabilize β-catenin. *Igfr1* expression is also diminished in these deficient mice, suggesting that feedback interactions may exist between these two signaling pathways. Rspo1 and Wnt4 also contribute in the early coelomic proliferation and double mutant null mice display a significant reduction in cell proliferation between embryo days 10.5 and 11.5 (97).

In humans, mutations in *RSPO1* are associated with a syndrome characterized with palmoplantar hyperkeratosis, predisposition to squamous cell carcinoma and 46,XX DSD (98). Regarding *WNT4*, dominant heterozygous mutations have been reported in association with anomalies of Müllerian structures, androgen excess and different virilization degrees (97). A homozygous mutation was found in three XX foetuses with an embryonic lethal syndrome (SERKAL, OMIM 611812) encompassing female-to male sex reversal and kidney, adrenal and lung dysgenesis (99).

RSPO1 and WNT4 activate other backup pathways that ensure the ovarian fate, including the insulin signalling pathway, *Runx1* gene (Runt-related transcription factor 1), which maintains Wnt4 expression (100), Fst (Follistatin) and Bmp2 (Bone Morphogenetic Protein 2.

1.3.2. Forkhead transcription factors

FOXL2 (Forkhead Box L2) is part of the forkhead box family of transcription factors and is required to maintain the ovarian phenotype at the postnatal stage. In mice, *Foxl2* is one of the earliest markers of ovarian development, however in *Foxl2*$^{-/-}$ mice the male differentiation pathway is started in the ovaries, where cells acquire the Sertoli cells' characteristics and *Sox9* expression is observed (101). Similarly, lack of Dmrt1 in Sertoli cells resulted in the upregulation of *Foxl2* and transdifferentiation into Granulosa and theca cells (73). These data support the hypothesis that the development and maintenance of the gonad is a biological process regulated by a double-repressive system.

In human, FOXL2 protein is expressed in the mesenchyme of developing eyelids, fetal and adult Granulosa cells of the ovary, in the developing pituitary and adult gonadotroph and thyrotroph cells. *FOXL2* heterozygous mutations are responsible for the blepharophimosis ptosis epicanthus inversus syndrome (BPES, OMIM 110100), a dominant condition characterized by palpebral malformations associated with premature ovarian failure (POF). For the moment few sequence variants have been found in patients with POF (101). Although *FOXL2* does not play a central role in sex differentiation in humans and mice, it seems to regulate the maintenance of gonad fate. Chromatin immunoprecipitation experiments showed that TESCO element, which regulates the expression of *Sox9*, is reactivated after deletion of *Foxl2*. Moreover, *FOXL2* antagonises Wt1 and represses the expression of Sf1 (102).

Other forkhead transcription factor demonstrated to have a role in ovarian function is FOXO3 (Forkhead Box O3). This function was highlighted when *Foxo3a* female KO mice showed early formation of ovarian follicles, followed by oocyte death and infertility (101), whereas constitutive expression of the gene demonstrated to cause delay in follicular development, oocyte growth and finally infertility. Further studies, revealed that Foxo3a reduced the expression of *Bmp15* (Bone Morphogenetic Protein 15) (103). Although few gene variant have been found in POF presenting patients the pathologic role of *FOXO3* remains unknown (101).

In addition to these genes necessary for the determination of the ovary, other genes are required for the ovarian follicle development and are reviewed later (See 1.4.5, page 33).

1.4. SEX DIFFERENTIATION

In the developing gonads Sertoli and Granulosa cells activate the formation of the steroidogenic Leydig (XY gonad) and theca (XX gonad) cells, which produce the necessary hormones for the subsequent correct differentiation of the fetus. The steroid hormones responsible for the phenotype of internal and external genitalia, as well as other reproductive tissues, are called androgens in males and oestrogens in females. At this time two duct precursor systems, the mesonephric (Wolffian) and the paramesonephric (Müllerian) co-exist, as well as two mesodermal swellings that develop to the urethral folds and labioscrotal swellings (8). These precursor structures differentiate into the internal and external genitalia under the influence of hormones secreted by the developing gonad.

Sertoli cells begin the secretion of anti-Müllerian hormone which promotes the regression of the Müllerian structures and initiate further testicular development with

the stimulation of Leydig cells, via a hedgehog signalling pathway. AMH expression has been detected by the 9th week of gestation (104). Leydig cells produce androgens and insulin-like peptide 3 (INSL3) at 8-9 weeks of development. Androgens released by the testes, lead to the formation of male internal (epididymis, vas deferens, seminal vesicle) and external genitalia (penis, scrotum). Testosterone is responsible of developing Wolffian duct into male internal structures while INSL3 stimulates testicular descent. Furthermore, 5α-reductase enzyme transforms testosterone to dihydrotestosterone (DHT), promoting the development of the penis, scrotum and prostate (105). Defects in testosterone and dihydrotestosterone biosynthesis and the androgen receptor are associated with XY DSDs.

1.4.1. Steroidogenesis

Steroid hormone biosynthesis comprises mineralocorticoid, glucocorticoid and sex steroid production from cholesterol in the adrenal glands and sex steroid production in the ovaries and testes (Figure 4). All steroidogenic processes take place in the adrenal cortex, where each zone is responsible for the synthesis of a specific set of steroid hormones, the mineralocorticoids in the outer *zona glomerulosa*, glucocorticoids in *zona fasciculata*, and androgen precursors, androstenedione and dehydroepiandrosterone (DHEA) in *zona reticularis* (106). Steroid production is regulated by external stimuli, such as adrenocorticotropic hormone (ACTH) released in the anterior pituitary. ACTH regulates cortisol synthesis and increases slightly mineralocorticoid and adrenal androgen synthesis via the cAMP (cyclic Adenosine monophosphate) mediated protein kinase A (PKA) pathway which activates StAR protein and SF1 (107), as well as angiotensin II and potassium, which selectively increase mineralocorticoid synthesis via protein kinase C (PKC) pathway. In parallel, the transcription of steroidogenic genes (*CYP11A1, HSD3B2, CYP17A1, CYP21A2,* and *CYP11B1*) and co-factors relevant to glucocorticoid synthesis increases.

Enzymes involved in steroidogenesis can be divided into two major groups: cytochrome P450 (CYP) and hydroxysteroid dehydrogenases/ketosteroid reductase (HSDs/KSR). Cytochrome P450 enzymes are a group of oxidative enzymes that contain a heme group. They are termed depending on their intracellular localization and the mechanism by which they obtain electrons from NADPH (Nicotinamide adenine dinucleotide phosphate hydrogen). Type 1 are targeted in the mitochondria and receive electrons from ferredoxin and ferredoxin reductase, whereas type 2 are targeted to the endoplasmic reticulum and get electrons from P450 oxidoreductase (POR) Each P450 enzyme can metabolize more than one substrate and is involved in broad steps of oxidations in steroidogenesis. Six enzymes take part in steroidogenesis. Mitochondrial P450 cholesterol side-chain cleavage (scc) enzymes (encoded by *CYP11A1*) catalyse 20,22-desmolase activities, isozymes P450c11 (*CYP11B1*) and P450c11β, 11β-hydroxylase activity and P450c11AS (aldosterone synthase) (*CYP11B2*)

18-hydroxylase and 18-methyl oxidase activities. In the endoplasmic reticulum P450c17 (*CYP17A1*) catalyses both 17α-hydroxylase and 17,20-lyase activities, P450c21 (*CYP21A2*) catalyses 21-hydroxylation in the synthesis of glucocorticoids and mineralocorticoids, and P450 aromatase (*CYP19A1*) the aromatization of androgens to oestrogens. The regulation of the steroidogenesis is quantitatively determined by the production of these enzymes, especially P450scc. On the other side, HSD/KSR enzymes catalyse *in vitro* reversible reactions whereas, *in vivo* they continue in an oxidative or reductive mode. They are classified into dehydrogenases, which oxidize hydroxysteroids to ketosteroids using NAD+ (Nicotinamide adenine dinucleotide+) and reductases, which reduce ketosteroids to hydroxysteroids consuming NADPH cofactor. Relevant in steroid synthesis are the 3-HSD type 2 (*HSD3B2*), the 11-HSD type 1 and type 2 (*HSD11B1* and *HSD11B2*), and some of 17-HSDs (108).

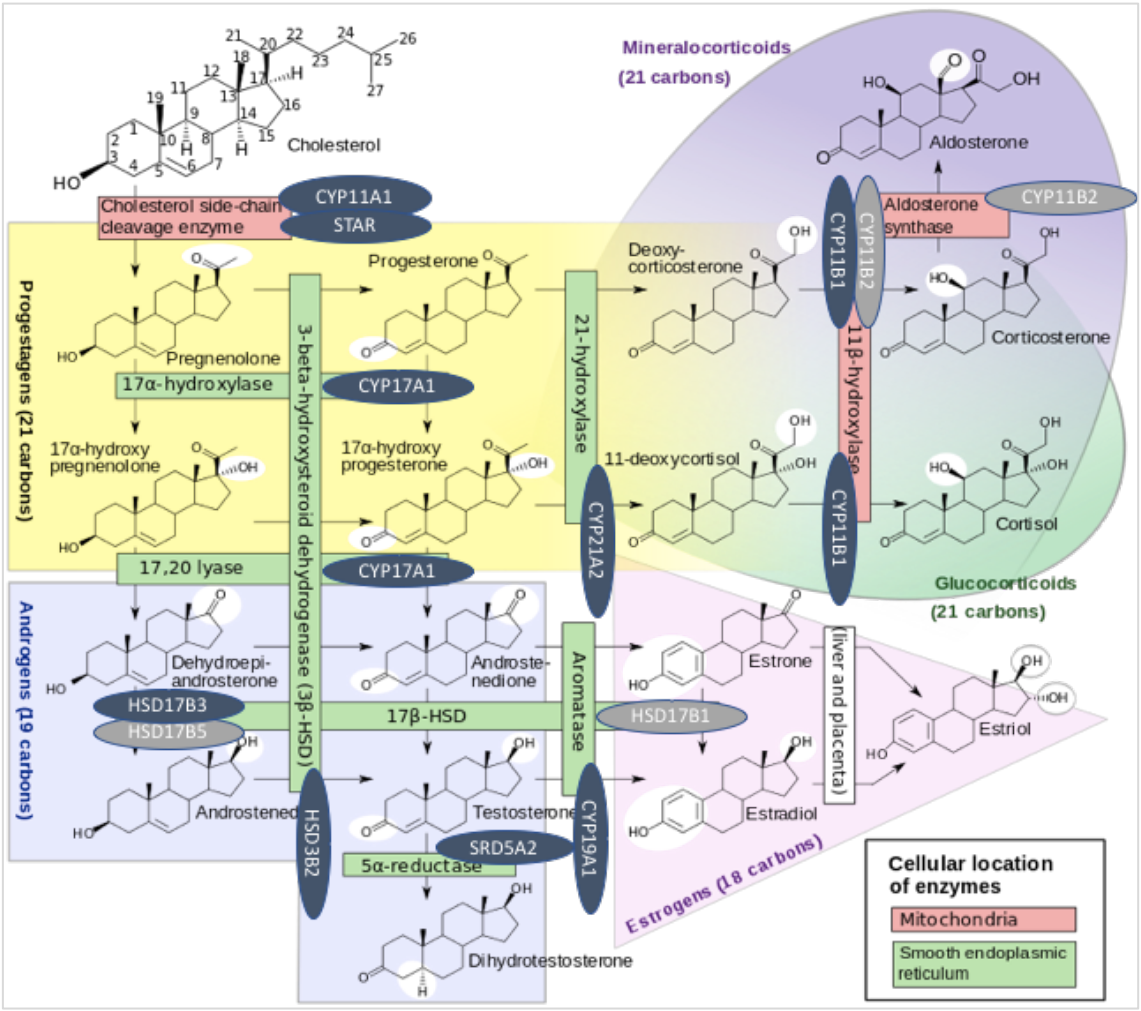

Figure 4. Classic pathway of the adrenal and gonadal steroidogenesis. Arrows indicate chemical reactions and boxes are the enzymes involved in catalysing the respective reaction. Genes encoding the enzymes are adjacent to the boxes. Genes included in the panel are shown in dark grey. Modified from Häggström M., 2014 (109).

Testosterone is required, as well as FSH, for optimal sperm production and sexual function in male gonad. LH (Luteinizing hormone) stimulates the enzymes

involved in testosterone synthesis in Leydig cells, including 17,20-lyase which directs the biosynthesis of steroids toward the sex hormones. Meanwhile androstenedione is converted to testosterone by 17β-hydroxysteroid dehydrogenase (17β-HSD) type 3 or 5 in Leydig cells (110), whereas type 2, found in prostate and placenta, performs the opposite reaction. Aromatase, which is expressed at low levels in Leydig, Sertoli, and germ cells converts testosterone to oestradiol. This step seems to be necessary for the correct initiation of spermatogenesis and mitosis of spermatogonia (111). Dihydrotestosterone (DHT) is the most potent androgen and is formed from testosterone by steroid 5α-reductase in the epididymis and prostate mainly. It has important physiological roles in maintaining sexual function.

In the ovary, steroid hormones are produced for sexual and reproductive functions. Steroidogenesis takes place mainly in theca cells which produce progesterone and androgens that act as precursors for oestrogen synthesis in the Granulosa cells. Once androstenedione and testosterone diffuse into the cells, they are transformed to oestradiol, mainly via the action of aromatase and 17β-HSD types 1 and 7 (110). During follicle maturation, LH and FSH upregulate aromatase and oestrogen synthesis increases. At this point ovulation is triggered due to the positive feedback created by oestrogens and responsible for the LH and FSH secretion. Then, follicle enters the luteal phase and develops to corpus luteum which predominantly synthesizes progesterone. The reduction of LH, aromatase expression and oestrogen production, as well as an increase in CYP11A1 and 3β-HSD activities, stimulate the synthesis of progesterone and the process of follicle rupture begins (112).

Genetic variants in the genes involved in the initial steps of steroidogenesis (Figure 4) shared by the adrenal cortex and the gonads, affect synthesis of steroids in both organs and may lead to DSD and congenital adrenal hyperplasia (Table 1).

Congenital adrenal hyperplasia (CAH) encompass a group of inherited autosomal recessive diseases characterized by impaired adrenal steroid synthesis and in some cases impaired gonadal steroid production with associated genital ambiguity. These inborn errors of steroidogenesis in which cortisol synthesis is impaired and the consequent overproduction of ACTH results in adrenal hyperplasia have been classified as CAH, being 21-hydroxilase (CYP21A2) deficiency the most common. But might also be caused by deficiencies of the following enzymes or cofactors: 17-hydroxylase (CYP17A1), 11-hydroxylase (CYP11B1), 3-HSD type 2 (3βHSD2), P450 oxidoreductase (POR), StAR (lipoid CAH) and P450 side-chain cleavage enzyme (CYP11A1). In the different types of CAH the adrenal steroidogenesis is interrupted at distinct points then, clinical manifestations are closely related to the type and severity of impairment. Genotype-phenotype correlation is good (113).

Table 1. DSD Phenotype due to inborn errors of the synthesis of steroids.

Disorder (OMIM)/Gene	DSD phenotype		CAH	Others
	46,XY	46,XX		
LCAH (201710)/*STAR*	CF: Feminization, GI and infertility; NCF: HH.	CF: absence of pubertal development and POF; NCF: none.	Yes	
P450scc syndrome (118485)/*CYP11A1*	CF: Feminization and GI; NCF: none.	CF: absence of pubertal development and POF. NCF: none.	Yes	
3βHSD2 def (201810)/*HSD3B2*	CF: Undervirilization; NCF: None	CF: Virilization; NCF: adrenarche, PCOS-like	Yes	
17αOH and 17,20 lyase def (202110)/*CYP17A1*	Feminization and GI	Absence of pubertal development, GI.	Rare	Hypertension and hypokalemic alkalosis
21OH def (201910)/*CYP21A2*	CF: precocious puberty; NCF: none.	CF: Virilization; NCF: none.	Yes	Rapid skeletal growth
11βOH def (202010)/*CYP11B1*	Precoious puberty	Virilization	Yes	Hypertension
P450 oxidoreductase deficiency (124015)/*POR*	Feminization and GI	Virilization, PCOS-like	Variable	Antley-Bixler syndrome in infants; Maternal virilization during pregnancy; changes in drug metabolism
17βHSD3 def (264300)/*HSD17B3*	Virilization and gynecomastia at puberty	None	No	
Aromatase def (613546)/*CYP19A1*	None	Virilization in newborns, delayed puberty, HH, multicystyc ovaries, primary amenorrhea at puberty.	No	Maternal virilization during pregnancy
Steroid 5-α reductase def (607306)/*SRD5A2*	Virilization and gynecomastia at puberty	None	No	

CAH, congenital adrenal hyperplasia; CF, classic form; Def, deficiency; DSD, disorder of sex development; GI, gonadal insufficiency; HH, hypergonadotropic hypogonadism; HSD, hydroxisteroid dehydrogenase; LCAH, lipoid congenital adrenal hyperplasia; NCF, non-classic form; OH, hydroxylase; PCOS, Polycystic Ovary Syndrome; POF, primary ovarian failure; SW, salt wasting.

An alternative pathway, called the “backdoor pathway” exists to activate androgen biosynthesis and might play a role in the physiology of CAH (106, 114). This pathway exemplifies a series of enzymatic reactions to transform 17-hydroxyprogesterone (17OHP) to dihydrotestosterone without testosterone as an intermediate. It seems to be active during fetal life in healthy individuals and might take part in male sex development (115).

1.4.1.1. Steroidogenic Acute Regulatory Protein (*STAR*)

Located on 8p11.23, codifies for the steroidogenic acute regulatory protein (StAR), which transports cholesterol from the cytosol to the inner membrane of the mitochondria where it becomes the substrate for the P450 side-chain cleavage/adrenodoxin/adrenodoxin reductase system (CYP11A1-FDX-FDXR) to be converted to pregnenolone (106).

Variants in the *STAR* gene cause lipoic congenital adrenal hyperplasia (LCAH, OMIM 201710), a severe disorder characterized by the impaired synthesis of all adrenal and gonadal steroid hormones due to loss of StAR action. It is believed to result from a two-hit event, a first genetic loss of steroidogenesis and the consequent cellular damage from the accumulation of cholesterol and cholesterol esters. Affected infants usually have adrenal insufficiency and the lack of gonadal steroid hormones leads to 46,XY DSD and infertility. 46,XX females develop secondary sexual characteristics, anovulatory cycles and progressive hypergonadotropic hypogonadism later. A milder form of LCAH, called non-classic LCAH (NC-LCAH) cause a partial loss of StAR activity in patients with primary adrenal insufficiency after infancy, including minimal mineralocorticoid deficiency. 46,XY cases have male external genitalia and mild hypergonadotropic hypogonadism (106). Most disease-causing gene variants are located between exons 5 and 7, encoding for C-terminal and the cholesterol binding site and are present in a homozygote or compound heterozygote state. *In vitro* studies have probed that mutations in the carboxil-terminal drive to diminished or absent function, while proteins lacking the N-terminal domain, including a mitochondrial target sequence, are able to stimulate steroidogenesis (116).

1.4.1.2. Cytochrome P450 Family 11 Subfamily A Member 1 (*CYP11A1*)

P450 cholesterol side-chain cleavage (P450scc) is the only enzyme that catalyzes the conversion of cholesterol to pregnenolone. It is expressed in the steroidogenic cell types of the adrenal, gonad and placenta, as well as in other tissues (106) and its transcription is regulated in a tissue-specific manner and hormonally, such as the action of SF1 in gonads and adrenal (117).

The cytochrome P450scc is encoded by the *CYP11A1* gene, which lies on chromosome 15q23-q24. It was thought that loss-of-function mutations in *CYP11A1*, required for the synthesis of progesterone would be incompatible with human gestation, however, it caused impaired production of gonadal and adrenal steroids in utero and after

birth, driving to a phenotype similar to LCAH with adrenal insufficiency and 46,XY sex reversal (118). *In vitro* analyses have shown that most mutants have enzymatic residual activity, although frameshift gene sequences have no activity (119). Clinical and biochemical findings in patients with LCAH due to *STAR* mutations or *CYP11A1* are similar.

1.4.1.3. Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta- and Steroid Delta-Isomerase 2 (*HSD3B2*)

The *HSD3B2* gene in 1p12 is highly expressed in the adrenals and gonads. Deficiency in 3β-hydroxysteroid dehydrogenase type 2 (3βHSD2) causes reduced synthesis of progesterone and 17-hydroxypreogesterone, precursors of aldosterone and cortisol, respectively, as well as androstenedione, testosterone and oestrogens. In contrast, renin, ACTH and dehydroepiandrosterone (DHEA) are increased. The later DHEA is converted to testosterone by 3βHSD1 (3β-hydroxysteroid dehydrogenase type 1), which is expressed in placenta and peripheral tissues (120).

Patients with *HSD3B2* gene variants are characterized by impaired steroid synthesis in the gonads and adrenal glands, and variable clinical presentation depending on the alteration. Generally, severe forms result in salt-wasting phenotype caused by frameshift mutations, in-frame deletions, and nonsense mutations. On the contrary, missense variants retain a residual enzymatic activity and lead to a non-salt-wasting phenotype (121).

Most patients present salt-wasting adrenal crisis in both sexes during childhood. High levels of androstenedione guide to relatively increased testosterone in females, but not in males. In 46,XY neonates, testosterone deficiency causes genital ambiguity, whereas in 46,XX the relative high testosterone levels cause virilization with clitoromegaly and partial labioscrotal fusion. After childhood, females present with precocious puberty, acne, hirsutism and menstrual irregularities. Non-classic 3βHSD2 deficiency is extremely rare (121).

1.4.1.4. Cytochrome P450 Family 17 Subfamily A Member 1 (*CYP17A1*)

The *CYP17A1* gene codifies for an enzyme with both 17α-hydroxylase and 17,20-lyase activities, principally in the adrenal and gonads. Hence, sequence mutation in the gene, located on 10q24.32, impair adrenal and gonadal sex steroid production (Figure 4), originating sexual infantilism and puberty failure (122).

Deficient 17β-hydroxylase activity results in decreased cortisol synthesis, and elevated concentrations of 11-deoxycorticosterone (DOC) and corticosterone mediated by ACTH. The overproduction of DOC in the *zona fasciculata* causes sodium retention, hypertension and hypokalemia which suppress plasma-renin activity and aldosterone secretion in a variable mode and the presence of corticosterone prevents the adrenal crisis due to the glucocorticoid activity. As a result of the gonadal sex steroid synthesis deficiency,

due to the obstruction of the dehydroepiandrosterone (DHEA), female patients are phenotypically normal with lack of pubertal development and adrenarche whereas 46,XY have absent or incomplete development of external genitalia. Both 46,XX and 46,XY patients present with hypergonadotropic hypogonadism, and low-renin hypertension. Among the identified alterations, 4 are recurrent (106) and include all type of changes in which all the enzymatic activity is impaired or reduced to 80%.

Selective 17,20-lyase-deficiency has been reported and is caused by mutations in cytochrome b5, the cofactor needed to exert 17,20-lyase activity (120).

1.4.1.5. Cytochrome P450 Family 21 Subfamily A Member 2 (*CYP21A2*)

The *CYP21A2* gene encoding for steroid 21-hydroxylase (21OH) lies in the middle of the HLA (Human leukocyte antigen) *locus* in 6p21, next to the nonfunctional pseudogene *CYP21A1P*. There is high degree of sequence similarity between these two genes, thus meiotic recombination is frequent and about 95% of cases with 21OH deficiency develop when *CYP21A1P* replaces the corresponding *locus* of *CYP21A2* gene reducing the expression of the protein and its function (123).

21-hydroxylase deficiency results in reduced glucocorticoid and mineralocorticoid formation and elevated 17-hydroxyprogesterone (17OHP), as well as other precursors. The lack of negative feedback of cortisol on ACTH (adrenocorticotropic hormone) production amplifies the adrenal androgen production. The severity of the CAH depends on the residual CYP21A2 function. Classic form of 21OH deficiency is divided in salt wasting and simple virilising depending on the severity of aldosterone insufficiency. Mutations that result in the salt wasting phenotype are completely inactivating and lead to an adrenal crisis if neonatal screening is not performed. Patients with simple virializing form retain 1-2% of the enzymatic activity and the neonatal crisis is prevented because minimal aldosterone is produced. In both forms, androgen excess results in virilization of external genitalia and rapid skeletal growth in 46,XX cases, and in males, simple virilizing form drives to precocious puberty as well as rapid skeletal growth. In the non-classic form, about 50% of the enzyme activity is maintained and do not show adrenal insufficiency. Patients might have partial glucocorticoid deficiency and present with milder androgen excess or asymptomatic. Females have normal genitalia (124).

In patients with *CYP21A2* sequence variants, the “backdoor pathway” contributes to the virilization of female external genitalia through a 17OHP excess. The pathway is reduced after the first year of life, which explains the normal growth of the patients during the first year (125).

1.4.1.6. Cytochrome P450 Family 11 Subfamily B Member 1 (*CYP11B1*)

CYP11B1, located on 8q24.3, encodes the steroid 11β-hydroxylase, which converts the 11-deoxycortisol to cortisol and deoxycorticosterone to corticosterone in the adrenal *zona fasciculata* under the regulation of ACTH. Mutations in *CYP11B1* results in reduced corticosterone and cortisol synthesis, with successive ACTH and androgens increase. Elevated levels of deoxycorticosterone after neonatal period inhibit the renin-angiotensin system and drive to extracellular fluid volume expansion, hypertension, diminish plasma-renin activity and low aldosterone concentrations. The hypertension, rather than salt loss is the clinical difference between patients with 11β-hydroxylase and 21-hydroxylase deficiency. Affected females are virilized due to fetal ACTH excess. Although mutations in *CYP11B1* usually result in classic CAH phenotype because an absent or minimal enzyme activity, sequence variants causing the non-classic form have also been described (106).

1.4.1.7. Cofactor defects: Cytochrome P450 Oxidoreductase (*POR*) and cytochrome b5 (*CYB5*)

Cytochrome P450 Oxidoreductase (POR) plays a key role in electron transport in the endoplasmic reticulum, and several enzymes including 17α-hydroxylase, 21-hydroxylase, and aromatase depend on POR as a cofactor. Although *POR*-null mice die in utero (106), some enzymatic function is retained in humans and patients present CAH with DSD in both sexes. Deficiency is characterized by partial fail of P450c17 activity, with or without associated deficient activity of P450c21 and P450 aromatase. The phenotypes vary from women with amenorrhea and polycystic ovaries, virilized 46,XX newborn or androgen deficiency in 46,XY newborn to important hormone disturbances that lead to abnormal genitalia in both sexes. Additionally, the “backdoor pathway” of androgen synthesis contributes to the prenatal virilization of affected females (126). Mineralocorticoid deficiency is not observed in these patients because 17α-hydroxylase increases production of mineralocorticoid intermediates, thus affected adults might develop hypertension (24, 127). POR acts as an electron donor of other not steroidogenic enzymes and skeletal dysplasia and changes in drug metabolism is often observed in patients (128). Moreover, newborns with POR deficiency sometimes have a skeletal defect called Antley–Bixler syndrome (ABS, OMIM 201750), which is characterized by craniosynostosis, brachycephaly, radio-ulnar or radio-humeral synostosis, bowed femora, arachnodactyl, midface hypoplasia, proptosis and choanal stenosis (129).

Regarding cytochrome b5 (CYB5), it facilitates the allosteric interaction between POR and CYP17A1. A missense and a premature stop coding *CYB5A* (Cytochrome B5 Type A) gene variants have been found in 46,XY DSD cases, all with low androgens and gonadal defect but normal mineralocorticoid and glucocorticoid production (120).

1.4.1.8. Hydroxysteroid 17-beta dehydrogenase 3 (*HSD17B3*)

HSD17B3 gene, localized in 9q22.32, encodes the 17β-HSD3 (17β-Hydroxysteroid dehydrogenase) enzyme. It is mainly expressed in male gonad and reduces DHEA and androstenedione to serve as precursors of testosterone (130, 131). In contrast, estrogens are poor substrates for the human 17βHSD3 enzyme. The enzymatic deficiency is caused by homozygous or compound heterozygous variants in *HSD17B3* and is characterized by reduced sex hormones production and normal adrenal secretion of glucocorticoids and mineralocorticoids.

Only 46,XY infants with a 17βHSD3 deficiency manifest a DSD. The scarce amount of produced testosterone during fetal development is suggested to be sufficient to trigger male internal genitalia, or alternatively other 17βHSD isoenzymes are able to synthesize testosterone from androstenediones, but not enough to compensate the deficiency. 46,XY children with 17βHSD3 deficiency present completely undervirilized with a blind vaginal pouch and undescended testes or varying degrees of hypospadias and micropenis with intra-abdominal or inguinal testes. Wolffian derivatives including the epididymides, vas deferens, seminal vesicles, and ejaculatory ducts are present (131).They are mostly assigned as females when born. As in patients with *SRD5A2* gene variants, these children begin to virilize at puberty (131). 46,XX with mutations in *HSD17B3* show a female phenotype and produce normal levels of androgens and estrogens, which suggests that testosterone is produced in the ovary even in the absence of 17βHSD3 enzyme expression (106).

The clinical presentation is not different from that of partial androgen insensitivity syndrome (PAIS) and 5α-reductase type 2 deficiency, thus molecular testing supports the diagnosis. Biochemical diagnosis is based on a low testosterone/androstenedione ratio (T/A<0.8).

1.4.1.9. Cytochrome P450 Family 19 Subfamily A Member 1 (*CYP19A1*)

The P450aro or aromatase is encoded by *CYP19A1* gene on chromosome 15q21.1. This enzyme catalyzes a complex series of reactions in three precursors androstenedione, testosterone and 16α-hydroxyandrostenedione into estrone, estradiol and estriol, respectively and is expressed in steroidogenic tissues, like ovaries, testes, placenta, brain and non-steroidogenic tissues, such as bone and fat. The transcription of the gene is tissue-specific and is regulated through different promoters and transcriptional start sites (132).

Identified mutations in the *CYP19A1* include missense, nonsense, deletions and insertions, splice site and also large intragenic deletions and are mainly found in exons 9 and 10. Clinical features of patients vary depending on gender, age and enzymatic activity. Aromatase deficiency drives to androgen excess in the mother and the fetus, causing maternal virilization and masculinization in the external genitalia of the 46,XX fetus,

whereas no changes are observed in males at birth In adolescent girls, P450 aromatase deficiency may lead to delayed puberty, hypergonadotropic hypogonadism, multicystic ovaries and primary amenorrhea due to lack of estrogens. Signs of virilization may also be present. In both sexes, the subsequent estrogen deficiency causes delayed epiphyseal closure, eunuchoid habitus and osteoporosis (132).

1.4.1.10. Steroid 5 Alpha-Reductase 2 (*SRD5A2*)

Two isoenzymes convert testosterone to dihydrotestosterone (DHT), which has more affinity for the androgen receptor. During embryogenesis, DHT acts to form the male urethra and prostate at the urogenital sinus and induces swelling and folding at the genital tubercle to create the penis and scrotum (133). The 5α-reductase type I, encoded by *SRD5A1* gene on chromosome 5p15 is expressed in peripheral tissues, such as the skin and hair follicles, whereas 5α-reductase type II enzyme is encoded by *SRD5A2* on chromosome 2p23 and is predominantly found in stromal cells of internal and external reproductive organs and prostate.

Up to date, human mutations in *SRD5A1* have not been associated to DSD. But affected 46,XY with inactivating sequence variants in *SRD5A2* lead to steroid 5α-reductase 2 deficiency, an autosomal recessive form of 46,XY DSD, which results in a wide spectrum of phenotypes. Newborns present with ambiguous external genitalia, micropenis, scrotal hypospadias, cryptorchidism or inguinal testis and prostatic hypoplasia. They are generally raised as females. At puberty, these patients are virilized and present gynecomastia due to 5α-reductase type I activity. Most patients with *SRD5A2* deficiency are homozygous or compound-heterozygous (35%) for inactivating mutations. Heterozygous mutations have been also reported, proposing the dominant effect of certain gene variants (133).

Both genes have a complex developmental regulation of expression.. 5α-reductase type I is expressed again in nongenital skin and liver after puberty, suggesting that it may be responsible of the pubertal virilization observed in the patients with 5α-reductase deficiency. Biochemical diagnosis is suggested when high testosterone/DHT ratio is observed although both isoenzymes reduce a variety of steroids in their degradative pathways and can be seen in a whole steroid profile (120), (106).

1.4.2. Insulin Like 3 (*INSL3*) and Relaxin-family peptide receptor 2 (*RXFP2*)

In humans and other mammals, testis descent follows two distinct steps: the intra-abdominal and the inguino-scrotal stage. *INSL3* (Insulin like factor 3), expressed in Leydig cells of the foetal and adult testis, was proposed to regulate the intra-abdominal step. *Insl3*-null mice developed bilateral cryptorchidism with testes located in the abdominal cavity. Normal testicular descent was achieved when *INSL3* was overexpressed in pancreatic β-cells in these KO mice (134). Subsequently, it was shown that Insl3 acts via its G-protein-coupled

receptor RXFP2 (Relaxin-family peptide receptor 2). *RXFP2* gene, located on chromosome 13, encodes a protein with an extracellular ligand-binding domain, seven transmembrane domains and one intracellular domain. In mice, gene changes cause cryptorchidism, with bilateral intra-abdominal testes (135).

In humans, several point alterations have been identified within the coding region and the promoter of *INSL3* in DSD patients presenting with cryptorchidism (136). In contrast, studies investigating *RXFP2* were scarce and have focused on the single found variant Thr222Pro in exon 8, which reduces *RXFP2* expression on the cell surface membrane on patients with undescended testes (137). Recently, another missense variant was detected in 4 members of the same family presenting with cryptorchidism. Functional analysis showed that the mutant protein failed to bind INSL3 or respond to the ligand with cAMP signalling (135).

On the other side, the regulation of *INSL3* is poorly known. Oestradiol and diethylstilboestrol may downregulate the expression of *INSL3* in foetal Leydig cells which explains the lack of testicular descent in male embryos after exposure to oestrogen excess (138). It has been hypothesized that SF1 binds to one of the three binding sites in the gene and acts as a transcriptional activator of *INSL3* promoter. However, *in vivo* studies demonstrated that the regulation of *INSL3* expression and downregulation by oestrogens requires additional transcription factors (139).

1.4.3. Androgen receptor (*AR*)

Androgens action, testosterone and dihydrotestosterone (DHT), via AR is essential for male sexual development before birth and for secondary male sexual development at puberty, while in females androgens take part in sexual development at puberty and in adulthood (140). The *AR* gene, located on Xq11-12 encodes a transcription factor with three major domains: amino-terminal domain (NTD), a DNA-binding domain (DBD) containing two zinc fingers and a hinge region that connects to the ligand-binding domain (LBD) on the carboxil-terminal. More than 500 variants have been described varying from coding sequence (CDS) changes, splicing, 5'UTR (untranslated region) and partial or complete gene deletions (141).

Androgen insensitivity syndrome (AIS, OMIM 300068) is a congenital disorder that manifests as a result of cellular resistance to testosterone and dihydrotestosterone. Mutations in the *AR* gene are associated with a wide spectrum that result in the formation of a feminine phenotype in a 46,XY individual and ranges from complete androgen insensitivity syndrome (CAIS) to partial (PAIS) or mild (MAIS). Sequence variants have been identified also in isolated hypospadias (140).

AR is also expressed in the developing Granulosa cells and as shown in $Ar^{-/-}$ transgenic female mice presenting a phenotype similar to that of primary ovarian insufficiency (POI), it's required for folliculogenesis. In humans, the CAG repetition in exon 1 of the gene has been proposed to cause POI (142).

1.4.4. Luteinizing Hormone/Choriogonadotropin Receptor (*LHCGR*) in males

The production of androgens is under the control of human chorionic gonadotropin (hCG) and luteinizing hormone (LH) during fetal and postnatal life, respectivelly, that act by stimulating the transmembrane receptor LHCGR. In differentiating Leydig cells the secretion of androgens seems to be autonomous, later hCG produced by the placenta highly stimulates their differentiation and the androgen production, until second and third trimester of pregnancy when the levels decrease. At pubertal stage, LH through LHCGR drives to the secretion of androgens that give rise to pubertal changes (143).

Defects in the LH receptor functionality cause impaired regulation of Leydig cell function and androgen production. LHCGR receptor is a member of the superfamily of G protein-coupled receptors, characterized by the presence of an extracellular hormone-binding domain which consists of a leucine-rich repeat domain flanked by two cysteine clusters, and a last exon of the gene codifying for a seven transmembrane helix domain (Figure 5) involved in the G protein activation. After ligand binding to the receptor, G proteins are activated in the intracellular membrane and an increase in cAMP levels activate of cAMP-dependent protein kinase. The variants on the different domains in the *LHCGR* lead to distinct effects on LH receptor function. Then, missense alterations may alter the function of the protein inhibiting the signal transduction, while large deletions and nonsense changes cause a complete defective protein and lack of hormone binding. Both forms of inactivating mutants diminish the efficiency of the LH receptor transduction signal and are supposed to cause the undervirilization of the 46,XY male, known as Leydig cell hypoplasia (LCH, OMIM 238320) (143).

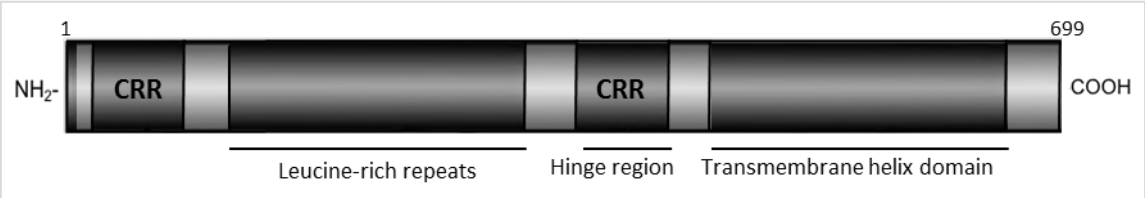

Figure 5. Scheme of the structure of the LHCGR protein. LHCGR comprises 4 main structures: A Cysteine-rich region (CRR) in the N-terminal, nine Leucine-rich repeats, a second CRR or hinge region and the 7 transmembrane domains next to the carboxi terminal. Modified from Xu Y., *et al* (144).

The complete form of Leydig cell hypoplasia or type 1 LCH, is described as 46,XY female external genitalia, lack of pubertal development, small undescended testes with seminiferous tubes, absence of mature-type Leydig cells and presence of rudimentary epidydimis and vas deferens. Biochemical analyses indicate elevated gonadotropins, mainly

LH, with low testosterone and no response to hCG stimulation test. In contrast, the partial form or Leydig cell hypoplasia type 2, encompasses a wide spectrum with predominantly male external genitalia, micropenis with or without hypospadias and cryptorchidic or scrotal testes. At puberty, these patients are partially virilized and testes acquire a normal size while penis length is impaired. At this time, LH levels increase and testosterone ranges from infant to adult male levels (143). Homozygous or compound heterozygous inactivating variants are located along the gene and all type of changes have been identified, such as missense, nonsense and complete gene deletion. The mechanisms of inactivation may cause a misfolding of the receptor and a consequent reduced transport to the cell surface or decrease the ability to activate the signal transduction. The severity of the phenotype is well correlated with the residual activity of the receptor. For example a compound heterozygous patient presenting a small deletion in the intracellular region together with a missense variant in the transmembrane domain led to a more severe perineoscrotal hypospadias than two missense changes in the last exon causing micropenis (145). On the other side, female siblings of LCH cases present with primary amenorrhea, oligoamenorrhea, anovulation or infertility (146).

In contrast, activating heterozygous variants have been described in male-limited precocious puberty or testotoxicosis (OMIM 176410) which is characterized by early puberty in boys, usually before 4 years of age, penis growth, pubic hair and other puberty signs. Gonadotropins are decreased indicating a peripheral cause of precocious puberty, while testosterone levels are high, in contrast to Leydig cell hypoplasia. *In vitro* expression analysis of the first identified activating gene variant, showed an increase of 10 to 20-fold in the basal activity of the LH receptor, amplifying the cAMP signal. Although the sixth helix and the third intracellular loop of the transmembrane domain seemed to be the hot spot of the activating mutations, further genetic analysis identified them in other regions (147). Indeed, molecular modelling studies revealed that changes in the interaction of conserved polar amino acids drive to an alternative conformation of the protein that allows the binding of the G protein to other amino acids side chains (148).

Defects in *AMH* gene, also involved in male differentiation, have been described before (See 1.2.9, page 15). Other genes that take park in male differentiation, such as oestrogen recpetors and FSHR are explained later (See 1.4.5.4, page 36 and 1.4.5.1, page 35).

1.4.5. Formation of the ovarian follicle and differentiation of the female genitalia

As mentioned before (See 1.3, page 18) other genes are required for the development of the ovarian follicle and differentiation of the female genitalia. The formation of the ovary is slower than testicular process, and is identifiable at the 11th week of gestation. In mice, the first sign of ovarian differentiation is at embryo day 13.5. At this time, germ cells enter meiosis and then arrest at prophase I to begin the development of

primordial follicles at time of birth (149). Primordial follicles comprise a single oocyte enclosed with a layer of somatic Granulosa cells, separated from the ovarian interstitium by a basal lamina. As follicles grow to secondary and antral follicle, the Granulosa cells change to a cuboidal shape and proliferate into multilayers. Steroidogenic theca cells differentiate at the secondary follicle phase and surround the Granulosa cell layers outside the basal lamina (Figure 6).

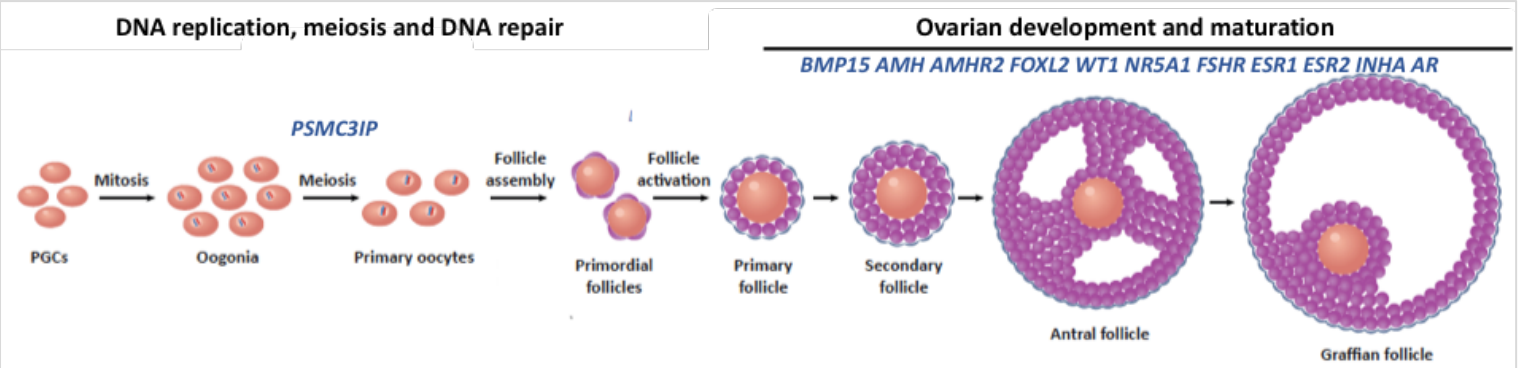

Figure 6. Schematic representation of oogenesis and ovarian folliculogenesis. Only genes sequenced in the targeted gene panel have been included in the image. Modified from Jiao 2018 (150).

Once formed, the Granulosa cells begin the expression of the Hedgehog signalling pathway which starts the differentiation of theca cells. For that, DHH (Desert Hedgehog) binds HHIP (hedgehog interacting protein 1) and membrane receptors PTCH1 and PTCH2 (Patched 1 and Patched 2) on the theca cells. GLI (glioma-associated oncogene) zinc finger transcription factor mediates the activation of the hedgehog signalling cascade (151). Progenitors of the theca cells derive from two cell populations. The transcriptional profiles of the cells coming from the mesonephros show high expression levels of genes involved in steroidogenesis, such as *Star, Cyp11a1 and Cyp17a1, Hsd3b2* and *Lhcgr*. In contrast, genital ridge derived cells revealed higher expression of *Esr1* (oestrogen receptor 1) and *Wt1*, amongst others (152). Theca cells produce androgens that are converted to oestrogens by Granulosa cells and support vascularization of the follicle. Secreted oestrogens allow the development of the Müllerian duct into uterus, Fallopian tubes and the upper part of the vagina. In the presence of androgens, the XX fetus lead to normal progress of fused urethral folds and labioscrotal swelling that develop to labia minora and majora respectively, the genital tubercle forms the clitoris and the vaginal plate creates the lower part of the vagina. Signs of this growth are present at 20 weeks of gestation.

Sequence variations in the genes involved in hormonal signalling and folliculogenesis end in ovarian disorders, such as ovarian dysgenesis, premature ovarian failure (POF), polycystic ovary syndrome (PCOS) or cancer. Hereafter, I describe several genes (*LHCGR, FSHR, BMP15, GDF9, INHA, ESR1, ESR2, HSD17B4* and *PSMC3IP*) in which genetic changes have been associated to ovarian disorders.

1.4.5.1. Gonadotropin receptors in females

As I have described before (See 1.4.4, page 32), FSHR (Follicle Stimulating Hormone Receptor) and LHCGR are glycoprotein hormone receptors that belong to the membrane-bound G-protein coupled receptors (GPCRs) family. These receptors, as well as their binding hormones FSH and LH, are involved in regulating hormonal signalling in males and females. In females, both receptors are expressed in the Granulosa cells when follicles progress from the primary to the secondary stage, together with oestrogen and androgen receptors. FSHR and LHCGR produce cAMP as the main second messenger, which in turn, stimulates ovarian follicular steroidogenesis, favouring puberty and menstrual cycles (153). Patients with pathogenic variants in these receptor genes form ovarian follicles up to the preantral stage, but further maturation is blocked.

The *FSHR* gene is located on 2p16.3 and encodes a 695 amino acid protein characterized by an extracellular domain, seven transmembrane segments and an intracellular tail. In female mice, the disruption of *Fshr* leads to infertility while males show small testes but remain fertile, supporting that Fshr plays a major role in the development of the female gonad. Inactivating mutations in the extracellular domain lead to a complete FSH resistance with hypergonadotrophic hypogonadism and 46,XX ovarian dysgenesis, such as the p.Ala189Val missense change reported in Finnish population (153). This substitution inhibits the FSHR cell-surface trafficking and therefore mutant FSH is retained inside the cell. Other mutations throughout the receptor have been noted in women with primary ovarian insufficiency (POI), secondary amenorrhea and high FSH levels.

LHCGR inactivating variants also cause primary ovarian insufficiency in 46,XX. Patients present with primary amenorrhea or oligoamenorrhea and infertility. The hypergonadotropic hypogonadism is characterized by high LH biochemical levels (143).

1.4.5.2. Bone morphogenetic protein 15 (BMP15) and Growth differentiation factor 9 (*GDF9*)

Bone morphogenetic protein 15 (BMP15) and growth differentiation factor 9 (GDF9) are members of the transforming growth factor-beta (TGFβ) superfamily that act synergistically to regulate folliculogenesis and ovulation. BMP15 and GDF9 have similar protein structure, expression patterns and interact with each other to regulate Granulosa cells. GDF9 and BMP15 bind to their corresponding transmembrane serine-threonine kinase receptors and finally drive to the regulation of target genes related to ovulation and luteinisation (154), such as KITL (Ligand for the receptor-type protein-tyrosine kinase) important in the growth of oocytes. Moreover, BMP15 reduces progesterone production by downregulating the expression of *STAR*, while GDF9 interacts with pituitary gonadotropins to diminish progesterone and thus, inhibit follicular luteinisation (154).

BMP15 is an X-linked gene that encodes a growth/differentiation factor secreted from oocytes. It is involved in either follicular development, regulation of several processes in Granulosa cells, prevention of apoptosis and regulation of FSH sensitivity. Studies in different animal models confirmed the role in folliculogenesis and ovulation although Bmp15 seems to be more critical in mono-ovulating species, like human or sheep rather than in poly-ovulating ones, as mice (155).. In women, variants in *BMP15* have been associated to primary and secondary amenorrhea due to primary ovarian insufficiency (POI) and polycystic ovary syndrome (PCOS) (155). Recently a *BMP15* gene duplication was reported in a patient with 46,XX/45,X0 karyotype and spontaneous menarche, which suggested a gene dosage contribution that enable an amount of follicles to reach puberty (156).

GDF9 gene is located on 5q31.1 and is linked to the follicular development between the primary to secondary follicle stage. Several studies in different cohorts have related missense variants with POI and PCOS (157)..

1.4.5.3. Inhibin Subunit Alpha (*INHA*)

Inhibin is a heterodimeric glycoprotein structurally related to the transforming growth factor-beta (TGFβ) superfamily. It is formed of inhibin A homodimers, codified by *INHA,* and inhibin B or activin codified by *INHBA* (Inhibin Subunit Beta A) and *INHBB* (Inhibin Subunit Beta B) genes. The principal function of inhibin in women is the regulation of pituitary FSH secretion and is an ovarian reserve marker. *In vivo*, inhibin functions as a tumour suppressor, as Inha-null transgenic mice developed stromal/Granulosa cell tumours with increased FSH levels and infertility(158).

INHA gene maps to 2q35 and the first supporting evidence of the relation between INHA and primary ovarian insufficiency in humans was revealed when a translocation in the gene *locus* was found (158). However, the recurrent p.Ala257Thr substitution in different populations has been also reported in controls (159).

1.4.5.4. Oestrogen receptor α and β

Oestrogens control development, cell differentiation and maintenance of homeostasis in adults. All the estrogenic functions are done by binding and activating their nuclear receptors, Oestrogen receptor α (ESR1) and β (ESR2). Mouse modelling studies suggested that ESR2 is a dominant negative regulator of ESR1-dependent transcription as a way to maintain the balance between the two receptors. *Esr1*-null male mice showed infertility, hyperplasia, adipocyte hypertrophy, insulin resistance and glucose intolerance while female animals had hypoplasic uterus and multicystic ovaries. *Esr2* deletions in animal models demonstrated different phenotypes, such as infertility without sex reversal in females (160) or normal phenotype but smaller litters, whereas male KO rodents were fertile and lacked reproductive tract abnormalities. This last model suggested that Esr2 is only

essential for ovulation but not for sexual differentiation. In humans, *ESR2* is more expressed in the Granulosa cell of the ovary in the developing follicle than *ESR1*, and also in the fetal ovary during gonadal development (161). However the identification of an *ESR2* mutation in a 46,XY DSD patient presenting with gonadal dysgenesis demonstrated a role in early testicular development (160). However, the potential ESR2 functions in testes and prostate need to be probed.

In humans *ESR1* gene alterations cause oestrogen resistance, high oestrogen levels in male and females and delayed puberty, infantile uterus, primary amenorrhea and ovarian cyst in women (162). Two heterozygous variants in *ESR2* have been reported in two 46,XX young woman. First was diagnosed with primary amenorrhea and second had streak gonads, infantile uterus, lack of puberty and breast development (161). Molecular modelling in the last case demonstrated that the mutant disrupted estradiol-dependent signalling and failed to interact with its coactivator.

Other previously mentioned genes that take part in the follicle formation and maturation are: *AR, NR5A1, WT1, NR0B1, MAMLD1, AMH* and *AMHR2.*

1.4.5.5. Other genes related to ovarian dysgenesis as part of a phenotypic spectrum

Perrault syndrome is an autosomal recessive condition characterized by ovarian failure in females and deafness in both males and females. Some women also present neurological features (Perrault syndrome type II), including developmental delay or intellectual disability, cerebellar ataxia, motor and sensory peripheral neuropathy (163).

Mutations in *HSD17B4* (Hydroxysteroid 17-Beta Dehydrogenase 4*)*, which encodes a 17-beta-estradiol dehydrogenase involved in peroxisomal fatty acid beta-oxidation and in *HARS2* (Histidyl-TRNA Synthetase 2, Mitochondrial), encoding the mitochondrial histidyl-tRNA synthetase, have been detected in patients with Perrault syndrome (OMIM 233400) (163). Regarding the *HSD17B4* gene, two compound heterozygous variants were found in two sisters with clinical features of DBP (D-bifunctional protein, OMIM 261515) deficiency, together with ovarian dysgenesis. In contrast, linkage-analysis and gene sequencing in five affected family members identified two heterozygous variants in *HARS2*. HARS2 is required for protein translation within the mitochondria and reduced activity was found in both mutants. However, the lack of *HSD17B4* and *HARS2* sequence changes in other cohorts evidenced the heterogenic cause of this disease (164). Further whole exome sequencing and linkage-analysis discovered other underlying genes causing the disease: *C10orf2* (chromosome 10 open reading frame 2), *LARS2* (leucyl-tRNA synthetase 2, mitochondrial) and *CLPP* (caseinolytic mitochondrial matrix peptidase proteolytic subunit), essential for normal mitochondrial function (163).

1.4.6. Genes affecting DNA replication, meiosis and DNA repair in the female gonad formation

Gene sequence variants implicated in creation and repair of DNA double-strand breaks for recombination, damage control, cell cycle progression or development of the synaptonemal complex have been also associated with defects of the female gonad formation.

PSMC3IP (Proteasome 26S Subunit, ATPase, 3-Interacting Protein) codifies a highly expressed nuclear protein in foetal and adult gonads, spleen and thymus in both human and murine models. It has a leucine zipper domain critical for nuclear receptor binding, a coiled-coil domain in the middle of the protein and a C-terminal acidic region that acts as a ligand-dependent transcriptional coactivation domain (165). As indicated by its yeast ortholog Hop2, it coactivates DMC1 (DNA Meiotic Recombinase 1) and RAD51 (DNA repair protein RAD51 homolog 1) proteins, fundamentals for homologous pairing and recombination in meiosis (166) and either mutations or deletions in Hop2 stop entry into meiosis I. In male mouse meiosis was blocked but since Sertoli, Leydig cells and spermatogonia were present, pubertal development was normal (165). However, in the female *Psmc3ip* KO mouse model, ovarian size was reduced and germ cells were lacking. The importance of the *PSMC3IP* gene in the germ cell development of both sexes was then revealed.

PSMC3IP binds DNA-binding domain of glucocorticoid receptor, oestrogen receptor 1 and 2, androgen receptor, progesterone receptor and thyroid hormone receptor beta, acting as a coactivator of hormone-dependent transcription (165).

In humans the first mutation was identified in 5 members of a consanguineous family affected by 46,XX gonadal dysgenesis. It was a 3-bp C-terminal deletion that disturbed the coactivation of ESR1, which affected the follicular environment during formation in the fetal ovary (167). More recently, a homozygous stop mutation has been reported in a consanguineous Yemeni family in 4 sisters presenting with POI and a brother with azoospermia. The carboxyl-terminal variation inhibited the association with single strand DNA and the proteins required for homologous recombination (166).

1.5. OPPOSING INTERACTIONS TO CORRECTLY DEVELOP AND MAINTAIN THE GONADS

A complex interaction is needed between the female and male pathways to correctly trigger the initiation, development and maintenance of gonadal differentiation. Data suggest that pathways act antagonistically, suppressing the alternate fate through a sequence of repressive feedback loops during sex differentiation, but also in adulthood.

During embryonic development, the most important antagonism occurs between the SRY/SOX9/FGF9 and WNT4/RSPO1/β-catenin signalling pathways (Figure 7) (48). *Fgf9*-null or *Fgfr2*-null transgenic mice resulted in the increased expression of *Wnt4* and male-to-female sex reversal, whereas the deletion of *WNt4, Rspo1* or B-catenin showed female-to-male sex reversal (168). In the male, *SRY* suppresses the activation of WNT transcription factor β-catenin. Additionally, FGF9 supports the transcriptional activity of *SOX9* and together, downregulate female-specific gene *WNT4* leading towards testicular development (50). CBX2 might also participate in this feedback loop upregulating *SOX9* and *SOX3* in the testicular Sertoli-like cells and inhibiting *FZD1* (Frizzled Class Receptor 1) and *FOXL2* involved in ovarian development (169). By contrast, during sex determination in the ovary, β-catenin previously activated by WNT4 and RSPO1, limits the expression of *Sox9* and therefore, suppresses the SOX9/FGF9 feedback loop (95). WNT4 signalling increases the transcription of *FST* (Follistatin) too, which antagonises Activin to inhibit the formation of the testes-specific vasculature (170). *MAP3K1* seems to play an important role in the maintenance of the male pathway, as missense mutations in the gene tilt the balance to ovarian-determining fate (82).

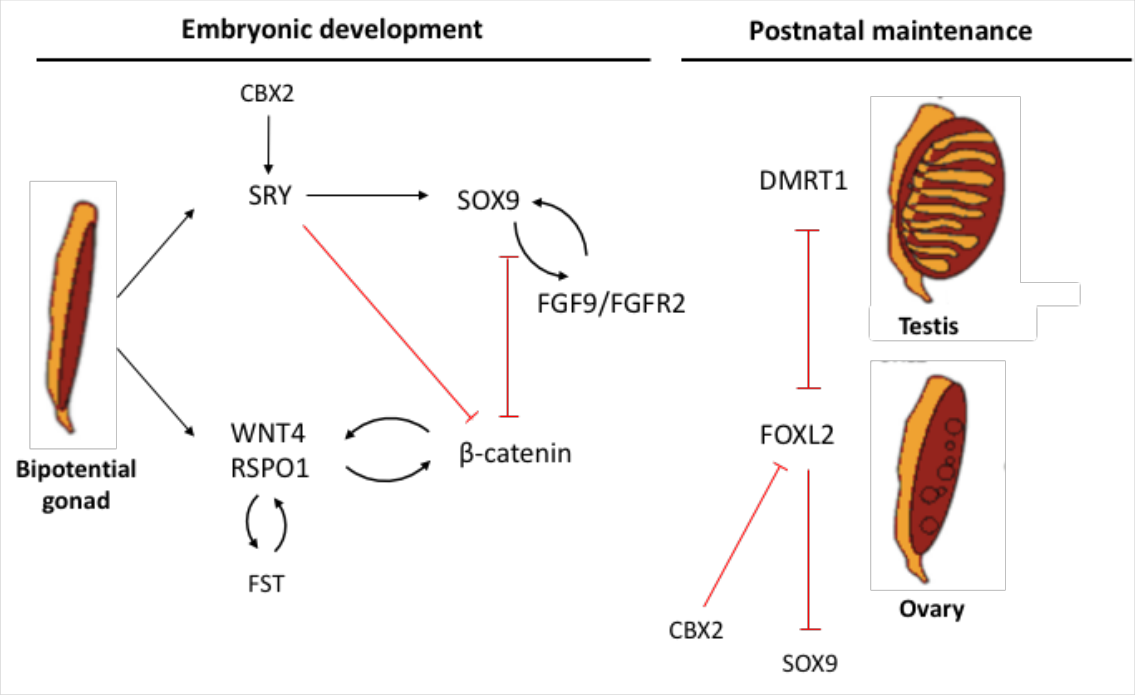

Figure 7. Simplification of the molecular signalling during sex differentiation of the gonads into testes and ovaries and maintenance of the cell fate. Black arrows indicate activation of a downstream target. Red lines ending in bars show repression of a downstream target. Modified from Ohnesorg T., 2014 (2).

Sex identity of the gonadal cells must be maintained postnatally to avoid transdifferentiation. At least in mice, this is achieved by DMRT1 and FOXL2 transcription factors. In *Dmrt1* KO XY adult, transdifferentiation of Sertoli cells into Granulosa-like cells occurred at 4 weeks postnatally because ablation of *Dmrt1* resulted in the activation of ovary-promoting genes, *Foxl2* and in the downregulation of *Sox9* (73). In normal ovary, expression of *FOXL2* in cooperation with ESR1 and ESR2 is shown to bind to the TESCO

enhancer of *Sox9* and inhibit its transcription. In *Foxl2* deleted XX adult mice ovaries, *Sox9* is upregulated and somatic cells (Granulosa and theca) convert into Sertoli and Leydig-like cells. Furthermore, complete deletion of *Esr1/Esr2* or *Cyp19a1* also resulted in transdifferentiation in the adult ovary, as well as germ cell loss (171). These results suggest that maintenance of testis or ovary fate is an active process in adult life, however transdifferentiation in the human gonad has not been clearly probed yet.

2. DISORDERS OF SEX DEVELOPMENT

2.1. DEFINITION AND PREVALENCE

Disorders of sex development (DSD) are a heterogeneous group of congenital conditions in which development of chromosomal, gonadal, or anatomical sex is atypical. Many causes are genetically determined and progression in this understanding led to the re-examination of the nomenclature in 2006 and a new classification of DSD was proposed (Table 2) (172).

Table 2. Classification of Disorders of Sex Development according to the Chicago consensus (172).

46,XY DSD	Disorders of testicular development	Complete gonadal dysgenesis
		Partial gonadal dysgenesis
		Gonadal regression
		Ovotesticular DSD
	Disorders of androgen synthesis or action	Androgen biosynthesis defects (i.e. 5α-reductase deficiency)
		Defects in androgen action (AIS)
		LH receptor defects (i.e. Leydig cell hypoplasia)
		Persistent Müllerian duct syndrome (Variants in *AMH* or *AMHR2* genes)
	Others	Severe hypospadias, cryptorchidism, complex syndromic disorders
46,XX DSD	Disorders of ovarian development	Ovotesticular DSD
		Testicular DSD
		Gonadal dysgenesis
	Disorders of androgen excess	Fetal (i.e. 21-hydroxylase deficiency)
		Fetoplacental (Aromatase deficiency, POR deficiency)
		Maternal (i.e. Luteoma)
	Others	Mayer-Rokitanski-Kuster-Hauser syndrome, complex syndromic disorders
Sex chromosome DSD	45,X0	Turner syndrome and variants
	47,XXY	Klinefelter syndrome and variants
	45,X0/46,XY	Mixed gonadal dysgenesis, ovotesticular DSD
	46,XX/46,XY	Chimerism, ovotesticular DSD

DSD, disorders of sex development

The wide extent of different phenotypes ranges from minor genital abnormalities, such as hypospadias, to ambiguous genitalia and complete gonadal dysgenesis. In general, ambiguous genitalia are a relatively rare situation reported to occur in 1:4,500 life newborns (172), but considering all types of atypical genitalia together with chromosomal, gonadal and genital conditions, the prevalence of DSDs increases to 5 in 1000 births, being boys with hypospadias the 73% of these cases (173). About 75% of the children with atypical genitalia have a 46,XY karyotype, 10-15% a 46,XX and the remainder will have structural anomalies of the sex chromosomes (174). Among the 46,XX DSD, the most frequent genetic condition is congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency, which occurs in its classic form in about 1:14,000-1:18,000 infants while nonclassic form is stimated to have a prevalence of 1:500 to 1:1000 in the general white population. CAH accounts for 90-95% of the individuals with a 46,XX DSD (113). In addition, DSDs are sometimes associated with syndromes or multiple congenital malformations (175). Expected clinical and biochemical features of each DSD is summarized in Supplementary data 1.

2.2. CLASSIFICATION OF DSD BASED ON THE CHICAGO CONSENSUS

2.2.1. 46,XY DSD: Disorders of testicular development

2.2.1.1. Complete and partial gonadal dysgenesis

Gonadal dysgenesis encompass several clinical conditions caused by the irregular development of the fetal gonad in 46,XY. Complete gonadal dysgenesis is characterized by female external and internal genitalia, absence of secondary sexual characteristics, normal or tall stature, eunuchoid habitus without other Turner syndrome features and bilateral dysgenetic gonads. Streak gonads don’t produce adequate testosterone for the development of male internal structures, or AMH with the subsequent persistence of Müllerian ducts that develop to hypoplastic or normal uterus and Fallopian tubes. In some cases clitoromegaly has been observed. These patients are raised as girls and are usually diagnosed at puberty due to infantilism and primary amenorrhea. On the other side, partial gonadal dysgenesis is described with distinct degrees of testicular development and function, as well as a wide spectrum of atypical genitalia and the presence or absence of Müllerian ducts. In both forms gonadotropins are increased, mainly FSH, and testosterone is at a prepuberal range in the complete form while in partial gonadal dysgenesis it varies from prepuberal levels to normal adult levels (175). Estradiol levels are decreased in both forms.

46,XY gonadal dysgenesis has been estimated to occur in 1 out of 10,000 births (172, 176). Regarding the genetic etiology, it is highly heterogeneous and might result from a defect in any gene contributing to gonadal development. However, alterationss in *SRY* are the most frequent cause of gonadal dysgenesis (20%) (177), followed by gene variants in

MAP3K1 (18%) and *NR5A1* (13%). Few are caused by other rare changes in *SOX9* or *NR0B1* (7, 105). Other genes causing 46,XY DSD gonadal dysgenesis are listed in Table 4 (Page 79).

2.2.1.2. Gonadal regression

Testicular regression syndrome (OMIM 273250), also termed vanishing testes, is a condition defined as the absence of testicular tissue in 46,XY individuals with a male phenotype. These individuals are born with male external genitalia, indicating a correct functioning testicular tissue during the first weeks of pregnancy. Micropenis is also present in 50% of the patients and all have hypergonadotropic hypogonadism and infertility (178).

It has been proposed that testicular regression syndrome happens due to vascular accidents, such as testicular torsion. If this occurs in the second half of gestation, penile growth will be impaired due to testosterone deficiency, leading to micropenis. Otherwise the penile length is normal when the accident is close to term or after birth. Thus, the manifestations of testicular regression are associated with the intrauterine accident. Its prevalence is about 1 in 20,000 males In contrast, the coexistence of anorchia with genital ambiguity in a number of familial cases led to the hypothesis that some forms of the condition might have a genetic cause and be part of the clinical spectrum of 46,XY gonadal dysgenesis. However, molecular analyses have failed to identify any gene in association to anorchia but *NR5A1* (178).

2.2.1.3. Gonadal dysgenesis and related syndromes

Several syndromes are associated with 46,XY gonadal dysgenesis in humans, caused by variants in genes involved in gonadal determination.

2.2.1.3.1. WAGR syndrome

WAGR syndrome or Wilms' tumor, aniridia, genitourinary malformation and mental retardation (OMIM 612469) is a contiguous gene syndrome caused by deletions in 11p13 encompassing both *WT1* and *PAX6* genes. Wilms' tumor is developed in approximately 50% of the patients. The diagnosis is based on the presentation of sporadic aniridia in the newborn along with genital anomalies, although these are not always present. In older children, clinical diagnosis is made when aniridia and one of the other features are present (179).

2.2.1.3.2. Denys-Drash and Frasier syndromes

Denys-Drash syndrome (DDS) is characterized by Wilms' tumor in the first decade of life, rapid progressive glomerular disease, and genitourinary abnormalities, mainly 46,XY DSD. Gonadal development is impaired variably, resulting in a heterogeneous spectrum. DDS

is associated with *WT1* gene mutations in the DNA-binding domain, affecting the transcription factor activity in a dominant negative manner (180).

Frasier syndrome (FS) describes the combination of female to ambiguous genitalia in 46,XY males, renal failure in the second decade of life, streak gonads and predisposition to gonadoblastoma. Heterozygous variants in the intron 9 of the *WT1* gene are the most common cause (See 1.1.2, page 5). Splice site variants drive to an imbalance in the expression of the WT1 (-KTS) isoform and underexpression of WT1 (+KTS), indicating that a precise balance between the WT1 isoforms is necessary for the normal function of the protein (10). Since patients with Frasier syndrome do not produce nonfunctional proteins the clinical course of the nephropathy is slowly progressive compared to Denys-Drash syndrome, and require more years to reach complete renal failure. On the other side, Wilms' tumour is not normally observed in cases with Frasier syndrome. It is thought that the -KTS isoform which is predicted to interact with DNA, is responsible for the tumour suppressor activity and therefore protects these patients against the development of the tumour (15).

Specific localization of the genetic changes has been helpful in the differentiation of these two entities in many cases.

2.2.1.4. Ovotesticular DSD

OT DSD in 46,XY DSD patients is explained later (See 2.2.4.1, page 47).

2.2.2. 46,XY DSD: Disorders in androgen synthesis or action

2.2.2.1. Androgen biosynthesis defect

Defects in the enzymes involved in the adrenal and androgen steroidogenesis result in 46,XY DSD. The clinical and molecular characteristics leading to a defect in the synthesis of testosterone and dihydrotestosterone have been explained before (See 1.4.1, page 21).

2.2.2.2. Defects in androgen action

Androgen insensitivity syndrome (AIS) is a disorder of androgen action due to a reduced or absent functionality of the androgen receptor (AR). It is the most frequent known monogenic cause of 46,XY DSD, estimated to be present in 1:20,000 to 1:64,000 male births (140) and is a X-linked recessive condition. The variable phenotypic expression due to the transcriptional activity of the AR has permitted the classification of AIS into complete (CAIS), partial (PAIS) or mild (MAIS) androgen insensitivity syndrome. CAIS phenotype is characterized by female external genitalia, inguinal or abdominal testes, complete breast development and scarce or absent axillary and pubic hair. The absolute resistance to androgens limits differentiation to male external genitalia and virilization, but testis continue

producing AMH and thus, inhibits female internal genitalia and leads to primary amenorrhea. On the other side, PAIS presents either with male phenotype and minor undervirilization, such as hypospadias, or female phenotype with ambiguous genitalia and gynecomastia at puberty, while MAIS is characterized by normal male external genitalia and infertility.

The clinical diagnosis of CAIS is easily confirmed, as about 85% of the cases harbor an alteration in the *AR* coding sequence (CDS), whereas the definitive diagnosis in PAIS is further complicated because less than 30% of the cases present an *AR* gene variant. (7). To elucidate the possibility of a clinical diagnosis of AIS without a variation in the *AR* coding sequence, Hornig *et al* analyzed the genital fibroblast of 46 individuals with a clinical suspicion of AIS and no molecular proof on the CDS using an APOD-assay tool. They found a subset of individuals with an androgen response, but 17 out of 46 presented a functionally proven androgen resistance compared to control groups. They proposed the term AIS type II for this new class of androgen resistance. In addition, some of these AIS type II patients had AR gene variants outside the coding sequence (181).

2.2.2.3. LH receptor defects

Defects caused due to *LHCGR* gene varitants have been described before (See 1.4.4, page 32).

2.2.2.4. Disorders of AMH and AMH receptor

Persistent müllerian duct syndrome (PMDS, OMIM 261550) is an autosomal recessive disorder defined as the presence of müllerian derivatives, uterus, and Fallopian tubes in otherwise normally masculinized 46,XY subjects. Bilateral intra-abdominal testes are found in the majority of the cases (40-60%). However, other cases have been reported with a scrotal testis and a contralateral inguinal hernia containing a testis, uterus and tubes (20-30%) and a hernial sac containing both testes and the müllerian ducts (25%). Most of the cases are diagnosed because of bilateral or unilateral cryptorchidism with or without inguinal hernia. In adulthood the diagnosis of PMDS is made due to the degeneration of the testes or the müllerian derivatives. About 33% of the patients above 18 years have undergone either unilateral or bilateral malignant testicular degeneration, then early orchidopexy is recommended, although it's not completely effective (182).

Homozygous or compound heterozygous variants in *AMH* and *AMHR2* result in an AMH deficiency or resistance, which is the cause of PMDS. There are not anatomical differences between the patients with *AMH* (PMDS type 1) or *AMHR2* mutations (PMDS type 2). However AMH serum levels are different, in those harbouring *AMH* sequence changes anti-Müllerian hormone levels are low or undetectable while in those with *AMHR2*

alterations hormone levels are normal or increased.. Molecular diagnosis is confirmed in approximately 88% of the cases with a persistent müllerian duct syndrome (182).

2.2.3. 46,XY DSD: Others

2.2.3.1. Hypospadias and cryptorchidism

Hypospadias is a congenital anomaly in which the urethral opening is not correctly placed at the tip of the penis. After cryptorchidism, is the most common congenital disorder in boys, with an incidence of 1 in 200-300. Hypospadias occurs when the urethral closure is abnormal or incomplete during the first weeks of embryonic development. Defects in the genes or pathways involved in the phallus development and urethral closure give rise to the atypical location of the urethral opening on the ventral part of the penis and are classified regarding this location into distal or anterior, midshaft or middle and proximal or posterior. This anomaly may be associated with a ventral curvature of the penis, known as chordae and abnormal foreskin (183).

A number of genes have been widely linked to hypospadias, although only the 30% of hypospadias cases have a clear genetic cause. Then, it was suggested that the basis of this anomaly is a combination of genetic susceptibility and environmental influences. In most of the cases, hypospadias appears as an isolated condition, but it can be associated with other anomalies of the urogenital tract mainly, such as undescended testes and micropenis, WAGR syndrome or Denys-Drash syndrome. Despite most of the cases of isolated hypospadias are idiopathic, familial aggregation has been found in about 10% of the cases (183).. Heritability of the condition is above 57% and depends on the severity of the hypospadias. Anterior or middle hypospadias are more frequent in families than proximal hypospadias. In humans several hypospadias association studies have been performed in large cohorts and expression analyses in the affected tissue. Sequence variants have been identified in patients with hypospadias in genes and pathways included in the early development of the genital tubercle, such as the hedgehog signalling pathway or *WT1* gene; the testes signalling pathway, like *MAP3K1*, *MAMLD1* and *NR5A1*; in enzymes necessary androgen synthesis, as *HSD3B2* or in *AR* gene. Regarding the connection between environmental factors and the occurrence of hypospadias, it is well known that an elevated exposure to oestrogenic and anti-androgenic compounds that interfere with the creation or metabolism of endogenous hormones contribute to the incidence of the anomaly. For example, the exposure to EDCs (Endocrine-disrupting chemical) during foetal life is correlated with the risk of hypospadias. Chemical compounds such as dichlorodiphenyltrichloroethane (DDT) block AR affecting the testosterone production, while bisphenol A downregulates the expression of *WT1, LHCGR, HSD17B3* and *SRD5A2* (183).

Cryptorchidism (OMIM 219050) or failure of testicular descent is the most common genitourinary defect and is found in 1.6%-9.0% of newborn boys (135). It can be classified as

unilateral or bilateral depending on the affected testes and can be associated to additional anomalies. In the intra-abdominal phase, occurring between gestation weeks 8-15, INSL3 stimulates the gubernaculum, or genito inguinal ligament, to enlarge and drives the testis to the abdomen, specifically to the caudal part of the abdomen. During the inguinoscrotal phase, between 25-35 gestation weeks, the gubernaculum, derived from the primitive mesenchyme, guides the descent of the testes to the scrotum. Lack of testosterone production by foetal Leydig cells or absence of a functional androgen receptor, stops the testicular descent. Anomalies in the first step of testicular descent are uncommon, whereas those in the inguinoscrotal phase and dependent on androgen action, account for most of the cases (184).*INSL3* and *RXFP2* genes have been implicated in testicular descent and cryptorchidism due to work in mouse models (135).

2.2.3.2. Hypogonadotropic hypogonadism and related gene variants

The increase of the hypothalamo-pituitary-gonadal (HPG) axis release at the beginning of puberty brings secondary sexual features and a mature reproductive system. Absence of puberty due to anatomical or functional defects that end in decreased gonadotropin releasing hormone (GnRH) and/or gonadotropin secretion is termed hypogonadotropic hypogonadism (HH). When the HH has no evident cause is called idiopathic HH (IHH) and can be divided into two major categories: Kallmann syndrome (KS) and normosmic IHH (nIHH). The first is characterized by HH and anosmia due to the interrupted migration of GnRH specific neurons and the olfactory receptor neurons into the hypothalamus. It is often associated with additional congenital anomalies such as cleft palate, unilateral renal agenesis, split hands and feet, short metacarpals, deafness, and synkinesis. nIHH refers to the IHH cases without anosmia and gives rise from the dysfunction of the GnRH neurons in the hypothalamus. However, these two entities seem to be overlapped in some cases. Male infants need the HPG axis active during gestation (between the 16th-22nd weeks) for normal virilization of the 46,XY fetus, then boys with IHH are born with micropenis and/or cryptorchidism. In contrast, there is no clinical manifestation before the first teen years in girls. Overall, pubertal delay is the most typical feature of IHH and is characterized by the lack of breast development in a girl at age 13 or a testicular volume less than 4mL in a boy age 14 (185). . In early infancy, between four to sixteen weeks, a temporary reactivation of the HPG axis happens (minipuberty) which is used as a diagnosis tool (186).

Up to 50% of the IHH are known to be caused by gene sequence defects and around 50 genes have been reported. Oligogenic inheritance in a patient is thought to account for nearly 20% of the cases (185). *GNRHR (*Gonadotropin Releasing Hormone Receptor) and *TACR3* (Tachykinin Receptor 3) gene variants are the two most common causative genes observed in nIHH, but also *KISS1R* and *KISS1*. *KISS1* (Kisspeptin) and *KISS1R* (KISS1 Receptor) gene alterations affect the pulse release of GnRH rather than the migration of GnRH neurons, thus resulting in nIHH exclusively (187). The *KISS1R* encodes the kisspeptin receptor

for the ligand kisspeptin, a secreted neuropeptide that acts as an upstream regulator of the GnRH neurons.

2.2.4. 46,XX DSD: Disorders of ovarian development

2.2.4.1. Ovotesticular DSD

Ovotesticular (OT) DSD is considered as the presence of both testicular and ovarian tissue in a single patient, either in the same gonad, in different gonads or as an ovary and a contralateral testis. In the gonadal biopsy, ovarian tissue is often normal and follicular growth, ovulation (50% of the cases) and oestradiol secretion that inhibits spermatogenesis and leads to Leydig's cell hyperplasia in the testicular tissues is observed (151). Infants are born with ambiguous genitalia and are usually assigned as males but at puberty breast development and menarche is possible due to oestrogen secretion in the ovarian tissue (188). The incidence of OT DSD is about 1 in 100,000 births and only the 65% have a 46,XX karyotype, whereas a 10% correspond to 46,XY and the remaining have sex chromosome mosaicisms (151). Among the sex chromosome abnormalities 46,XX/46,XY chimerism (12.8%) is the most frequent, followed by 46,XX/47,XXY (5.6%) and 45,X0/46,XY (3.5%) mosaicism (189).

From a molecular point of view, *SRY* gene is localized in 10% of 46,XX OT DSD due to a translocation but in the majority of the cases the genetic diagnosis remains unknown (190). Heterozygous and homozygous variants have been detected in *WNT4* in patients with 46,XX OT DSD and testicular DSD as part of a complex syndrome and also loss-of-function sequence changes in *RSPO1* (99, 191, 192). Moreover, copy number gains were described upstream the regulatory region of *SOX9* in three patients with 46,XX OT DSD (44) and an inversion of the long arm of chromosome 22 in another *SRY*-negative 46,XX OT patient (193), increasing the genes involved in the development of female-to-male sex reversal. A *de novo* gain in Xq27.1, including *SOX3* gene was recently discovered in a 46,XX patient with ambiguous genitalia and bilateral ovotestes (55). Additionally, some well-known genes promoting the ovarian and testicular pathways are also able to cause OT DSD, for example *NR0B1* and *NR5A1* (194, 195).

In contrast, few causes of 46,XY OT DSD have been noticed, such as variants in the *SRY* gene or *SOX9*. A partial deletion in the *DMRT1* gene was also detected in a 46,XY OT DSD (196). In patients with chimerism or mosaicism, the coexistence of both gonads explains the phenotype.

2.2.4.2. Testicular DSD

Testicular (T) DSD is described in individuals with a 46,XX karyotype, male external genitalia that varies from normal to ambiguous, the absence of Müllerians remnants and

two testicles. These patients are often infertile and the presence of hypospadias is reported in 10-15% of the cases. At birth are referred as ambiguous genitalia while after puberty, they present small testes, gynecomastia and azoospermia. The total incidence is 1:20,000-25,000 in newborn (151).

The 90% of the cases are caused by the translocation of the *SRY* gene to the X chromosome, which is sufficient to drive to male differentiation, short stature and produce gynecomastia. Other gain-of-function alterations have been found in the upstream region of *SOX9* (45). Loss-of-function in genes involved in ovarian development such as the aforementioned *RSPO1* and *WNT4* also contribute to T DSD (190).

2.2.4.3. Ovarian dysgenesis

Females with a 46,XX karyotype and ovarian dysgenesis are born without ambiguity, but present later with absent or delayed puberty referred as primary or secondary amenorrhea. External genitalia are formed normally. Since streak gonads are found, oestrogens are not produced and the gonadotropins secretion is not inhibited by the hypothalamic-pituitary-gonadal axis, causing elevated gonadotropins levels which result in failure to begin puberty and menarche (43).

Indeed, phenotype of ovarian dysgenesis is not uniform and variable expression has been observed. In several families, one affected sister presented with secondary amenorrhea due to ovarian hypoplasia while another had streak gonads and primary amenorrhea (197). 46,XX gonadal dysgenesis is very rare in childhood and might happen due to a sequence variant in genes involved in ovarian development in an autosomal recessive manner. Many of the 46,XX ovarian dysgenesis cases happen due to disturbances in hormonal signaling and folliculogenesis. As previously mentioned, inactivating mutations in *FSHR* and *LHCGR* genes disrupt the correct development of the ovarian follicle while *BMP15* and *PSMC3IP* act as modulators of FSH and other nuclear receptors.

Primary ovarian insufficiency (POI) is a term used to include a spectrum of ovarian dysfunctions such as ovarian dysgenesis. Indeed, the distinction between the two entities is difficult, so the inconveniency to identify the 46,XX ovarian dysgenesis-specific genes. Among the newly identified genes in association with ovarian dysgenesis and POI, *NOBOX* and *FIGLA* seem to be the more representative. The Newborn Ovary Homeobox-Encoding Gene (*NOBOX*) is expressed in germ cell cysts, primordial and developing oocytes in mouse models and humans. Loss-of-function variants in *NOBOX* accounted for 6.2, 6.5 and 5.6% of the cases in three POI cohorts (198). Folliculogenesis specific bHLH transcription factor or *FIGLA* has a specific role in the development of primordial follicle and in the synchronization of the genes in the *zona pellucida.* In 2008, an study revealed the presence of *FIGLA* gene variants in the 4% of chinese women with sporadic primary ovarian insufficiency (199).

2.2.5. 46,XX DSD: Disorders of androgen excess

2.2.5.1. Fetal androgen excess due to inborn errors of steroidogenesis

In the virilizing form of CAH, androgen excess results in 46,XX DSD. As mentioned before, deficiencies caused by mutations in *CYP21A2, CYP11B1* and *HSD3B2* lead to specific manifestations in each patient, depending on the severity of the enzymatic defect (See 1.4.1, page 21).

2.2.5.2. Fetoplacental androgen excess

The steroids that are produced during fetal development are transferred to the placenta and transformed intro androgens and subsequently to oestrogens. Deficiencies in the aromatase enzyme or in the POR cofactor due to genetic variants result in accumulation of the precursors and masculinization of the female fetuses.

2.2.5.3. Excess maternal androgen production

The origin of the androgen excess or hyperandrogenism in the mother during pregnancy may be due to an ovarian tumor (luteoma), polycystic ovary syndrome (PCOS) or adrenal tumors. Luteomas are benign tumors of the ovary during pregnancy that produce masculinization of the fetus in about 65% of the mothers that are virilized (200). In most cases the tumors are small, don't produce virilization and regress after pregnancy. PCOS is mainly characterized by androgen excess and reproductive defects associated to metabolic abnormalities, like insulin resistance. Carcinomas and adenomas of the maternal adrenal gland are very rare and may lead to virilization of the mother, among other clinical features such as hypertension or diabetes mellitus, but also ambiguous genitalia due to virilization of 46,XX infants (201).

2.2.6. 46,XX DSD: Others

2.2.6.1. Non-CAH monogenic primary adrenal insufficiency

Primary adrenal insufficiency is characterized by the impaired production of glucocorticoids, mineralocorticoids and hypersecretion of ACTH and could be classified as CAH and non-CAH monogenic primary adrenal insufficiency (PAI), in which adrenals are hypoplastic or normoplastic. Non-CAH monogenic PAI is a group of heterogeneous disorders caused by adrenal hypoplasia, ACTH resistance or impaired adrenal redox homeostasis (202).

The best known transcription factors implicated in the defective organogenesis of the adrenal gland are DAX1 and SF1, although adrenal insufficiency is a rare condition in patients harboring a *NR5A1* gene variant. DAX1 is expressed in the subcapsular region of the adrenal

glands and is involved in the maintenance of adrenocortical progenitor cells, while the expression of SF1 begins in the adrenocortical primordium and continues to steroidogenic tissues. The age of onset of patients with *NR0B1* gene mutations is younger than the rest. On the other hand, ACTH resistance occurs due to genetic defects in the ACTH signal transduction and depending on the residual signal transduction activity the size of the adrenal glands is normal or small. Up to date two genes have been identified in the ACTH signal transduction, *MC2R (*Melanocortin 2 Receptor) encoding the ACTH receptor and *MRAP (*Melanocortin 2 Receptor Accessory Protein*),* an accessory protein of melanocortin 2-receptor (202). Finally, adrenal redox homeostasis presents, usually, with normal aldosterone production, as ACTH resistance. Array-based analyses in consanguineous patients with familial glucocorticoid deficiency (OMIM 202200) allowed the identification of responsible gene *NNT* (Nicotinamide nucleotide transhydrogenase), which codifies for a protein in the transfer system of the respiratory chain in the mitochondria (203).

In general, patients with non-CAH monogenic PAI present symptoms related to adrenal insufficiency, such as vomiting and skin pigmentation, but also salt-wasting manifestation including dehydration and low blood pressure because aldosterone production is affected even in some patients with ACTH resistance and impaired adrenal redox homeostasis. Genital anomalies are frequently observed in reported PAI patients.

2.2.6.2. Mayer-Rokitansky-Küster-Hauser syndrome

Mayer-Rokitansky-Küster-Hauser (MRKH, OMIM 277000) syndrome is characterized by the congenital absence of the upper two-thirds of the vagina due to interrupted embryonic development of the Müllerian ducts. Otherwise, patients have a 46,XX karyotype and develop secondary sex characteristics normally. At puberty, they present amenorrhea, infertility and normal female phenotype and external genitalia. Anomalies of the genital track may range from upper vagina atresia to Müllerian agenesis and is classified as MRKH type I when is isolated or as type II when is associated to malformations, such as renal defect or skeletal malformations. The incidence of this syndrome is 1 in 4,500 newborn females.

Candidate gene sequencing identified *WNT4* as the monogenic cause for MRKH but aCGH (array-based Comparative Genomic Hybridization) analyses allowed the recognition of duplications and deletions in other genes. The most frequently affected chromosomal regions are 17q12, encompassing *LHX1* (LIM Homeobox 1) and *HNF1* (Hepatocyte Nuclear Factor 1) genes (6%), and 22q11 (4%), followed by 16p11.2 (1%) and 1q21.1 (1%) where rearrangements in *TBX6* (T-Box Transcription Factor 6), *RBM8A* (RNA Binding Motif Protein 8A) and in the Xpter pseudoautosomal region 1, including *SHOX* (Short stature Homeobox) genes have been reported. Nevertheless the etiology of the syndrome remains unknown because most of the cases are sporadic, although segregation analyses in few families have shown an autosomal dominant inheritance with incomplete penetrance. Oligogenic or polygenic inheritance has also been suggested (204). More recently, 4 new point variants

have been described in *MYCBP2* (MYC Binding Protein 2), *NAV3* (Neuron Navigator 3)and *PTPN3* (Protein Tyrosine Phosphatase Non-Receptor Type 3) in a Japanese cohort (205).

2.2.7. Sex chromosome DSD

2.2.7.1. 45,X0: Turner syndrome and variants

Turner syndrome is associated with a complete or partial missing of the X chromosome and is a rare condition in females characterized by hypergonadotropic hypogonadism, infertility, short stature, endocrine and metabolic illnesses and augmented risk of autoimmune disease among others. The syndrome occurs with a prevalence of 1:2,000 women in different ethnic populations (206).

Different karyotypes have been observed in women with Turner syndrome. Nearly 50% of them present 45,X0 (complete loss of one X chromosome), 15-25% 45,X0/46,XX mosaicism, an isochromosome of p or q arm (20%) and less, an X ring chromosome. Hypergonadotropic hypogonadism is present in almost all patients and drives either to primary or secondary amenorrhea and then, to infertility, possibly due to an accelerate loss of oocytes from the ovaries since fetal development. The onset of puberty depends on the karyotype as only 2-3% of women with 45,X0 karyotype present menstrual cycles. Breast development at puberty occurs in 21-50% of the patients and menarche in 15-30%. Moreover, spontaneous puberty has been correlated with viable follicles, younger age, mosaic karyotypes and normal levels of FSH and AMH Spontaneous pregnancy happens in about 6% of these individuals, mainly in mosaic Turner syndrome, although the rate of natural miscarriage is high (206).

SHOX is the only gene related to Turner syndrome. This gene belongs to the paired homeobox family and is located in the pseudoautosomal region 1 (PAR1) of X and Y chromosomes. *SHOX* (Short stature homeobox) controls proliferation and maturation of chondrocytes in the growth plate, regulates the expression of NPPB (Natriuretic Peptide B) and FGFR3 (Fibroblast Growth Factor Receptor 3), and interacts with transcription factors SOX5, SOX6 and SOX9.. Then, the reduced expression of SHOX elucidates the growth deficit, as women with the syndrome are about 20 cm shorter than their target height. Although other genes like *TIMP1* (Tissue inhibitor of matrix metalloproteinase 1), may be implicated in the frequent bicuspid aortic valves and aortic dilation observed in Turner syndrome individuals, further studies are required. It is now emerging the possible impact of epigenetic changes and variants in RNA expression, as well as protein-protein interactions in the development of the pathogenesis of Turner syndrome, however the genomic mechanisms need to be further studied. (206).

2.2.7.2. 47,XXY: Klinefelter syndrome and variants

Klinefelter syndrome is the most common male chromosomal disorder. It is reported to occur in 1 out of 500 or 1000 male births and is mostly diagnosed later in life. Moreover, about 64% remain undiagnosed . The 90% of the cases present a 47,XXY karyotype while the remaining include 46, XY/47, XXY mosaicism, 48, XXXY or 49, XXXXY aneuploidies and structural irregularities in the X chromosomes, such as 47,iXq,Y. This genetic background is based on the failure of chromosomal separation or non-disjunction during meiosis I and II or mitosis, which leads to the presence of an extra X-chromosome.

The clinical features in patients with Klinefelter syndrome are highly heterogeneous and depend on the supernumerary X chromosome and the consequences of hypogonadism. Common presentation is testicular failure with small gonads, hypergonadotropic hypogonadism, gynecomastia, sparse body hair, signs of androgen deficiency, azoospermia, difficulties in language processing, social and learning disabilities and tall stature with eunuchoid body. Infants are generally born with a normal male phenotype, although cryptorchidism and inguinal hernias have been reported too (207). It has been suggested that the phenotypic variability of the syndrome depends on the expression of altered genes, androgen deficiencies, variants in *AR* such as the number of CAG repetitions or random inactivation of the additional X-chromosome material. Certainly, few studies have reported the presence of complete or partial androgen insensitivity syndrome due to *AR* gene variants in 47,XXY patients (207).

2.2.7.3. 45,X0/46,XY: Mixed gonadal dysgenesis and ovotesticular DSD

The 45,X0/46,XY karyotype has an estimated incidence of 1 to 15,000 births and the chromosome constitution arises from the mitotic errors that take place during early divisions. Among the different phenotypes that present with the 45,X0/46,XY karyotype, mixed gonadal dysgenesis (MGD) is characterized by ambiguous genitalia due to a streak gonad and one or two dysgenetic testes. Undifferentiated gonadal tissue (UGT) has been also noted. This, has the same characteristics as streak tissue but contains germ cells that are at increased risk of neoplastic transformation.The risk of developing a malignant germ cell tumor, is increased in patients with a DSD and a specific part of the Y chromosome in their karyotypes, more specifically the aberrant expression of the *TSPY* (Testis Specific Protein Y) gene (208). On the other side, the development of the male reproductive tract is going to be determined by the ratio of germ line cells expressing the XY genotype, and then completely male and female phenotypes might be observed (209). Phenotypes similar to partial gonadal dysgenesis can also result from 45,X0/46,XY karyotype. Recently, a positive correlation between *AR* activity and external virilization in 45,X0/46,XY patients has been described (210).

Structural rearrangements of the Y chromosome, including deletions, ring chromosomes and isochromosomes lead to 45,X0/46,XY mosaicism and have been described in mixed gonadal dysgenesis, but also in Turner syndrome and infertility (211). Yp (short arm) deletions including *SRY* gene directly affect testis differentiation leading to streak gonads and a female phenotype, whereas deletions of the long arm (Yq11), especially involving the Azoospermia Factor regions (AZF) AZFa, AZFb and AZFc on Yq11 lead to male infertility (212).

The incidence of OT DSD in individuals with 45,X0/46,XY mosaicism is low and additionally, the distinction between OT DSD and MGD may be complicated. Clinical and molecular characteristics of the OT DSD have been already explained (See 2.2.4.1, page 47).

2.2.7.4. 46,XX/46,XY: chimeric and ovotesticular DSD

Clinical and molecular characteristics of the chimeric and ovotesticular DSD have been already explained (See 2.2.4.1, page 47).

2.3. MANAGEMENT OF DSD AND THE RISK TO DEVELOP A GERM CELL TUMOR

Caring for patients with DSDs require a multidisciplinary team (172), that integrates medical and other professional and social science disciplines (i.e. Psychologist). Patients with DSD should be followed-up to guarantee the correct transition of care from childhood and adolescence to adult life. When referred to clinician, the clinical evaluation of individuals should include family and prenatal history, the appearance of the genitals and external phenotype, molecular diagnosis and the use of techniques like ultrasonography or laparoscopy to assess gonadal location and description of the tissue. Molecular diagnosis includes karyotyping and gene analysis. Biochemical investigations are still used as a tool for the identification of the mechanism. In newborns, the measuring of 17-hydroxyprogesterone levels is used to discard CAH and facilitates diagnosis of overvirilized 46,XX. Hormone therapy is going to be necessary in most patients with DSD. During childhood, the treatment is critical because hormone therapy leads to irreversible effects and induction of puberty is performed according to practices used for adolescents with pituitary or gonadal failure (213). Surgical treatments are performed to allow the correct development of external genitalia and remove the internal structures. Regarding the sex of rearing, it should be considered the development of the patient at puberty and future fertility. Even though no recommendations exist for sex assignment in newborns with DSD, the social sex is favoured (172).

Overall, reaching a molecular diagnosis is useful in relation to possible gender assignment, evaluation of gonadal and adrenal function, risk of gonadal cancer, infertility and other consequences (174). Some groups of DSD patients are more disposed to the

development of a germ cell tumor. Phenotypes with gonadal dysgenesis in 46,XY and sex chromosomal DSD are highly heterogeneous and the resulting gonad depends on the genetic variant. Normally, the earlier the gonadal development is disturbed the less the gonad is differentiated into a testis or an ovary. The different configurations that are recognized are: normal testes with occasionally immature germ cells, dysgenetic testes, primitive sex cords including germ cells, streak gonads with some testicular or ovarian differentiation and undifferentiated gonadal tissue containing isolated germ cells (214). A combination of these configurations may be found in the same patient within the same gonad or the opposite one.

In undervirilized 46,XY DSD, the male gonad is a testis with immature germ cells that are derived to apoptosis and therefore postpubertal gonads lack germ cells unless a tumor has surfaced. Germ cell tumors, mainly carcinoma *in situ* (CIS) and gonadoblastoma (GB), emerge with increased frequency in patients with a specific region on Y chromosome, namely the testis specific protein Y (TSPY). Then, overvirilized 46,XX DSD patients are not at risk. Immature germ cells express OCT3/4 (Octamer-Binding Protein 4), a transcription factor expressed in fetal undifferentiated germ cells but not later. Immunohistochemical studies revealed high positivity for OCT3/4 and TSPY expression in the germ cells of CIS and gonadoblastoma and has been proposed to be of pathogenetic relevance in the formation of germinal cell tumors in these patients (214).

Moreover, gonadal differentiation also determines the tumor risk. Mostly, undifferentiated gonadal tissue gives rise to germ cell tumors. Immunohistochemistry is used to evaluate the number and maturation of germ cells according to their location in the seminiferous tubule. Thus, early CIS or GB is discarded or an estimation of the risk for the development of an invasive germ cell tumour is made. The incidence of a germ cell tumor in DSD varied between 5 to 7% but more recent reviews reported that the risk is extremely low in CAIS compared to PAIS or patients with *HSD17B3* gene variants. Overall, the incidence is about 2.3% in underviriled patients and as gonads are well differentiated CIS is only described in these populations. However, in cases with gonadal dysgenesis the risk increases up to 12% and GB is developed, although CIS has also been reported. Usually, gonadoblastoma appears in these patients before puberty, which indicates the poor role of pubertal hormones in its development. However, the risk of malignancy in patients with 46,XY gonadal dysgenesis is dependent on the underlying genetic cause (214).

In the past, early gonadectomy was performed in 46,XY patients before puberty to prevent degeneration of dysgenetic tissue. Now, tumor risk is predicted based on medical analysis, such as the presence of OCT3/4 positive cells and gonadal biopsies and in some cases, such as in CAIS, gonads are retained until late adolescence to take advantage of the secretion of hormones that allow the development of secondary sex characteristics (215). Table 3 summarizes the recommendations for a stratified treatment model of DSD patients based on defined parameters.

Table 3. Schematic representation of the stratified model for gonadal management in the DSD. Modified from Cools *et al*, 2014 (215).

Gonadal dysgenesis 45,X0/46,XY and 46,XY	Boys	**Mild undervirilization:**
		Orchidopexy
		Self-examination and annual US (postpubertal)
		Biopsy: 1 prepubertal (during orchidopexy) and 1 postpubertal
		CIS/GB: gonadectomy
		Ambiguous genitalia: Low threshold to gonadectomy
	Girls	**CGD:** Bilateral gonadectomy before puberty
		PGD: Bilateral gonadectomy at diagnosis
	Unclear gender identity	Bilateral biopsy
		Low threshold to gonadectomy
Undervirilization 46,XY AIS and testosterone synthesis defects	Boys	**PAIS and testosterone synthesis defect:**
		Orchidopexy
		Self-examination and annual US (postpubertal)
		Biopsy: 1 prepubertal and postpubertal (bilateral biopsy). Repeat after 10 years (?)
		CIS: gonadectomy
		Gonadectomy in case of breast development or testosterone supplementation (?)
	Girls	**CAIS:** postpubertal gonadectomy or follow-up
		PAIS and testosterone synthesis defect: prepubertal gonadectomy
	Unclear gender identity	**PAIS and testosterone synthesis defect:**
		Bilateral biopsy
		Low threshold to gonadectomy

CAIS, complete androgen insensitivity syndrome; CGD, complete gonadal dysgenesis; CIS, non-invasive precursors carcinoma *in situ*; GB, gonadoblastoma; PAIS, partial androgen insensitivity syndrome; PGD, partial gonadal dysgenesis; US, ultrasound.

3. NEXT GENERATION SEQUENCING AND ITS USE IN DSD

Advances in DNA-based techniques have improved diagnosis and management of human genetic diseases, including disorders of sex development.

Since the highly variable aetiology of DSD, a large number of genes have been considered causative and thus, the arrival of next generation sequencing (NGS) completely changed the diagnostic strategies from the gene-by-gene sequencing to the massively parallel sequencing. Gene panel sequencing, in which the genes of interest are sequenced, has become the first approach to be used for DSD diagnosis. It combines a fast analysis of a large number of genes with adequate gene coverage. As targeted panels are restricted to known or candidate genes and lack consistent methods for copy number variant (CNV) analysis, molecular diagnosis is limited. Reduced sequencing costs and increasing bioinformatics expertise, have improved the practice of whole exome sequencing (WES) and whole genome sequencing (WGS) approaches. WES and WGS are more flexible than DSD-related gene panel and allow the identification of new genes. While whole exome sequencing is mainly limited to the detection of single-nucleotide variants (SNV) and small

indels in the total exons present in the human genome, WGS will eventually become first-line clinical test as it detects all variant types in both coding and non-coding regions (216).

Few studies have screened a large targeted gene panel or have guided the sequencing of the whole exome to analyse known and candidate DSD genes. Baxter *et al* as a continuation of a previous work, screened forty-seven 46,XY DSD patients and evaluated 64 known genes (81), while another study reported the findings in the *NR5A1* gene in thirteen 46,XX testicular and ovotesticular DSD (194). Bigger targeted gene panels have been used for the analysis of either small (217, 218) or larger cohorts, as the study by Eggers *et al* in which gene panel including 64 well-known DSD and 967 candidate genes was designed to test 326 international DSD individuals (12) More recently, the efficiency of the method has been reinforced with the use of targeted panels of 30, 56 and 67 genes (93, 219, 220). The diagnostic rate varied from 20.7% to 38.6% in these analyses, supporting the use of targeted NGS as a first-tier diagnostic tool for DSD. (218) (12)

WES and WGS facilitate the discovery of new disease genes or genomic loci involved in the pathogenesis of DSD. Moreover, new roles for well-known DSD genes are increasingly described. This further advance also increases the amount of data and complicates the interpretation of the results, mainly when detecting variants of unknown significance (VUS). (195). In some occasions the interpretation of a variant is direct, such as novel variants in a well-known gene. However, when the relationship between the variant and the phenotype is not clear, functional assays are used to determine the effects of an alteration on a gene (195). In order to avoid incidental findings or findings that cannot be interpreted, the European Society of Human Genetics (ESHG) as well as the current guidelines for genetic testing in DSD, have recommended the use of a targeted approach first, when the clinical and biochemical results are suggestive for specific genes, and whole exome or genome sequencing if this is not the case. A complete list of genes that contribute in human DSD was recently reported by the international registry DSDnet and included 62 genes in 46,XY DSD and 61 in 46,XX (174). (174).

As well as in many other developmental disorders, gene dosage effects contribute to the pathogenesis of a big percentage of DSD patients. As previously stated, CNV detection in the whole genome is limited to WGS, although several algorithms for assessment on WES and gene panel data are available (216). Studies of gene copy number were initially done with karyotype banding, which only detected large chromosome rearrangements, and FISH (Fluorescence in situ hybridization), which was employed to identify smaller CNV. Nowadays, detection of intragenic and whole gene CNV may be done by MLPA (Multiplex Ligation-dependent Probe Amplification), but aCGH (array-based Comparative Genomic Hybridization) are extensively used to detect small to large-scale duplications and deletions (<5 Mb) along the genome. The intragenic deletions are not always detected by Sanger sequencing (221) and therefore these techniques have become part of the routine investigations. Besides the diagnostic potential shown by the aCGH (26, 28, 222), novel

gonad genes can also be detected (223). Additionally, arrays have been crucial in the analysis of regulatory regions in genes such as *SOX9* and *GATA4 (26)* (44).

Previous to new molecular diagnosis techniques, the investigation of DSD involved the Sanger sequencing of a small number of known genes, depending on the disease phenotype, and in combination with a MLPA for the analysis of bigger rearrangements. Many variants in gonad determining and differentiating genes, such as *SRY* or *SOX9*, were discovered by sequencing candidate genes and screening the appropriate DSD patients. However, the difficulty to define many DSD phenotypes complicated the choice (195) and large number of patients did not achieve a molecular diagnosis. Currently, candidate gene sequencing is mainly used when the massively parallel sequencing fails to interpret a result as in complex rearrangements, to sequence a region of interest in segregation analysis and to validate the variants found by NGS (174). The heterogeneous nature of DSD, as well as the increasing evidence of an oligogenic mode of inheritance has led to the replacement of individual gene sequencing also when clinical and biochemical profiles point to a specific gene (25, 224). Nevertheless, the sequencing of larger cohorts with novel genetic and genomic technologies is expanding our knowledge on the genetic changes leading to DSD, and may improve the understanding of the observed phenotypes (195). The use of these fast diagnostic tools enables better diagnoses and has changed the clinical diagnostic pathway in DSD. The traditional approach consisted on clinical examination including biochemical investigations, karyotype determination and testing for CNV in known DSD genes, mainly when DSD were associated with other malformations (222). Gene sequencing, both by Sanger or NGS, was often the last step and was based on the gathered phenotypic information (225). However, genome wide sequencing has supported a parallel approach in which biochemical and molecular evaluations are complementary (Figure 8). Nevertheless, the best strategy should be determined based on karyotyping, clinical and hormonal phenotype and family history (174, 226).

Technology is constantly changing and rapid advances in machines and platforms occur, coverage is improved and new bioinformatic tools are created for the interpretation of variants identified. Current molecular techniques, in the form of aCGH or NGS allow a faster analysis of patients when targeted to known genes, while broaden approaches (WES/WGS) increase diagnostic rates and provide the opportunity to discover novel genes, developmental pathways and regulatory mechanisms in undiagnosed DSD patients, improving the understanding of molecular aetiology. Although the diagnostic process of rare diseases has been enhanced with the current molecular analyses, these results need to be placed in the correct clinical context. In the past, genetics were performed to confirm the biochemical findings, nowadays, NGS allows sequencing while essential clinical tests are done and thus, saving healthcare system's resources. In any case, a complete examination including physical and biochemical testing should be placed to functionally assess the relevance of the gene mutation (7). Genome-wide sequencing has improved diagnostics, however about 50% of individuals with 46,XY DSD do not reach a genetic finding, which

highlights the importance of underlying molecular mechanisms (12, 226). Beside the psychological circumstances that patients need to face, DSD may be associated to other conditions such as cardiovascular disease, obesity or cancer (226), that hinder the management of DSD and a well-founded diagnosis in individuals with such complex conditions. The advances in the understanding of molecular mechanisms, technology and expertise enable a better clinical practice in people with DSD, allowing the clinician to predict the prognosis and long-term outcome.

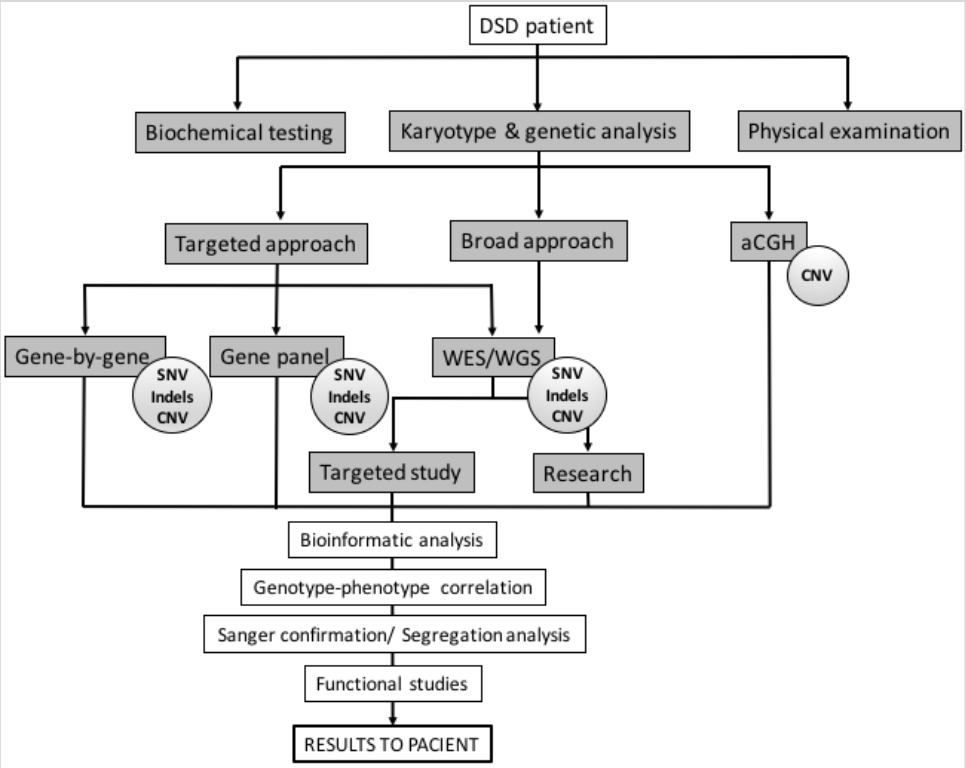

Figure 8. Diagnostic approach in which the patient's clinical and biochemical features are studied in parallel with the genetic results. Modified from the recommendations given by the DSDnet (174).

HYPOTHESIS AND OBJECTIVES

HYPOTHESIS

Disorders of Sex Development encompass a heterogeneous group of conditions in which multidisciplinary care is essential. Clinical examination, biochemical tests and karyotype determination, along with genetic investigations have a significant impact to assess the aetiology of the disease.

Genetic testing prior to novel technologies was dependent on the clinical information and the suspicion of the involvement of a specific gene. With the arrival of next generation sequencing, new genes and pathways are being implicated in the pathogenesis and the genetic basis is gradually being elucidated. Previous studies have analysed cohorts of individuals with DSD, though reaching a specific molecular diagnosis remains challenging in nearly 50% of the 46,XY DSDs.

Therefore, the hypothesis of this present work is that molecular analysis either by a gene-by-gene approach or using massively parallel sequencing technologies will identify the genetic cause in these patients and together with the clinical diagnosis determine a phenotype-genotype correlation.

MAIN OBJECTIVE

Clinical and molecular characterization of patients with a disorder of sex development using a candidate gene approach or a targeted gene sequencing panel in those without a previously causative genetic variant identified in a traditionally associated gene.

Specific objectives

1. Implement a single-gene and a targeted gene sequencing panel analysis to genetically characterize a cohort of DSD patients /Improve the genetic characterization of patients with disorders of sex development using a single-gene and a targeted gene panel as a diagnostic tool.

2. Identify genetic variants that explain the development of the disease in patients with DSD using a single-gene approach and a targeted gene panel.

3. Functional characterization of identified variants either by the single-gene testing or next generation sequencing to stablish the impact of the genetic change in the pathogenesis of the disease.

4. Establish a phenotype-genotype correlation between the clinical features of the patient and the genetic defect.

PATIENTS AND METHODS

4. STUDY DESIGN

Clinical and molecular analyses were performed in all the included patients.

4.1. DESCRIPTION OF THE PATIENTS

Clinical characterization of each proband, and relatives when applicable, was done after examination of the clinical data sheet collected in our Institute. For biochemical and hormonal parameters, values were compared based on age and sex-appropriate range of the hospital of origin according to the reference values recommended by the Spanish Society for Paediatric Endocrinology (SEEP, Sociedad Española de Endocrinología Pediátrica) working group on DSD in 2017 (227).

Clinical and biochemical data of the 125 patients with DSD diagnosis that have been included in this study are summarized in Supplementary data 2 and 3.

4.2. MOLECULAR STUDY

The molecular analysis was carried out in three different phases (Figure 9).

4.2.1. Candidate gene sequencing approach for DSD-related genes

Candidate genes are chosen on the basis of either their functional relevance to the disease pathogenesis or their locations within implicated chromosomal regions. This approach has been applied for gene-disease research, among others. The first gene selected for testing is the more related to the disease. If no alteration is found, a second gene is chosen and successively, until a mutation explaining the phenotype of the patient is found. The success of this approach depends upon the correct choice of the genes to be studied and is limited by its support on existing knowledge about the theoretical biology of disease.

Among included patients, 36 were genetically diagnosed for classical DSD-associated genes by a candidate gene approach using traditional Sanger sequencing technique, Multiplex Ligation-dependent Probe Amplification (MLPA), Quantitative Multiplex Polymerase chain reaction of Short Fluorescent Fragments (QMPSF) or a Comparative Genome Hybridization array (aCGH), depending on the suspected gene variants causing the phenotype of the individual.

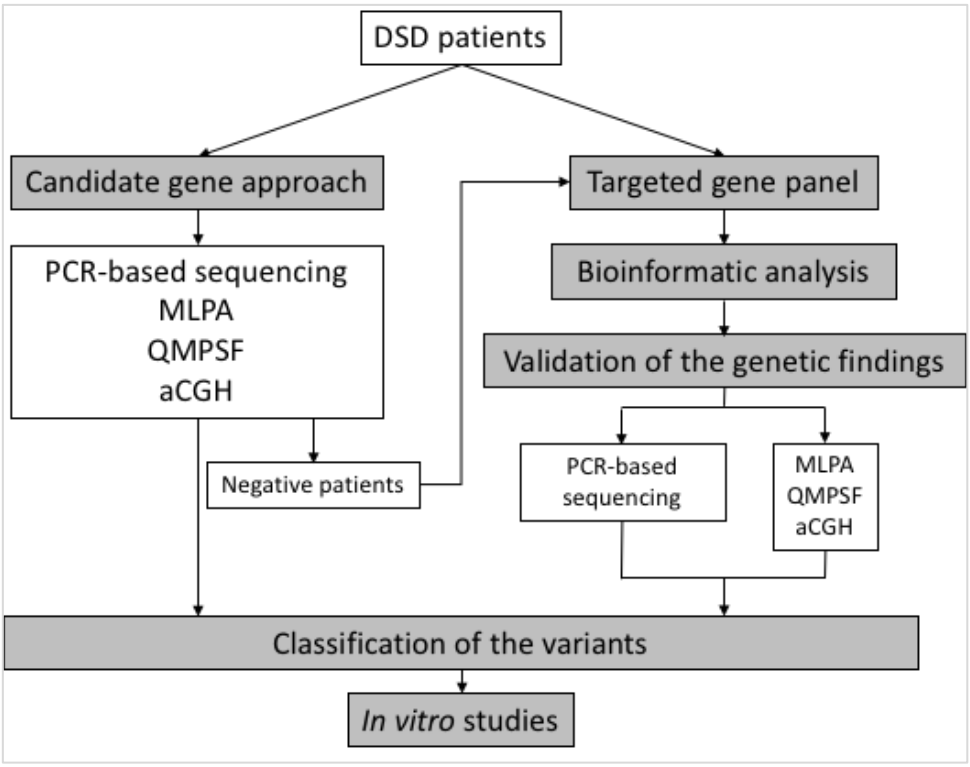

Figure 9. Schematic representation of the molecular study developed in this work.

4.2.2. Next generation Sequencing: targeted gene panel for DSD

After the implementation of next generation sequencing (NGS) technology in our Institute, patients collected since 2015 were analysed with a customized gene panel for DSD. The testing panel was designed to contain a total of 48 genes, associated with sex determination, sex differentiation and hypogonadism. The list of the genes in the customized panel can be found in Table 4 (Page 79).

Individuals with a previous negative result in the gene-by-gene study were also included in this procedure. Altogether, 89 subjects were studied in the panel.

Among included clinical cases 11 positive controls with variants confirmed by Sanger sequencing were incorporated to the NGS study to validate the assay. These controls comprised 8 index cases diagnosed with DSD and 2 relatives of an affected patient, included due to the absence of DNA of the proband. These cases presented with changes in the *SRY*, *HSD17B3*, *AR*, *NR5A1 and NR0B1* genes. Additionally, 1 subject diagnosed with Wilms' tumour harbouring a known mutation in the *WT1* gene was included. *WT1* is a well-known gene that is also involved in the gonadal development in 46,XY DSD and is comprised in the customized panel. Four negative controls were also tested to ensure the efficacy of the methodology.

Variants of interest identified by panel analysis were verified by Sanger sequencing.

An aCGH, MLPA or Fluorescence *In Situ* Hybridization (FISH) was performed for complete genetic diagnosis of some patients in which deletions, amplifications or complex rearrangements were suspected, as well as to validate the findings of the NGS technique.

For both approaches, candidate gene and NGS approach, parents and family members were tested to establish the mode of inheritance, when available.

4.2.3. Functional characterization of likely pathogenic variants or variants of unknown significance

The main objective of these functional studies was to analyse the potentially pathogenic alterations mapped in three different genes in cells and determine the effect in the pathogenesis of the disease.

5. CLINICAL DATA OF THE PATIENTS

In this study, we have performed the clinical and genetic analysis of patients affected by a Disorder of Sex Development (DSD). A total of 125 index individuals with a clinical diagnosis of DSD were evaluated between 2002 and 2018.

This multicentre study included patients attending an Endocrinology or Paediatric Endocrinology service from several Spanish and one Swiss hospital (Supplementary data 4).

The inclusion criteria were as follow:

- Clinical diagnosis of a DSD, according to the consensus statement on management of intersex disorders (172)
- Complete clinical data sheet of the patient
- Informed consent from patients and family

Clinical data were provided by the different clinicians. Clinical data sheet included age at consultation, assigned gender, karyotype of the patient and family background comprising consanguinity of the parents, virilisation during pregnancy or infertility. At examination, data about external genitalia, Tanner stage in puberty, biochemical and hormonal values as well as treatment and progression were gathered. If performed, comments on Magnetic Resonance Imaging (MRI) and ultrasound (US) images were achieved, and reports from the gonadal biopsy were included. Lastly, previous molecular studies were also annotated. The clinical data sheet is included as Supplementary data 5.

Written informed consent was obtained at the respective hospitals involved from all subjects and their family members after full explanation of the purpose and nature of all the procedures used. The study was approved by the corresponding ethical committees.

6. SAMPLE COLLECTION AND STORAGE

Blood samples or DNA (Deoxyribonucleic Acid) from the patients and family members were obtained during routine examination at the different hospitals, and together with the clinical data and informed consent, were collected at the Biocruces Bizkaia Health Research Institute (Barakaldo, Spain).

Two different blood samples containers were used depending of the molecular study to be performed:

- Blood sample on EDTA-containing (Ethylenediaminetetraacetic acid) tubes were used for the extraction of genomic DNA and its application on candidate gene sequencing approach and NGS, as well as validation of the variants by Sanger sequencing and CNV (copy number variant) by MLPA and aCGH.
- Blood samples on Sodium Heparin were used for the FISH technique in which lymphocytes life long enough to yield a proper cell culture.

Each sample was registered in the laboratory and identified with a family name. Then, blood samples on EDTA-containing tubes were stored at -40°C until use. On the contrary, samples on Sodium Heparine were kept at room temperature and procedure for fixing cells was done within the same day the sample was obtained. Slides with the fixed lymphocytes reared at metaphase were stored at -20°C until use.

7. MOLECULAR STUDY

7.1. GENOMIC DNA EXTRACTION

Genomic DNA was extracted from peripheral blood leukocytes using either the manual QIAamp DNA Blood Mini kit (Qiagen NV, Venlo, Netherlands) or the automated MagPurix 12S system from Zinexts (Zinexts Life Science Corp., New Taipei City, Taiwan).

7.1.1. Manual DNA extraction

The purification procedure comprises 4 steps and is carried out using spin columns. After cellular lysis, DNA is adsorbed onto the membrane of the column and the washing steps remove contaminants. Finally, purified DNA is eluted.

This method is designed for a purification of an average of 6µg of total DNA from 200µL of whole human blood.

Reagents and materials

- Absolute ethanol (C_2H_5OH) (Thermo Fisher Scientific, Cat. No. BP2818100).

- QIAamp DNA Blood Mini kit (50u) (Qiagen NV, Cat. No 51104), containing QIAamp Mini Spin Columns (50), Collection Tubes (2mL) (150), Buffer AL (12mL), Buffer AW1 (19mL), Buffer AW2 (13mL), Buffer AE (15mL), QIAGEN® Protease (1 vial) and Protease Solvent (1.2mL).
- Centrifuge Heraeus Pico 17 (Thermo Fisher Scientific, Waltham, MA, USA, Cat. No. 75002491).
- Heating block (JP Selecta S.A, Abrera, Barcelona, Cat. No. 7462200).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Schwabach, Germany, Cat. No. 541-10000-00).

Procedure

1. Pipet 20µl QIAGEN® Protease into a 1.5mL tube.
2. Add 200µl blood sample to the tube.
3. Add 200µl Buffer AL to the sample. Mix by vortexing for 15 seconds.
4. Incubate at 56°C for 10 minutes. Spin the samples to remove drops from the inside of the lid.
5. Add 200µl ethanol.
6. Vortex for 15 seconds. Spin the samples to remove drops from the inside of the lid.
7. Transfer the mixture to a QIAamp Mini spin column.
8. Centrifuge at 8000 rpm for 1 min. Place the column in a new tube and discard the tube containing the filtrate.
9. Add 500µl Buffer AW1 and centrifuge at 8000 rpm for 1 min. Discard the filtrate.
10. Add 500µl Buffer AW2 and centrifuge at 13000 rpm for 3 minutes. Place the column in a new tube and discard the tube containing the filtrate.
11. Centrifuge at 13000 rpm for 1 min.
12. Place the QIAamp Mini spin column in a clean 1.5mL microcentrifuge tube.
13. Add Buffer AE. For candidate gene sequencing, add 75µL of Buffer AE; however, for NGS analysis add 200µL of Buffer AE.
14. Incubate at room temperature for 5 minutes.
15. Centrifuge at 8000 rpm for 1 min.
16. Transfer the eluted DNA to a 1.5mL screw tube, assign a new DNA number for the laboratory register and store at 4°C or -20°C for posterior use.

7.1.2. Automated DNA extraction

The Magpurix 12S system consists of a robotic workstation for automated nucleic acid purification that uses pre-filled reagent cartridges and disposable consumables, with the flexibility of processing 1-12 samples per run. The MagPurix technology uses magnetic beads to purify nucleic acids from samples and follows steps of lysis, binding, washing and elution. The MagPurix Blood DNA Extraction Kit 200 (36 preps) (Zinexts Life Science Corp., Taiwan) for gDNA (genomic DNA) extraction from 100-400µL mammalian blood cells was used. This method allows for a DNA yield of 4-7µg.

Depending on the genetic approach to be performed, the required DNA concentration varies.

While 25-50ng/µL human genomic DNA is recommended for Polymerase Chain Reaction (PCR), the optimal concentration for NGS is about 100ng/µL. Therefore, for candidate gene sequencing, we isolated and purified 300µL from each blood sample to get 300µL eluted DNA; however, for NGS analysis the elution volume was 100µL.

Reagents and materials

- Magpurix 12S system (Zinexts Life Science Corp., Cat. No. ZP01001).
- MagPurix Blood DNA Extraction Kit 200 (36 preps) (Zinexts Life Science Corp., Cat. No. ZP02001), containing Reaction chamber (36 pieces, 6x6), Tip holder (36 pieces), Filtered tip (38 pieces), Piercing pin (38 pieces), Sample tube of 2mL (38 pieces), Elute tube of 1.5mL (38 pieces), Barcode paper and Reagent cartridge (36 pieces, 6x6). The Reagent cartridge includes:
    - Well 1: Proteinase K solution (40µl).
    - Well 2: Lysis Buffer 2 (1000µl).
    - Well 3: Binding Buffer 1 (600µl).
    - Well 4: Magnetic Bead Solution (800µl).
    - Well 5: Washing Buffer 1 (1000µl).
    - Well 6: Washing Buffer 2 (1000µl).
    - Well 7: Washing Buffer 3 (1000µl).
    - Well 8: Elution Buffer 1 (1000µl).
    - Well 9: Elution Buffer 2 (1000µl).
    - Well 10: Empty.

Procedure

1. Turn the power switch on. Open the sliding door and remove the sample rack from the instrument.
2. Load reagent cartridges and all plastic disposables, such as, reaction chamber, tip holder, piercing pin and filtered tip.
3. Load one reaction cartridge and one set of plastic disposable per sample.
4. Load sample tube and elute tube to sample rack on the bench.
5. Load 300µL of the blood sample to sample tube.
6. Place sample rack on the instrument platform. Close the door.
7. Scan the protocol barcodes to select purification protocol MagPurix Blood DNA Extraction Kit 200, sample volume 300µL and elute volume 300µL or 100µL.
8. Follow the instructions displayed on the screen to check the operating steps.
9. Push “Enter” to confirm and press “Start” button. The automated purification protocol begins and steps of the protocol.
10. At the end of the run, open the instrument door and remove the elute tubes containing the purified DNA.

11. Discard the used cartridges and all plastic consumables into biohazard waste. Close the instrument door and turn the power switch off.
12. Transfer the eluted DNA to a 1.5mL screw tube, assign a new DNA number for the laboratory register and store at 4°C or -20°C for posterior use.

7.2. EXTRACTED DNA QUANTIFICATION AND PURITY CHECKING

To check for concentration the NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific) and the Qubit® 2.0 Fluorometer (Thermo Fisher Scientific) were used. For measuring the quality of the extracted DNA samples the NanoDrop® ND-1000 spectrophotometer was used.

7.2.1. DNA quantification

DNA concentrations can be assessed using different methods, such as absorbance or optical density (Spectrophotometric method) and fluorescent DNA-binding dyes (Fluorometric method).

7.2.1.1. Spectrophotometric method

Nucleic acids absorb ultraviolet (UV) light due to the heterocyclic rings of the nucleotides among the DNA. The wavelength of maximum absorption for DNA is 260nm (λmax = 260nm) with a characteristic value for each base, that will allow the determination of the concentration.

At very high concentrations, especially if the material is scattering, the absorbance value and therefore the concentration can be inaccurate. UV absorbance measurements are not selective and cannot distinguish between single or double strand nucleic acids. Values are easily affected by other contaminants, such as free nucleotides, salts, and organic compounds, and variations in base composition.

Reagents and materials

- Buffer AE (15mL) from the QIAamp DNA Blood Mini kit (Qiagen NV).
- Milli-Q ultrapure water (Merck KGaA, Darmstadt, Germany).
- NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific).
- NanoDrop 1000 Spectrophotometer Software v3.8 (Thermo Fisher Scientific).

Procedure

1. Clean the pedestal by pipetting 2µL bidistilled water onto the measurement pedestal. Wipe both the upper and lower pedestals.

2. Pipet 1µL Milli-Q ultrapure water or AE buffer onto the pedestal to make the blank measurement.
3. Lower the sampling arm and initiate the measurement by clicking on the software.
4. When the measurement is complete, raise the sampling arm and clean.
5. Pipet 1µL sample directly onto the measurement pedestal and measure.
6. The software displays the results of the concentration measurement, as well as the purity checking values.
7. Repeat step 1 once all the samples have been measured.

7.2.1.2. Fluorometric method

The Qubit® 2.0 Fluorometer in combination with the Qubit® dsDNA BR Assay Kit yields specific double-stranded DNA (dsDNA) quantitation and is accurate for initial sample concentrations from 100pg/µL to 1000ng/µL. The basis of the Qubit™ assay is a molecular probe dye that emits fluorescent signals only when bound to dsDNA.

Reagents and materials

- Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Cat. No. Q32866).
- Qubit® dsDNA BR Assay Kit (500 preps) (Thermo Fisher Scientific, Cat. No. Q32853), containing Qubit® dsDNA BR Reagent (Component A) (1,25mL), Qubit® dsDNA BR Buffer (Component B) (250mL), Qubit® dsDNA BR Standard #1 (Component C) (5mL) and Qubit® dsDNA BR Standard #2 (Component D) (5mL).
- Qubit Assay tubes (500 tubes) (Thermo Fisher Scientific, Cat. No. Q32856).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).

Procedure

1. Set up the required number of 0.5mL Qubit® assay tubes for the two standards and samples. Label the tube lids.
2. Prepare the Qubit® working solution by diluting the Qubit® dsDNA BR Reagent 1:200 in Qubit® dsDNA BR Buffer.
3. Add 190µL of Qubit® working solution to each standard tube.
4. Add 10µL of each Qubit® standard to the appropriate tube, then mix by vortexing.
5. Add 198µL of Qubit® working solution to each assay tube.
6. Add 2µL sample to each assay tube, then mix by vortexing.
7. Incubate all tubes at room temperature for 2 minutes.
8. On the Home screen of the Qubit® 2.0 Fluorometer, press DNA and select dsDNA Broad Range.
9. Read the standards. Insert the tube containing Standard #1 into the sample chamber, close the lid, and then press Read. When the reading is complete remove Standard #1 and repeat with Standard #2.
10. When the calibration is complete, the instrument displays the Sample screen.

11. Insert a sample tube into the sample chamber, close the lid, and then press Read. When the reading is complete, remove the sample tube.
12. The instrument displays the results on the Sample screen.
13. Repeat step 12 until all samples have been read.

7.2.2. DNA purity checking

Absorbance measurements will measure any molecules absorbing at a specific wavelength. Nucleotides, Ribonucleic Acid (RNA), single-stranded DNA (ssDNA) and dsDNA will absorb at 260nm and contribute to the total absorbance.
The ratio of absorbance at 260nm and 280nm is used to assess the purity of DNA. A ratio of 1.8 is accepted as pure DNA. If the ratio is lower, it may indicate the presence of protein, phenol or other contaminants that absorb at nearly 280nm.

The 260/230 ratio is commonly in the range of 2.0 for pure DNA. If it is appreciably lower, it may indicate the presence of contaminants which absorb at 230nm, such as EDTA, carbohydrates and phenol.

7.3. POLYMERASE CHAIN REACTION

This widely used technique exponentially amplifies a specific segment of DNA to generate multiple copies of a DNA sequence.

The PCR comprises five steps developed at different temperatures in a thermocycler. The initialization step consists in the heat activation of the Taq polymerase at 94-95°C for 5-10 minutes. In the denaturing stage, the hydrogen bonds between complementary bases are broken at 90-94°C and this breaks the dsDNA into two ssDNA molecules. In the annealing stage, the temperature is lowered to 50-65°C and the primers bind to their complementary sequence of DNA. This forms the double strand site which the polymerase can bind to and allows the elongation of DNA. In the elongation stage, at 72 to 75°C, the DNA polymerase synthesizes a new DNA strand, complementary to the DNA template strand, by adding free dNTPs (Deoxynucleoside triphosphates) in the 5’ to 3’ direction. At each elongation stage, the number of DNA target sequences is doubled. The processes of denaturation, annealing and elongation constitute a single cycle and are repeated among 35 to 40 times, leading to exponential amplification of the specific DNA target region. A final elongation step is performed at 72 to 75°C for 5 to 10 minutes to ensure that any remaining ssDNA is elongated.

Besides DNA template containing the target region to amplify, several components and reagents are required for the development of the technique:

- Buffer solution: It provides the suitable environment for the activity and stability of the DNA polymerase. It contains Tris-HCl to maintain the proper pH at 8-8.3, KCl to increase the activity of the enzyme and $MgCl_2$ which provides the Mg^{2+} ions that work as cofactors of the polymerase.
- dNTPs: The main subject from which the new strand is synthesized by the polymerase. Adenine, guanine, cytosine and thymine deoxynucleosides are added in the same concentration.
- Two DNA primers: Specific short ssDNA fragments complementary to the 3' ends of each of the sense and antisense strands of the target DNA. We custom-made the oligonucleotides for each region of interest and add a M13 tail (Forward: 5'-TGTAAAACGACGGCCAGT-3' and reverse: 5'-CAGGAAACAGCTATGACC-3'). With this M13 tailed primers we can use the same primers for sequencing, ignoring the original sequence.
- Dimethyl sulfoxide (DMSO): An organosulfur compound that binds the cytoside residue, changing its conformations ant therefore, makes DNA more labile for heat denaturation.
- Heat-stable DNA polymerase (Taq polymerase): Originally isolated from *Thermus aquaticus* bacteria. This enzyme polymerizes new DNA strands and it remains intact during the high temperature DNA denaturation process. We used the KAPA Taq DNA Polymerase (Kapa Biosystems, Boston, MA, USA), of the thermophilic bacterium *Thermus aquaticus*, purified from recombinant *Escherichia coli* (*E. coli)*. It has 5'-3' polymerase and 5'-3' exonuclease activity.
- Milli-Q ultrapure water (Merck KGaA, Darmstadt, Germany).

We used PCR for the amplification of the genes of interest by the candidate gene approach, validation of variants found by NGS and amplification of regions not covered by the targeted gene panel.

Reagents and materials

- KAPA Taq PCR Kit (Kapa Biosystems, Cat. No. KK1014), containing KAPA Taq DNA Polymerase (5U/µL), KAPA Taq Buffer A (10X) KAPA Taq Buffer B (10X) and $MgCl_2$ (25mM).
- dNTP Mix (10mM) (5mL) (Merck KGaA, Cat. No. D7295-5ML).
- DMSO (C_2H_6OS) (50mL) (Merck KGaA, Cat. No. D8418-50ML).
- Primers pool (250 µM) (Integrated DNA Technologies Inc, Skokie, Il, USA).
- Milli-Q ultrapure water (Merck KGaA).
- PCR-strip tubes (Sarstedt AG&Co KG, Nümbrecht, Germany, Cat. No 72.982.002).
- PCR lid strips (Sarstedt AG&Co KG, Cat. No. 65.989.002).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).

Procedure

For each amplification, a mix containing Milli-Q water, Buffer A, dNTPs (10mM), primers at 25pM, Taq DNA polymerase and DNA (25-100ng) is prepared. DMSO is added when required. Conditions in the thermocycler are also different for each DNA amplification sequence. Annealing temperature of the different primers was optimized by doing a temperature gradient PCR.

The conditions for each amplification and primers can be found in Supplementary data 6.

7.4. AGAROSE GEL ELECTROPHORESIS

Agarose gel electrophoresis is the most common way of separating DNA fragments of varying sizes for visualization and purification. Nucleic acid molecules are separated by applying an electric field to move through a matrix of agarose, in which the pore size is determined by the percentage of agarose in the buffer. When applying an electric field, negatively charged DNA, due to the phosphate backbone, will migrate to the positively charged anode in neutral pH conditions given by 1X TAE (Tris base, acetic acid and EDTA) buffer.

Shorter DNA fragments migrate through the gel more quickly than longer ones. Consequently, the approximate length of a DNA fragment can be determined by running it on an agarose gel alongside a DNA ladder. After separation, the DNA molecules can be visualized under UV light after staining with the fluorophore GelRed™ (Biogen, Cambridge, MA, USA), an intercalating nucleic acid dye that when exposed to UV light will fluoresce with an orange colour after binding to DNA. Comparison of the band location with the molecular ruler confirms the specific amplification of the region of interest

Reagents and materials

- Agarose D-1 (Pronadisa, Madrid, Spain, Cat. No. 8016).
- 50X TAE Buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA, pH 8,3) (Bio-Rad, Hercules, CA, USA, Cat. No. 1610773).
- EZ Load™ 100bp Molecular Ruler (Bio-Rad, Cat. No. 1708352).
- 5X Nucleic Acid Sample Loading Buffer (Bio-Rad, Cat. No. 161-0767).
- GelRed™ (Biogen, Cambridge, MA, USA, Cat. No. BT41003).
- Milli-Q ultrapure water (Merck KGaA).
- Electrophoresis chamber Sub® Cell GT MINI (Bio-Rad, Hercules, CA, USA, Cat. No. 1704467).
- PowerPac™ 3000 Power Supply (Bio-Rad, Cat. No. 1655057).
- G: BOX Chemi 16 Bio Imaging system (Syngene, Cambridge, UK, Cat. No. SGBOX).
- GeneSnap Software v7.12 (Syngene).

- Sub-Cell GT UV-Transparent Gel Tray (Bio-Rad, Cat. No. 1704435).
- Fixed-Height Comb (Bio-Rad, Cat. No. 1704465).

Procedure

1. Melt the 1.5% agarose containing bottle (Supplementary data 7) and pour 50mL to a falcon. Add 2µL of GelRed™.
2. Pour the mixture into a gel tray.
3. Put the Fixed-Height Comb and cool.
4. Remove the Fixed-Height Comb and put the gel into the electrophoresis chamber. Add 1X TAE buffer (Supplementary data 7) until the gel is covered.
5. Mix 2µL of DNA products with 1µL of 5X Nucleic Acid Sample Loading Buffer.
6. Load 1µL of the EZ Load™ 100bp Molecular Ruler on the first well.
7. Continue loading 2µL of the DNA products mixed with loading buffer.
8. Run the electrophoresis at 110V for 30 minutes.
9. Visualize the PCR product in the imaging system and capture the image by using the GeneSnap Software v7.12.

7.5. EXOSAP-IT PURIFICATION

When the PCR amplification is complete, unconsumed dNTPs and primers remain in the PCR product mixture. ExoSAP-IT™ (Thermo Fisher Scientific) prepares PCR products for sequencing. It employs two hydrolytic enzymes, Exonuclease I and Shrimp Alkaline Phosphatase (SAP), to remove the remaining dNTPs and primers that would interfere with the sequencing reaction. Exonuclease I degrades residual primers and single stranded DNA while SAP hydrolyses the remaining dNTPs from the mixture.

Reagents and materials

- ExoSAP-IT™ (Thermo Fisher Scientific, Cat. No. 78202).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).

Procedure

1. Add 5µL of the PCR product to a 0.2mL tube.
2. Add 2µL ExoSAP-IT and mix by vortexing.
3. Incubate at the thermocycler at 37°C for 15 minutes and at 80°C for another 15 minutes.
   Store the samples at 4°C until use.

7.6. CAPILLARY SANGER SEQUENCING

The principles of DNA replication were used by Sanger *et al* (228) in the development of the process known as Sanger dideoxy sequencing. This process takes advantage of the ability of DNA polymerase to incorporate 2′, 3′-dideoxynucleotides, base analogues that lack the 3′-hydroxyl group essential in phosphodiester bond formation.

Sequencing requires a DNA template, a sequencing primer, a thermal stable DNA polymerase, dNTPs, dideoxynucleotides triphosphates (ddNTPs), and reaction buffer containing manganese (Mn^{2+}). Unlike Sanger's method, which used radioactive material, fluorescence-based cycle sequencing uses fluorescent dyes to label the extension products and the components are combined in a reaction that is subjected to cycles of annealing, extension, and denaturation in a thermal cycler. Thermal cycling of the sequencing reactions creates and amplifies extension products that are terminated (dye terminators) by one of the four ddNTPs. Because each dye label emits a unique wavelength when excited by light, the fluorescent dye on the extension product identifies the 3′ terminal ddNTP as A (Adenine), C (Cytosine), G (Guanine) or T (Thymine). The ratio of dNTPs to ddNTPs is optimized to produce a balanced population of long and short extension products.

During capillary electrophoresis, the extension products of the cycle sequencing reaction enter the capillary as a result of electrokinetic injection. A high voltage charge applied to the buffered sequencing reaction forces the negatively charged fragments into the capillaries. The extension products are separated by size based on their total charge. Before reaching the positive electrode, the fluorescently labelled DNA fragments move across the path of a laser beam, causing the dyes on the fragments to fluoresce. An optical detection device on the genetic analyser detects the fluorescence and the software converts the fluorescence signal to digital data.

7.6.1. Cycle sequencing DNA templates

Reagents and materials

- BigDye® Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific, Cat. No. 4337455), which includes 5X Sequencing Buffer (2 × 1mL) and BigDye® Terminator v3.1 Ready Reaction Mix (1 × 800µl).
  5X Sequencing Buffer, includes: Tris-HCl and $MgCl_2$.
  BigDye® Terminator v3.1 Ready Reaction Mix, contains: dideoxynucleotide triphosphates (ddATP, ddTTP, ddCTP, ddGTP), deoxynucleotide triphosphates (dATP, dTTP, dCTP, dGTP) and Taq polymerase enzyme.
- DMSO (50mL) (Merck KGaA, Cat. No. D8418-50ML).
- Specific primers for each fragment (2.5µM) or M13-tailed primers (2.5µM) (Integrated DNA Technologies Inc).
- Milli-Q ultrapure water (Merck KGaA).

- MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher Scientific, Cat. No. 4316813).
- MicroAmp™ Clear Adhesive Film (Thermo Fisher Scientific, Cat. No. 4306311).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).

Procedure

1. For each reaction add the following reagents to the MicroAmp® Optical 96-Well Reaction Plate:
    - BigDye® Terminator v3.1 Ready Reaction Mix: 1µL
    - 5X Sequencing Buffer: 1.5µL
    - Milli-Q ultrapure water: 3.5µL
    - DMSO: 0.5µl
2. Mix using a Vortex
3. Add 2.9µl purified PCR product and 1.3µL corresponding primers.
4. Mix using a Vortex.
5. Cover the plate with MicroAmp™ Clear Adhesive Film.
6. Place the plate in the GeneAmp® PCR System 9700 thermocycler and perform cycle sequencing as follows:
    - Initial denaturation: 94°C for 3 minutes
    - Repeat the following for 25 cycles:
        - Denaturation: 96°C for 10 seconds.
        - Annealing: 50°C for 5 seconds.
        - Elongation: 60°C for 4 minutes
    - Hold at 4°C until purification

7.6.2. Purification of extension products

The presence of both unlabelled and dye-labelled reaction components can interfere with electrokinetic injection, electrophoretic separation, and data analysis. Purification of extension products reduces this interference.

Alcohol-based nucleic acid precipitation techniques include a wide variety of methods, such as the choice of alcohol, concentration of salt, temperature and additives. Purification by Ethanol/EDTA/NaOAc was performed following the protocol by Applied Biosystems.

Reagents and materials

- EDTA ($C_{10}H_{16}N_2O_8$) (Merck KGaA, Cat. No. 431788).
- Sodium acetate (CH_3COONa) (NaOAc) (Merck KGaA, (Cat. No. S2889).
- Absolute ethanol (Thermo Fisher Scientific, Cat. No. BP2818100).

- Formamide (CH_3NO) (Thermo Fisher Scientific, Cat. No. 4311320).
- Milli-Q ultrapure water (Merck KGaA).
- Towel paper Scottfold (Kimberly-Clark, Cat. No. MRT116, Barcelona, Spain).
- Beckman Spinchron R centrifuge (Beckman Coulter, Cat. No. 358723, Brea, California, USA).

Procedure

1. Remove the 96-well reaction plate from the thermal cycler and centrifuge the plate at 100 x g for 1 minute. Then remove the seal.
2. Prepare 125mM EDTA and 3M NaOAc (Supplementary data 7)
3. Mix enough volume of 125mM EDTA and 3M NaOAc in 1:1 proportion. Add 2µL of the mix to each sample.
4. Add 25µL absolute ethanol to each well and mix by pipetting up and down.
5. Incubate the plate in darkness at room temperature for 15 minutes
6. Centrifuge the plate at 3000 x g at 4°C for 30 minutes.
7. Invert the plate onto a paper towel and spin the plate in the centrifuge.
8. Add 35µL 70% ethanol to each well.
9. Centrifuge the plate at 1650 x g at 4°C for 15 minutes.
10. Invert the plate onto a paper towel and centrifuge the plate at 190 x g for 4 minutes.
11. Add 12µL Milli-Q water.
12. Add 12µL formamide
13. Seal the plate with self-adhesive film, mix by vortexing.
14. Remove the self-adhesive film. Cover the plate with the specific lid and place the reaction plate in the genetic analyser.

7.6.3. Capillary electrophoresis and data analysis

Capillary electrophoresis was performed on an ABI PRISM 3130xl (Thermo Fisher Scientific) Genetic Analyser, following manufacturer's protocol.

Obtained chromatograms were analysed using Sequencing Analysis Software v5.2 (Thermo Fisher Scientific) and SeqScape® Software v3.1 (Thermo Fisher Scientific).

Reagents and materials

- 10X Genetic Analyser Running Buffer (Thermo Fisher Scientific, Cat. No. 402824).
- POP-7™ Polymer (Thermo Fisher Scientific, Cat. No. 4352759).
- ABI PRISM 3130xl Genetic Analyser (Thermo Fisher Scientific, Cat. No. 4359571).
- Sequencing Analysis Software v5.2 (Thermo Fisher Scientific).
- SeqScape® Software v3.1 (Thermo Fisher Scientific).

7.7. TARGETED GENE SEQUENCING PANEL

Next generation sequencing provides a high-throughput screening of multiple genes in a highly efficient manner. Although different platforms have been developed for the panel sequencing, they all include the following steps (Figure 10):

1. Design of the panel
2. Library preparation
3. Preparation of the template and enrichment
4. Sequencing
5. Data analysis

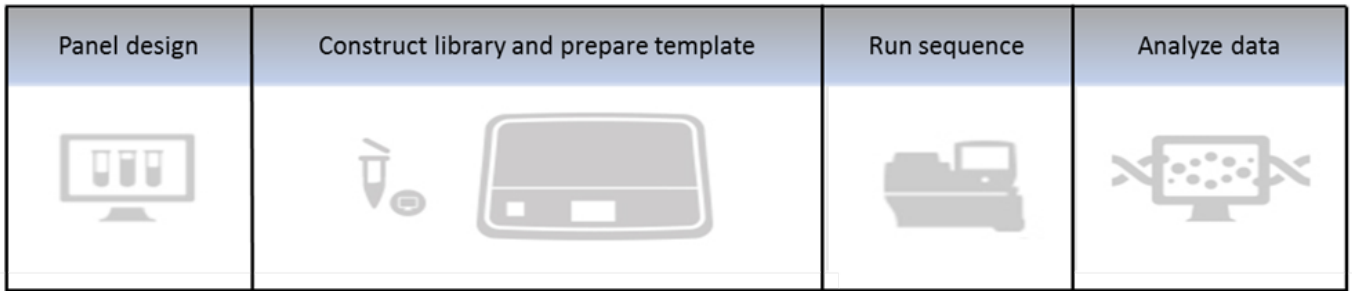

Figure 10. Diagram of the main steps involved in the targeted gene panel sequencing. Modified from Thermo Fisher Scientific (https://www.thermofisher.com/blog/).

We performed the targeted gene sequencing panel with Ion Torrent technology (Thermo Fisher Scientific).

7.7.1. Design of the panel

After the implementation of NGS technology in our Institute, patients collected since 2015 were analysed with the customized gene panel for DSD, unless the suspicion of the causative gene was clear and the methodology to be used was aCGH or MLPA.

We designed a targeted panel to sequence 48 genes associated with sex determination, sex differentiation and hypogonadism. Included genes were chosen after search in online databases and in the available literature, including HGMD (Human Gene Mutation database) (https://portal.biobase-international.com/hgmd/pro/start.php?), OMIM (Online Mendelian Inheritance in Man) (https://www.omim.org/) and Pubmed (https://www.ncbi.nlm.nih.gov/pubmed/). The list of the included genes in the customized panel as well as their associated role and phenotype can be found in Table 4.

Table 4. Included genes in the customized panel and their associated role and phenotype in DSD.

Gene (*locus*)	Alias	Transcript	Inheritance (OMIM)	Role in DSD	Associated phenotype
AMH (19p13.3)	Anti-Mullerian Hormone	NM_000479.3	AR (600957)	Sex Diff	PMDS, POI
AMHR2 (12q13.13)	Anti-Mullerian Hormone Receptor Type 2	NM_020547.2	AR (600956)	Sex Diff	PMDS, POI
AR (Xq12)	Androgen Receptor	NM_000044.3	X-linked (313700)	Sex Diff	46,XY DSD AIS, hypospadias, POI
ATRX (Xq21.1)	ATP-Dependent Helicase ATRX	NM_000489.4	X-inked (300032)	G Diff	46,XY DSD assoc with Alpha thalassemia X-linked intellectual disability syndr
BMP15 (Xp11.22)	Bone Morphogenetic Protein 15	NM_005448.2	X-linked (300247)	G Dev	POI w/wot primary amenorrhea
CBX2 (17q25.3)	Chromobox homolog 2	NM_005189.2	AR, AD (602770)	G Dev	46,XY DSD GD
CYP11A1 (15q24.1)	Cytochrome P450 Family 11 Subfamily A Member 1	NM_000781.2	AR, AD (118485)	Sex Diff	46,XY DSD CGD and adrenal insuf; Hypospadias and adrenal insuf
CYP11B1 (8q24.3)	Cytochrome P450 Family 11 Subfamily B Member 1	NM_000497.3	AR (610613)	Sex Diff	46,XX DSD, Steroid-11 beta-hydroxylase def, Non-classic Steroid-11 beta-hydroxylase def, CAH
CYP17A1 (10q24.32)	Cytochrome P450 Family 17 Subfamily A Member 1	NM_000102.3	AR (609300)	Sex Diff	46,XY DSD, 17-alpha-hydroxylase/17,20-lyase def, CAH
CYP19A1 (15q21.2)	Cytochrome P450 Family 19 Subfamily A Member 1	NM_000103.3	AR (107910)	Sex Diff	46,XY DSD aromatase def
CYP21A2 (6p21.33)	Cytochrome P450 Family 21 Subfamily A Member 2	NM_000500.7	AR (613815)	Sex Diff	21-hydroxylase def w/wo salt wasting; simple virilizing form
DHH (12q13.12)	Desert Hedgehog	NM_021044.2	AD, AR (605423)	G Dev	46,XY DSD GD
DMRT1 (9p24.3)	Doublesex And Mab-3 Related Transcription Factor 1	NM_021951.2	AD (602424)	G Dev	46,XY DSD GD, OT DSD
DMRT2 (9p24.3)	Doublesex And Mab-3 Related Transcription Factor 2	NM_181872.4	AD (604935)	G Dev	46,XY DSD GD
ESR1 (6q25.1-q25.2)	Oestrogen Receptor 1, Nuclear Receptor Subfamily 3 Group A Member 1	NM_001122740.1	AD, AR (133430)	G Dev	46,XY DSD, male infertility, PP, MRKS
ESR2 (14q23.2-q23.3)	Oestrogen Receptor 2, Nuclear Receptor Subfamily 3 Group A Member 2	NM_001437.2	AD, AR (601663)	G Dev	46,XY DSD GD, Hypospadias, primary amenorrea, 46,XX DSD GD
FGF9 (13q12.11)	Fibroblast Growth Factor 9	NM_002010.2	AD (600921)	G Dev	46,XX T DSD
FOXL2 (3q22.3)	Forkhead Box L2	NM_023067.3	AD (605597)	G Dev	POI with and without BPES
FOXO3 (6q21)	Forkhead Box O3	NM_001455.3	AD (602681)	G Dev	POI
FSHR (2p16.3)	Follicle Stimulating Hormone Receptor	NM_000145.3	AR (136435)	CHH	POI, hypergonadotropic hypogonadism
GATA4 (8p23.1)	GATA Binding Protein 4	NM_002052.3	AD (600576)	G Dev	46,XY DSD GD w/wo CHD

Table 4. Included genes in the customized panel and their associated role and phenotype in DSD (Continuation).

Gene (*locus*)	Alias	Transcript	Inheritance (OMIM)	Role in DSD	Associated phenotype
HARS2 (5q31.3)	Histidyl-TRNA Synthetase 2, Mitochondrial	NM_012208.3	AR (600783)	Sex Diff	Perrault syndr, including 46,XX DSD with ovarian dysgenesis
*HSD17B3 (*9q22.32)	Hydroxysteroid 17-Beta Dehydrogenase 3	NM_000197.1	AR (605573)	Sex Diff	46,XY DSD 17β-hydroxysteroid dehydrogenase 3 deficiency, hypospadias
HSD17B4 (5q23.1)	Hydroxysteroid 17-Beta Dehydrogenase 4	NM_000414.3	AR (601860)	Sex Diff	Perrault syndr, including 46,XX DSD with ovarian dysgenesis, and POI
HSD3B2 (1p12)	Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta and Steroid Delta-Isomerase 2	NM_000198.3	AR, AD (613890)	Sex Diff	46,XY DSD GD, CAH with 3β-hydroxysteroid dehydrogenase type 2def, Hypospadias
INHA (2q35)	Inhibin Subunit Alpha	NM_002191.3	AD (147380)	G Dev	POI, male infertility
INSL3 (19p13.11)	Insulin Like 3	NM_001265587.1	AD (146738)	Others	Cryptorchidism
KISS1 (1q32.1)	Kisspeptin-1	NM_002256.3	AR, AD (603286)	CHH	Puberty delay, HH
KISS1R (19p13.3)	KISS1 Receptor	NM_032551.4	AD (604161)	CHH	Puberty delay, HH
LHCGR (2p16.3)	Luteinizing Hormone/Choriogonadotropin Receptor	NM_000233.3	AR, AD (152790)	Sex Diff	PP, 46,XY DSD Leydig cell hypoplasia, amenorrhea
MAMLD1 (Xq28)	Mastermind Like Domain Containing 1, CXorf6	NM_001177465.2	X-linked (300120)	G Dev	46,XY DSD, 46,XX DSD, Hypospadias
MAP3K1 (5q11.2)	Mitogen-Activated Protein Kinase Kinase Kinase 1	NM_005921.1	AD (600982)	G Dev	46,XY DSD GD, hypospadias
NR0B1 (Xp21.2)	Nuclear Receptor Subfamily 0 Group B Member 1	NM_000475.4	X-linked (300473)	G Dev	Adrenal hypoplasia & HH, 46,XY DSD GD, hypospadias, 46,XX OT DSD
NR5A1 (9q33.3)	Nuclear Receptor Subfamily 5 Group A Member 1	NM_004959.4	AD, AR (184757)	G Dev	46,XY DSD GD, POI, 46,XX T and OT DSD, hypospadias
POR (7q11.23)	Cytochrome P450 Oxidoreductase	NM_000941.2	AR (124015)	Sex Diff	Hypospadias, Disordered steroidogenesis, Antley-Bixler syndr with genital anomalies & disordered steroidogenesis
PSMC3IP (17q21.2)	Proteasome 26S ATPase Subunit 3-Interacting Protein	NM_016556.3	AR (608665)	G Dev	POI
RSPO1 (1p34.3)	R-Spondin 1	NM_001242908.1	AR (609595)	G Dev	46,XX OT DSD with palmoplantar hyperkeratosis with squamous cell carcinoma of skin
RXFP2 (13q13.1)	Relaxin Family Peptide Receptor 2, GREAT, LGR8	NM_130806.3	AD (606655)	Others	Cryptorchidism

Table 4. Included genes in the customized panel and their associated role and phenotype in DSD (Continuation).

Gene (*locus*)	Alias	Transcript	Inheritance (OMIM)	Role in DSD	Associated phenotype
SOX3 (Xq27.1)	SRY (Sex Determining Region Y)-Box 3	NM_005634.2	X-linked (313430)	G Dev	46,XX OT and T DSD w/wo growth retardation and microcephaly
SOX9 (17q24.3)	SRY (Sex Determining Region Y)-Box 9	NM_000346.3	AD (608160)	G Dev	46,XY DSD GD, 46,XX TDSD, 46, XX OT DSD, 46,XY DSD with campomelic dysplasia
SRD5A2 (2p23.1)	Steroid 5 Alpha-Reductase 2	NM_000348.3	AR (607306)	Sex Diff	46,XY DSD Steroid 5-alpha-reductase def, Hypospadias
SRY (Yp11.2)	Sex Determining Region Y	NM_003140.2	Translocation,AD (480000)	G Dev	46,XX OT and Testicular DSD, 46,XY DSD GD
STAR (8p11.23)	Steroidogenic Acute Regulator	NM_000349.2	AR (600617)	Sex Diff	46,XY DSD, 46,XY DSD and LCAH, 46,XX DSD
TSPYL1 (6q22.1)	Testis-Specific Y-Encoded-Like Protein 1	NM_003309.3	AR (604714)	G Dev	46,XY DSD GD, 46,XY DSD and Sudden infant death
WNT4 (1p36.12)	Wingless-Type MMTV Integration Site Family, Member 4	NM_030761.4	AD (603490)	G Dev	46,XY DSD CGD, 46,XX OT and T DSD, 46,XX MRKS
WT1 (11p13)	Wilms Tumour 1	NM_024426.4	AD (607102)	G Dev	46,XY DSD GD, Frasier syndr, Denys-Drash syndr, hypospadias
WWOX (16q23.1-q23.2)	WW Domain Containing Oxidoreductase	NM_016373.2	AD (605131)	G Dev	46,XY DSD GD, 46,XX DSD GD
ZFPM2 (8q23.1)	Zinc Finger Protein, FOG Family Member 2, Friend Of GATA 2	NM_012082.3	AD (603693)	G Dev	46,XY DSD GD, hypospadias

AIS, Androgen insensitivity Syndrome; AD, Autosomal dominant; AR, Autosomal recessive; Assoc, associated; BPES, Blepharophimosis, ptosis epicanthus inversus syndrome; CAH, Congenital Adrenal Hyperplasia; CGD, Complete Gonadal dysgenesis; CHD, congenital heart defects; CHH, central causes of hypogonadism; Def, deficiency; Diff, differentiation; GD, gonadal dysgenesis; G Dev, gonadal development; HH, Hypogonadotropic hypogonadism; Insuf, insufficiency; LCAH, lipoid CAH; MRKS, Mayer-Rokitansky-Küster-Hauser syndrome; OMIM, Online Mendelian Inheritance on Man; OT, Ovotesticular; PGD, Partial Gonadal dysgenesis; PMDS, Persistent Müllerian Duct Syndrome; POI, Primary Ovarian Insufficiency; PP, precocious puberty; Syndr, syndrome; T, testicular; w/wo, with and without.

The customized gene panel was designed with the Ion AmpliSeq Designer tool v4.4.8 (https://www.ampliseq.com/browse.action) (Thermo Fisher Scientific). Oligonucleotides against the entire coding region, exon-intron boundaries (± 50 base pair, bp) and the UTR (Untranslated Region) regions were included. Performance of the customized panel at any given *locus* is highly dependent on the sequencing of the amplicons that span the location. The panel designer tool automatically selected the amplicons to be amplified in each gene. In order to increase the coverage of certain genes, we chose the location of the amplicons by writing the locus of interest. We selected the region of interest of the following genes: *AMH, CYP11B1, CYP21A2, FOXO3, GATA4, MAP3K1* and *POR*. The number of amplicons used to sequence each gene, missed base-pair and coverage is shown in Supplementary data 8.

The final panel spanned a 233.46kb (kilobase) region with an expected coverage of 98.55% for the targeted region (Table 5). The targeted regions that are not covered according to the Ion Ampliseq Designer tool are shown in Supplementary data 9. The final panel comprised 874 amplicons, from 125 to 375bp in size, divided in two pools of primers (441 and 433 amplicons per pool).

To demonstrate the analytical performance of the test, the targeted gene panel was validated by measuring sensitivity, specificity plus false positive rate, and repeatability.

7.7.2. Library preparation

The preparation of the library consists in the selection and multiplex amplification of a targeted region. We used the Ion AmpliSeq™ Library Kit 2.0 (Thermo Fisher Scientific) to prepare template genomic DNA for sequencing.

We used the protocol written in the “Ion AmpliSeq™ Library Kit 2.0 User Guide Targeted DNA and RNA Library Preparation for use with: Ion AmpliSeq™ Library Kit 2.0 Catalogue Numbers 4475345, 4480441, 4480442, A31133, A31136, A29751, 4482298, 4479790 Publication Number MAN0006735 Revision E.0”.

Table 5. Coverage (%) of the DSD-associated gene panel according to the Ion Ampliseq Designer tool. Target and missed regions are indicated in base pairs (bp) for each gene.

Genes	Target (bp)	Missed (bp)	Coverage (%)	Genes	Target (bp)	Missed (bp)	Coverage (%)
AMH	1830	466	74.5	*HSD3B2*	2099	0	100
AMHR2	2413	0	100	*INHA*	1531	0	100
AR	11288	0	100	*INSL3*	1050	0	100
ATRX	12970	45	99.6	*KISS1*	864	0	100
BMP15	1362	0	100	*KISS1R*	1875	33	98.2
CBX2	5246	11	99.8	*LHCGR*	3643	27	99.3
CYP11A1	2928	0	100	*MAMLD1*	5260	22	99.3
CYP11B1	2716	171	93.7	*MAP3K1*	7001	46	99.3
CYP17A1	2270	0	100	*NR0B1*	1685	0	100
CYP19A1	5064	0	100	*NR5A1*	3445	28	99.2
CYP21A2	2385	148	93.3	*POR*	3311	100	97
DHH	2084	0	100	*PSMC3IP*	2380	0	100
DMRT1	2472	3	99.9	*RSPO1*	3539	4	99.9
DMRT2	3135	19	99.4	*RXFP2*	3703	0	100
ESR1	7568	0	100	*SOX3*	2124	0	100
ESR2	4519	18	99.6	*SOX9*	4084	0	100
FGF9	4680	155	96.7	*SRD5A2*	2695	0	100
FOXL2	2967	104	96.5	*SRY*	3864	0	100
FOXO3	7338	435	94.1	*STAR*	3045	32	98.9
FSHR	3274	0	100	*TSPYL1*	5309	86	98.4
GATA4	3436	422	87.7	*WNT4*	4155	182	95.6
HARS2	3260	0	100	*WT1*	3864	0	100
HSD17B3	1684	0	100	*WWOX*	3373	0	100
HSD17B4	4188	0	100	*ZFPM2*	4880	0	100

Library preparation is divided into the following steps (Figure 11):

1. Amplification of gDNA targets by multiplex PCR: The Ion Ampliseq™ HiFi Mix in conjunction with the pools of primers was used to prepare the amplification of the selected regions by PCR. For each pool of primers we used 15ng of gDNA of the patient.
2. Partial digestion of primer sequences: The resulting DNA amplicons were treated with FuPa reagent, which partially removes PCR primers and repairs the endings of the fragments.
3. Ligate barcode adapters and purify: We ligated adapters and assigned a unique barcode to each library to identify the sample. Ligation products were purified using Agencourt AMPure XP beads.
4. Quantification by quantitative PCR (qPCR): The qPCR is a relative quantitation assay, in which the concentration of each library is calculated in relation to the values given in the DH10 dilutions in the standard curve. FAM™ dye is used as the probe reporter and ROX™ as the passive reference dye. To determine the concentration of each library we used the Ion Library TaqMan™ Quantitation kit. Each sample, standards and negative control were analysed in duplicates.

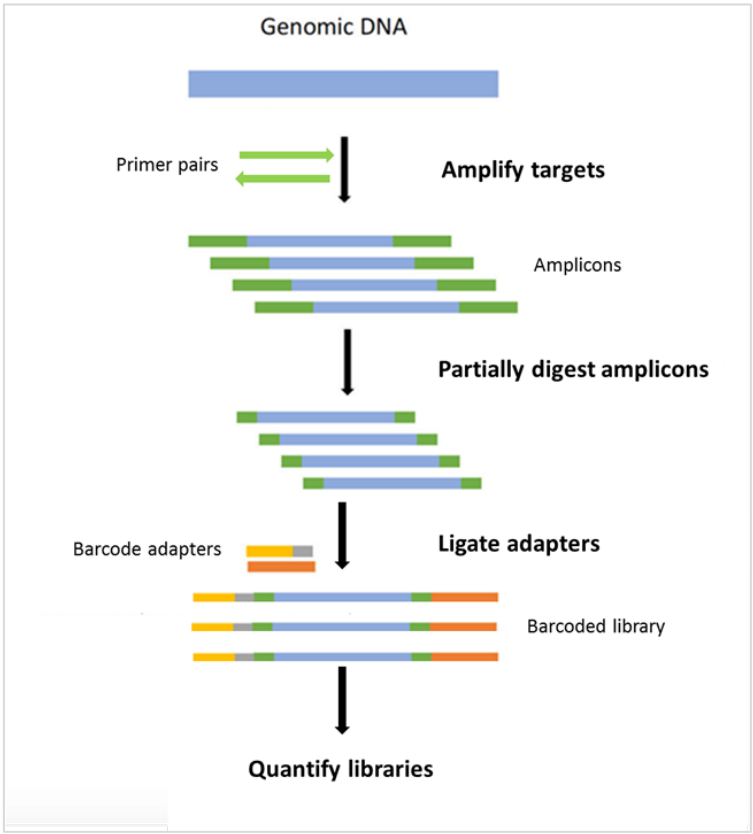

Figure 11. Diagram of the steps followed for library preparation. Modified from Thermo Fisher Scientific (https://www.thermofisher.com/blog/).

Reagents and materials

- Ion Ampliseq™ Library kit 2.0 – 96LV (Thermo Fisher Scientific, Cat. No. 4480441), which contains 5X Ion AmpliSeq™ HiFi Mix (1 x 384µL), FuPa Reagent (1 x 192µL), Switch solution (1 x 384µL), DNA Ligase (1 x 192µL), Ion Ampliseq Adapters (1 x 192µL), Platinum® PCR SuperMix HiFi (1 x 1.6mL), Library Amplification Primer Mix (1 x 192µL) and Low TE (1 x 12mL).
- 2X Ion Ampliseq™ Primer Pool (100nM) (Thermo Fisher Scientific).
- Ion Xpress™ Barcode Adapters 1-16 kits ((Thermo Fisher Scientific, Cat. No. 4471250), containing Ion Xpress™ Barcodes (16 tubes x 20µL) and Ion Xpress™ P1 Adapter (1 x 320µL).
- Agencourt AMPure® XP reagent (60mL) (Beckman Coulter, Cat. No. A63881).
- Ion Library TaqMan™ Quantitation kit (Thermo Fisher Scientific, Cat no. 4468802), containing 2X Ion Library TaqMan® qPCR Mix, 20X Ion Library TaqMan® Quantitation Assay and *E. coli* DH10B Control Library (68pM) (2 x 25µL).
- Absolute Ethanol (Thermo Fisher Scientific, Cat. No. BP2818100).
- DynaMag™ -2 magnet (Thermo Fisher Scientific, Cat. no. 12321D).
- DNA LoBind tubes (1.5mL) (Eppendorf AG, Hamburg, Germany, Cat. No. 022431021).
- Milli-Q ultrapure water (Merck KGaA).
- Veriti 96-well Thermal Cycler (Thermo Fisher Scientific, Cat. No. 4375786).
- Applied Biosystems 7300 Real-Time PCR System (Thermo Fisher Scientific).
- Corning™ PCR Microplates (Thermo Fisher Scientific, Cat. No. 12799438).
- Microseal® 'B' PCR Plate Sealing Film (Bio-Rad, Cat. No. MSA5001).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).

Procedure

1. Amplification of gDNA by multiplex PCR
    - Prepare dilution of input gDNA (5ng/µL) for each primer pool.
    - Prepare Mix. For each pool of primers and gDNA sample, add in a tube:
        - 5X Ion AmpliSeq™ HiFi Mix: 2µL
        - 2X Ion Ampliseq™ Primer Pool: 5µL
        - gDNA: 3µL
    - Mix using a vortex.
    - Place the plate in the Veriti 96-well Thermal Cycler and run the following program to amplify target genomic regions:
        - Activate enzyme: 99°C for 2 minutes.
        - Cycle 16 times: 99°C for 15 seconds and 60°C for 5 minutes.
        - Hold at 10°C.

    At this point PCR products could be stored at 10°C overnight or at -20°C for longer periods.

2. Digestion of primer sequences
    - Combine amplified libraries of each sample in one tube.
    - Add to each tube 2µL FuPa reagent and mix by pipetting.
    - Place the plate in the Veriti 96-well Thermal Cycler and run the following program to digest primer sequences:
        - 50°C for 10 minutes.
        - 55°C for 10 minutes.
        - 60°C for 20 minutes.
        - Hold at 10°C for up to 1 hour.
3. Ligate adapters and purify
    - Prepare diluted barcode adapter mix. Dilute Ion Xpress™ Barcodes and Ion Xpress™ P1 Adapter with nuclease-free water (1:4). Vortex
    - Vortex Switch solution and add 4µL to each tube.
    - Add to ach tube:
        - Diluted barcode adapter mix: 2µL.
        - DNA Ligase: 2µL.
    - Mix using a Vortex.
    - Place the plate in the Veriti 96-well Thermal Cycler and run the following program to make ligation reaction:
        - 22°C for 30 minutes.
        - 72°C for 10 minutes.
        - Hold at 10°C for up to 1 hour.

    At this point PCR products could be stored at -20°C until purification.
    - Bring Agencourt AMPure® XP reagent to room temperature. Vortex thoroughly to disperse the beads before use.
    - Prepare 70% ethanol dilution.
    - Transfer each sample to a Low Retention tube.
    - Add 45µL Agencourt AMPure® XP reagent to each tube. Incubate the mixture for 5 minutes at room temperature.
    - Place the tubes in the DynaMag™ -2 magnet and incubate until solution clears. Without disturbing the pellet, remove and discard the supernatant.
    - Add 150µL 70% ethanol and move the tube side-to-side to wash the beads. Remove and discard the supernatant.
    - Repeat previous step.
    - Keeping the tubes in the magnet, air-dry the beads at room temperature for 2 minutes. If necessary, remove all the ethanol droplets with the pipette.
    - Remove the tubes from the magnet and elute the library by adding 50µL Low TE to the pellet and vortex.
    - Incubate for 5 minutes at room temperature.

- Place the tubes in the magnetic rack until solution clears. Remove and transfer the supernatant to new Low Binding tubes. Discard the tubes with the pellet.

At this point unamplified libraries could be stored at -20°C until quantification.

4. Quantification by qPCR
    - Prepare 1/100 dilutions of each unamplified library. In a new 0.5mL tube add 2µL of each library to 198µL of nuclease-free water.
    - Prepare three 10-fold serial dilutions of *E. coli* DH10B Control Library (68pM) for the standard curve (6.8pM, 0.68pM and 0.068pM).
    - Prepare the reaction mix. For each library add in a tube:
        - 2X Ion Library TaqMan® qPCR Mix: 20µL
        - 20X Ion Library TaqMan® Quantitation Assay: 2µL
    - Dispense 11µL of the mix into the well of the PCR plate for each sample in duplicate.
    - Add 9µL of the diluted libraries, standards or Milli-Q water as a negative control to the corresponding well for a total volume of 20µL.
    - Program the real-time instrument as manufacturer's instructions. The cycling program is as follows:
        - Hold: 50°C for 2 minutes.
        - Hold: 95°C for 20 seconds.
        - Cycle 40 times: 95°C for 3 seconds and 60°C for 30 seconds.
    - Calculate the average concentration of the undiluted library. Amplified libraries typically have a yield of 100-500pM.
    - Dilute each library to 100pM with Low TE. Proceed to the template preparation or store library at -20°C until use.

7.7.3. Preparation of the template and enrichment

Library fragments are clonally amplified onto Ion Sphere Particles (ISPs) through emulsion PCR (EmPCR) and then enriched for template-positive ISPs. This step was performed using the Ion OneTouch 2 System (Thermo Fisher Scientific, Cat. No. 4474779), which consists of two modules: The Ion OneTouch 2 (OT2) Instrument and the Ion OneTouch ES (Enrichment System) (Figure 12).

We used the protocol written in the "Ion PGM™ Hi-Q™ View OT2 Kit user guide for use with: Ion OneTouch™ 2 System Catalogue Number A29900. Publication Number MAN0014579 Revision C.0".

We followed the next steps:

1. Prepare template: EmPCR is the method for template amplification that is used in this NGS-based sequencing platform. The EmPCR is based on the dilution and

compartmentalization of template molecules in water droplets in a water-in-oil emulsion. Ideally, each droplet contains a single template molecule.

2. Enrichment of the template: This allows the selection of the ISPs that contain the amplified library fragments to streptavidin coated magnetic beads, washing and denaturing the library strands to collect the template-positive ISPs. This step is made to improve the sequencing yield by using the Ion OneTouch ES.
3. Quality control: To determine the percentage of templated ISPs that have been enriched we used the Qubit®2.0 Fluorometer (Thermo Fisher Scientific, Cat. No. Q32866) and the Ion Sphere Quality Control kit (Thermo Fisher Scientific, Cat. No. 4468656), following manufacturers' instructions. We obtained >90% enrichment efficiency for all reactions

Figure 12. A, Image of the Ion OneTouch 2 system including the Ion OneTouch 2 instrument (right) and the Ion OneTouch ES instrument (left). B, Schematic representation of the Ion OneTouch reaction filter and C, its correct rotation with the amplification solution. Modified from Ion PGM™ Hi-Q™ OT2 Kit user guide.

Reagents and materials

- Ion PGM™ Hi-Q™ View OT2 Kit (Thermo Fisher Scientific, Cat. No. A29900) (8rxn), which contains:

- Ion PGM™ OT2 Supplies (Cat. No. A27744), containing: Ion OneTouch™ Reagent Tubes (2 tubes), Ion OneTouch™ Recovery Routers (8 routers), Ion OneTouch™ Recovery Tubes (16 tubes), Ion OneTouch™ Sipper Tubes (2 sipper tubes), Ion OneTouch™ Amplification Plate (8 plates), Ion OneTouch™ Cleaning Adapter (8 adapters), Ion OneTouch™ Reaction Filter (8 reaction filters) and tubes Ion OneTouch™ ES Supplies (12 pipette tips and 1 box of ES 8-well strips).
- Ion PGM™ Hi-Q™ View OT2 Reagents (Cat. No. A29811), containing: Ion PGM™ Hi-Q™ View Reagent Mix (8 × 800µL), Ion PGM™ Hi-Q™ View Enzyme Mix (1 x 400µL) and Ion PGM™ Hi-Q ™ View ISPs (1 x 800µL).
- Ion PGM™ Hi-Q™ OT2 Solutions (Cat. No. A27742), containing: Ion OneTouch™ Breaking Solution (2 × 1.2mL), Ion OneTouch™ Oil (1 x 450mL), Ion OneTouch™ Reaction Oil (1 x 25mL) Nuclease-free Water (1 x 30mL), Ion OneTouch™ Recovery Solution (1 x 350mL), Neutralization Solution (1 x 100µL), Ion OneTouch™ Wash Solution (1 x 16mL), MyOne™ Beads Wash Solution (2 tubes x 1.4ml) and Tween™ Solution (1 x 6mL).

- Ion Sphere Quality Control kit (Thermo Fisher Scientific, Cat. No. 4468656), containing: Ion Probes (20µL), Alexa Fluor™ 488 Calibration Standard (400µL), Alexa Fluor™ 647 Calibration Standard (400µL), Annealing Buffer (400µL) and Quality Control Wash Buffer (20mL).
- DNA LoBind tubes (1.5mL) (Eppendorf AG, Cat. No. 022431021).
- Sodium hydroxide (NaOH) (Merck KgaA, Cat. No. S8045).
- DynaMag™-2 magnet (Thermo Fisher Scientific, Cat. No. 12321D).
- Ion OneTouch 2 System (Thermo Fisher Scientific, Cat. No. 4474779).
- Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Cat. No. Q32866).
- Centrifuge Sorvall™ Legend™ Micro 17R (Thermo Fisher Scientific, Cat. No. 75002440).
- Minicentrifuge (Nippon Genetics Europe, Dueren, Germany, Cat. No. NG002R).

Procedure

1. Prepare template
    - Set up the Ion OneTouch ™ Instrument
        - Open lid and dispense 150µL Ion OneTouch™ Breaking Solution into each of the two Recovery Tubes. Insert the Recovery Tubes into each slot of the centrifuge and close the lid of the centrifuge.
        - Install the Amplification Plate, disposable tubes and injector following manufacturers' instructions.
        - Install the Ion OneTouch ™ Oil and Ion OneTouch Recovery Solution following manufacturers' instructions.

- Prepare the amplification solution
    - Prepare the Ion PGM™ Hi-Q™ View Reagent Mix at room temperature by vortexing for 30 seconds and centrifuging for 2 seconds. Keep it at room temperature.
    - Prepare the Ion PGM™ Hi-Q™ View Enzyme Mix. Centrifuge for 2 seconds and palace it on ice.
    - Place the Ion PGM™ Hi-Q™ View ISPs at room temperature for 30 minutes.
- Dilute the 100pM library to 8pM by adding 2µL of the library to 23µL of Nuclease free-water. Vortex the diluted library for 5 seconds, centrifuge for 2 seconds and place it on ice.
- Prepare the Ion PGM™ Hi-Q™ View ISPs by vortexing for 1 minute. Then centrifuge for 2 seconds and mix the ISPs by pipetting up and down.
- Add to the tube containing 800µL Ion PGM™ Hi-Q™ View Reagent the following components in order and pipet up and down to mix:
    - Nuclease-free Water: 25µL
    - Ion PGM™ Hi-Q™ View Enzyme: 50µL.
    - Diluted library (8pM): 25µL
    - Ion PGM™ Hi-Q™ View ISPs: 100µL
- Vortex the amplification solution at maximum speed for 5 seconds.
- Fill the Ion OneTouch™ Reaction Filter with the amplification solution and the Ion OneTouch™ Reaction Oil. Install the filled filter as indicated by manufacturers' (Figure 12).
- Run the Ion OneTouch™ Instrument, following manufacturers' instructions.
- Recover the template-positive ISPs.
    - Open lid and remove both Ion OneTouch™ Recovery Tubes from the instrument.
    - Remove all but 100µL of the Recovery Solution from each tube without disturbing the ISP pellet.

  At this point the Recovery Solution could be stored at 2 to 5°C for 3 days by adding 500µL of Ion OneTouch™ Wash Solution.
- Process the ISPs.
    - Add 500µL of Ion OneTouch™ Wash Solution to the Recovery Tubes and pipet up and down to disperse the ISPs.
    - Combine the two tubes into the suspension from one 1.5mL Eppendorf LoBind™ Tube.
    - Centrifuge for 2.5 minutes at 15.500 × g.
    - Remove all but 100µL of the Wash Solution from the tube by using a pipette without disturbing the pellet.

- Asses the quality of the unenriched template-positive ISPs by transferring 2µL to a tube and using the Qubit®2.0 Fluorometer. Follow manufacturers' instructions.

2. Enrichment of the template
    - Prepare fresh Melt-Off Solution (Supplementary data 7).
    - Resuspend Dynabeads™ MyOne™ Streptavidin C1 Beads:
        - Vortex the tube for 30 seconds, then centrifuge for 2 seconds.
        - Pipet the pellet of beads up and down.
        - Transfer 13µl of Dynabeads™ MyOne™ Streptavidin C1 Beads to a new 1.5mL Eppendorf LoBind™ Tube and place it on a DynaMAg™ -2 magnet for 2 minutes.
        - Remove and discard the supernantant without disturbing the pellet.
        - Add 130µl of MyOne™ Beads Wash Solution to the Dynabeads™ MyOne™ Streptavidin C1 Beads.
        - Remove the tube from the magnet, vortex for 30 seconds, and centrifuge for 2 seconds.
    - Fill the 8-well strip by adding to the different wells (Figure 13):
        - Well 1, Template-positive ISP sample: 100µL.
        - Well 2, Dynabeads™ MyOne™ Streptavidin C1 Beads resuspended in MyOne™ Beads Wash Solution: 130µL.
        - Well 3, 4 and 5, Ion OneTouch™ Wash Solution: 300µL.
        - Well 6, empty.
        - Well 7, Melt-Off Solution: 300µL

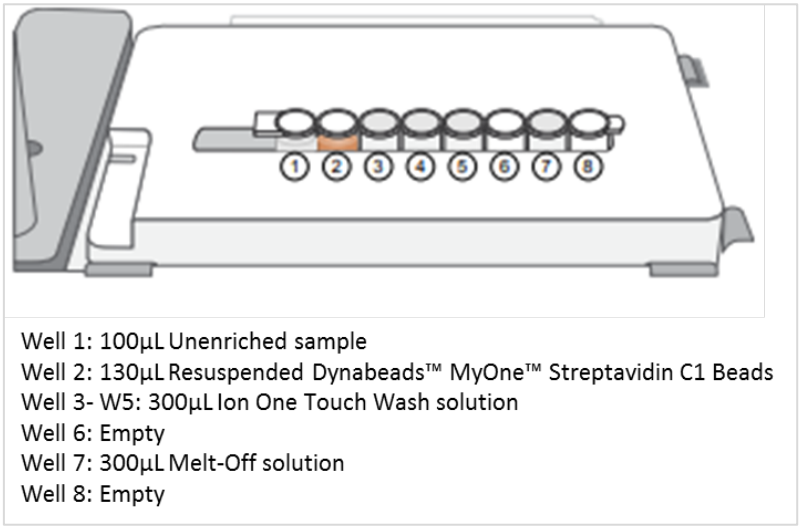

Figure 13. Schematic disposition of the 8-well strip on the Ion OneTouch ES. Modified from Ion PGM™ Hi-Q™ OT2 Kit user guide.

- Install the 8-well strip on the slot of the tray as indicated by manufacturers.
- Prepare the Ion OneTouch™ ES and perform the run following manufacturers' instructions.

3. Determination of the enrichment efficiency using the Qubit®2.0 Fluorometer. Follow manufacturers' instructions.
   At this point the enriched ISPs can be stored at 2 to 8°C for up to 3 days.

7.7.4. Sequencing

Ion Personal Genome Machine™ (PGM™) System (Thermo Fisher Scientific) is based on semiconductor sequencing technology. This method detects hydrogen ions that are released during the polymerization of DNA, in which a complementary strand is built based on the sequence of a template strand. Each microwell in the chip containing a template DNA strand is displayed with one of the four types of dNTPs, and if the introduced dNTP is complementary to the template nucleotide, it will be incorporated to the complementary strand and therefore, the sensor detects the liberation of a hydrogen ion and leads to an electronic signal.

We used the protocol written in "Ion PGM™ Hi-Q™ View Sequencing Kit user guide for use with: Ion PGM™ Hi-Q™ View Sequencing Kit Ion PGM™ System Ion 318™ Chip v2 BC Ion 316™ Chip v2 BC Ion 314™ Chip v2 BC Catalogue Number A30044 Publication Number MAN0014583 Revision C.0".

We followed the next steps:

1. Create a planned run: Perform the planning of all the settings used in a sequencing run in the Ion PGM™ Torrent Browser.
2. Clean and initialize the sequencer: The sequencer requires cleaning every time the instrument is initialized. Cleaning with Milli-Q water is done daily, when instrument is in use, and if more than 27 hours since the last initialization. Cleaning with chlorite solution is done once a week or if the sequencer has been left with reagents more than 48 hours.
3. Load the chip and sequence.

Reagents and materials

- Ion PGM™Hi-Q™ View Sequencing Kit (Thermo Fisher Scientific, Cat. No. A30044), includes the following components:
  - Ion PGM™ Sequencing Supplies (Part No. A25587), containing: Wash 1 Bottle (1 bottle x 250mL), Wash 3 Bottle (1 bottle x 250mL), Ion PGM™ Reagent Bottle Sipper Tubes (16 sipper tubes), Ion PGM™ Wash Bottle

Sipper Tubes (8 tubes for 250mL bottles), Ion PGM™ Wash Bottle Sipper Tubes (4 tubes for 2L) and Reagent Bottles (25 bottles x 50mL).

- Ion PGM™ Hi-Q™ View Sequencing Reagents (Part No. A30043), containing: Ion PGM™ Hi-Q™ View Sequencing Polymerase (1 tube x 36µL), Ion PGM™ Hi-Q™ View Sequencing Primer (1 tube x 144µL) and Ion PGM™ Hi-Q™ View Control Ion Sphere™ Particles (1 tube x 60µL).
- Ion PGM™ Hi-Q™ View Sequencing Solutions (Part No. A30275), containing: Ion PGM™ Hi-Q™ View Sequencing W2 Solution (4 bottles x 125mL), Ion Cleaning Tablet (4 tablets), Annealing Buffer (1 tube x 12mL) and Ion PGM™ Hi-Q™ View Sequencing W3 Solution (2 bottles x 100mL).
- Ion PGM™ Hi-Q™ Sequencing dNTPs (Part No. A25590), containing: Ion PGM™ Hi-Q™ Sequencing dGTP (1 x 80µL), Ion PGM™ Hi-Q™ Sequencing dCTP (1 x 80µL), Ion PGM™ Hi-Q™ Sequencing dATP (1 x 80µL) and Ion PGM™ Hi-Q™ Sequencing dTTP (1 x 80µL).
- Ion PGM™ Wash 2 Bottle Kit (Cat. No. A25591) includes the following components: Wash 2 Bottle (2L) and Wash 2 Bottle Conditioning Solution (1 bottle x 125mL).

- Sodium hydroxide (NaOH) (Merck KGaA, Cat. No. S8045).
- Ion 316™ Chip Kit v2 BC (8 pack) (Thermo Fisher Scientific, Cat. No. 4488149).
- Ion 318™ Chip v2 BC (8 pack) (Thermo Fisher Scientific, Cat. No. 4488150).
- Ion PGM™ System and accessories (Thermo Fisher Scientific, Cat. No. 4462921)
- Veriti 96-well Thermal Cycler (Thermo Fisher Scientific, Cat. No. 4375786).
- Minicentrifuge (Nippon Genetics Europe, Cat. No. NG002R).
- Ion PGM™ Torrent Server (Thermo Fisher Scientific, Cat. No. 4483643).

Procedure

1. Create a planned run

   To create the run plan, diverse setting need to be selected, such as the sequencing instrument, chip, library and template kit type and the plugins.
2. Clean and initialize the sequencer
   - Cleaning with Milli-Q water and chlorite solution is performed following manufacturers' instructions.
   - Before initialization
     - Remove the dNTP stock solutions to ice.
     - Rinse the Wash 2 Bottle 3 times with 200mL Milli-Q water. Rinse the Wash 1 and Wash 3 Bottles 3 times with 50mL of Milli-Q water.
     - Fill the Wash 2 bottle to the mold line with Milli-Q water. 18 MΩ water.

    - Add the entire bottle of Ion PGM™ Hi-Q™ View Sequencing W2 Solution to the Wash 2 Bottle.
    - Prepare 500µL of 100mM NaOH by diluting 50µL of NaOH 1M in 450µL nuclease-free water.
    - Add 70µL of 100mM NaOH to the Wash 2 Bottle. Mix the bottle by inversion 5 times.
    - Add 350µL of freshly prepared 100 mM NaOH to the Wash 1 Bottle.
    - Add Ion PGM™ Hi-Q™ View Sequencing W3 Solution to the 50mL line marked on the Wash 3 Bottle.
- Begin the initialization as indicated in the screen of the Ion PGM™ System.
- Prepare the Reagent bottles with dNTPs
    - Vortex each dNTP stock solution and centrifuge to collect the contents. Keep the dNTP stock solutions on ice throughout this procedure.
    - Transfer 20µL of each dNTP stock solution into its respective Reagent Bottle.
    - Store on ice until you are ready to attach it to the instrument.
- Attach the sipper tubes and Reagent bottles
    - After the initialization, remove the used sipper tubes and collection trays from the dNTP ports.
    - Insert a new sipper tube into each dNTP port and attach each prepared Reagent Bottle to the correct dNTP port.

3. Load the chip and sequence
    - Add controls to the enriched, template-positive Ion Sphere™ Particles (ISP)
        - Vortex the Control Ion Sphere™ Particles, then pulse-centrifuge in a picofuge for 2 seconds before taking aliquots.
        - Add 5µL Control ISPs directly to the entire volume of enriched, template positive ISPs (prepared using your template preparation method) in a 0.2mL non-polystyrene PCR tube.
    - Anneal the Sequencing Primer
        - Mix the tube containing the ISPs by pipetting up and down and centrifuge for 2 minutes at 15.500 x g.
        - Remove the supernatant from the top down and discard. Leave 15µL in the tube
        - Vortex the Sequencing Primer for 5 seconds, centrifuge in a picofuge for 5 seconds to collect the contents. Leave on ice until use.

- Add 12µL of Sequencing Primer to the ISPs. Mix by pipetting to disrupt the pellet.
- Place the tube in the thermal cycler and run the following program: 95°C for 2 minutes and then 37°C for 2 minutes. Remain the tube in the cycler at RT while the sequencing run is set up.

- Perform chip check to ensure the functionality of the chip, as indicated in the screen of the Ion PGM™ System.
- Bind the Sequencing Polymerase
  - Remove the Ion PGM™ Hi-Q™ View Sequencing Polymerase from storage and mix with the fingertip 4 times. Centrifuge for 3–5 seconds and place on ice.
  - Remove the ISPs from the thermal cycler, add 3µL of Ion PGM™ Hi-Q™ View Sequencing Polymerase to the ISPs and mix by pipetting. Incubate at room temperature for 5 minutes.
- Prepare and load the chip (Figure 14)
  - Remove the new chip from the Ion PGM™ Sequencer and insert a used chip in the clamp while loading the new chip.
  - Tilt the new chip at a 45° angle so that the loading port is the lower port. Insert the pipette tip firmly into the loading port, then remove as much liquid as possible from the loading port. Discard the liquid.
  - Place the chip upside-down in the minifuge bucket, then transfer the bucket with the chip tab pointing in. Centrifuge for 5 seconds to empty the chip completely. Remove the chip from the bucket.
  - Place the chip on a firm surface and insert the tip into the loading port of the chip.
  - Collect 30µL of prepared ISPs with a pipette and load the chip slowly by applying gentle pressure between the tip and chip and slowly dial down the pipette to deposit the ISPs. Avoid introducing bubbles into the chip and discard any displaced liquid from the other port of the chip.
  - Centrifuge the chip for 30 seconds with the chip tab pointing towards the centre of the minifuge.
  - Centrifuge the chip for 30 seconds with the chip tab pointing away from the centre of the minifuge. Remove the chip.
  - Tilt the chip 45 degrees. Slowly pipet the sample out and then back into the chip.
  - Slowly remove as much liquid as possible from the chip. Discard the liquid.

- Spin the chip upside-down for 5 seconds in the minifuge. Remove and discard any liquid. Proceed immediately to perform the run.

4. Perform the run
   - Select the planned run in the screen of the Ion PGM™ Sequencer.
   - Confirm the settings. Load and clamp the chip.
   - Perform the run.

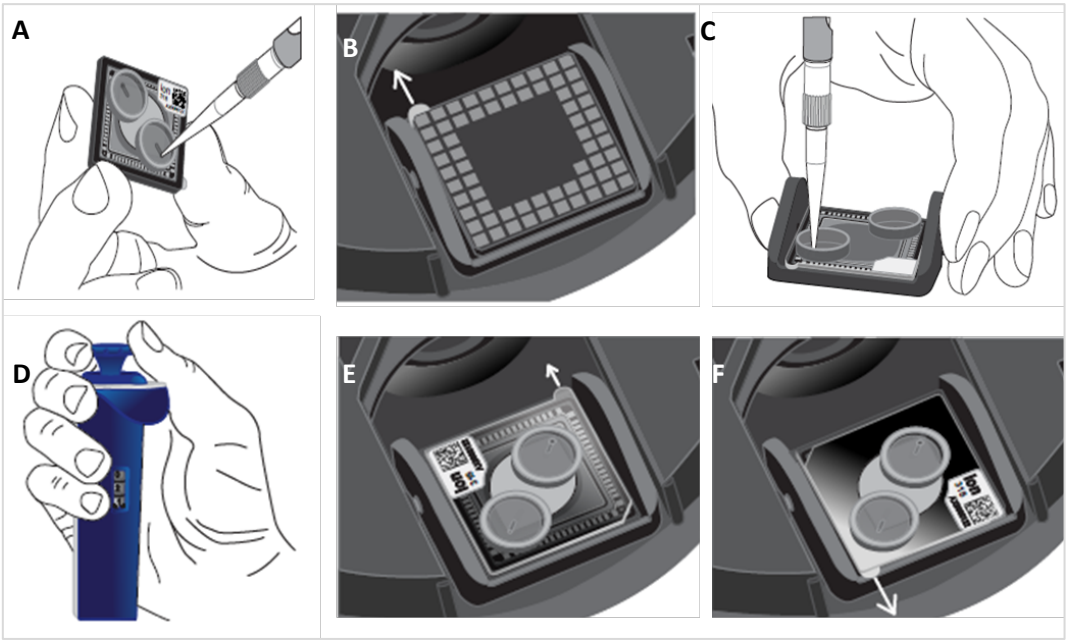

Figure 14. Images showing the preparation and loading of the chip. A, removal of the liquid. B, Chip upside-down in the minifuge bucket. C, load the chip on a flat surface. D, dial down the pipette to deposit the ISPs. E, Chip tab pointing towards the centre of the minifuge. F, Chip tab pointing away from the centre of the minifuge.

7.8. DETECTION OF COPY NUMBER VARIATIONS

Copy Number Variations (CNV) are a prominent source of genetic variation in human DNA and play a role in a wide range of disorders. We used four different techniques for the detection of deletions, amplifications and complex rearrangements. These are Multiplex Ligation-dependent Probe Amplification (MLPA), Quantitative Multiplex Polymerase chain reaction of Short Fluorescent Fragments (QMPSF), genetic markers or a Comparative Genome Hybridization array (aCGH). Furthermore, we analysed the relocation of the *SRY* gene by Fluorescence *in situ* hybridization (FISH). A summary of the several techniques used to detect CNV is shown below (Table 6).

Table 6. Summary of the different methods used to detect copy number variations.

Method	Commercial name	Genes
MLPA	SALSA MLPA probemix P185-C2 Intersex	*NR0B1, CXorf21, SOX9, SRY, ZFY, WNT4, NR5A1*
	SALSA MLPA P118 WT1 probemix	*WNT1*
QMPSF		*SRD5A2*
		AR
Genetic markers		*NR0B1*
aCGH	SurePrint G3 Human CGH Microarray 8x60K	Genome wide probes
FISH		*SRY*

aCGH, Comparative Genome Hybridization array; MLPA, Multiplex Ligation-dependent Probe Amplification; QMPSF, Quantitative Multiplex Polymerase chain reaction of Short Fluorescent Fragments.

7.8.1. Multiplex Ligation-dependent Probe Amplification

MLPA is a semi-quantitative technique, used to determine the copy number of DNA sequences in a single multiple PCR-based reaction. It is based on the amplification of several probes that detect a specific DNA sequence of approximately 60 nucleotides and its separation by capillary electrophoresis.

The first step comprises the denaturation of the sample DNA and adding of a mixture of probes. These probes are two oligonucleotides that hybridize to adjacent target sequences so that are ligated to create a single probe. Then, all ligated probes are amplified simultaneously using the same PCR primer pair. One of the primers is fluorescently labelled, which allows the visualization of the amplification products during fragment separation on a capillary electrophoresis instrument. PCR amplicons result in products between 64-500 nucleotides in length.

Compared to relative probe peak height in various control DNA samples, the relative height of each individual probe peak reflects the relative copy number of the corresponding target sequence in the sample. Thus, the deletion of one or more target sequence converts as a relative decrease in a peak height, while an increase in a relative peak height reflects amplification.

We used the following SALSA MLPA probemix products to detect CNVs in our population of interest.

1. SALSA MLPA probemix P185-C2 Intersex (MRC-Holland BV, Cat. No. P185-050R). It contains 44 probes with amplification products between 130 and 494nt for the following genes: NR0B1, CXorf21, SOX9, SRY, ZFY, WNT4 and NR5A1. Furthermore, it contains 10 control fragments.

2. SALSA MLPA P118 WT1 probemix (MRC-Holland BV, Cat. No. P118-050R), containing 34 MLPA probes with amplification products between 130 and 409nt and 12 control fragments. It spans the *WT1* and *AMER1* genes.

The included probes in the SALSA MLPA probemix P185-C2 Intersex and SALSA MLPA P118 WT1 probemix, as well as its gene location and partial sequence are shown on http://www.mlpa.com/.

The amount of used DNA for this methodology ranges between 50 and 250ng.

Reagents and materials

- SALSA MLPA probemix (MRC-Holland BV, Amsterdam, The Netherlands), containing SALSA MLPA Buffer (1x 180µl), SALSA Ligase-65 (1x 115µl), Ligase Buffer A (1x 360µl), Ligase Buffer B (1x 360µl), SALSA PCR Primer Mix (1x 240µl), SALSA Polymerase (1x 65µl) and Probemix (1x 80µl).
- Formamide (Thermo Fisher Scientific, Cat. No. 4311320).
- 10X Genetic Analyser Running Buffer (Thermo Fisher Scientific, Cat. No. 402824).
- POP-7™ Polymer (Thermo Fisher Scientific, Cat. No. 4352759).
- Milli-Q ultrapure water (Merck KgaA).
- ABI PRISM 3130xl Genetic Analyser (Thermo Fisher Scientific, Cat. No. 4359571).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).
- Beckman Spinchron R centrifuge (Beckman Coulter, Cat. No. 358723).
- MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher Scientific, Cat. No. 4316813).
- MicroAmp™ Clear Adhesive Film (Thermo Fisher Scientific, Cat. No. 4306311).
- GeneScan™ 500 ROX™ dye size Standard (Thermo Fisher Scientific, Cat.No. 401734).
- GeneMapper® Software v4.0 (Thermo Fisher Scientific).

Procedure

The protocol can be divided into the following steps:

1. DNA denaturation:
    - Add 5µL DNA sample to each tube.
    - Place the tubes in the thermocycler and run the following program:
        - DNA denaturation: 98°C for 5 minutes.
        - Hold at 25°C.
2. Hybridization reaction
    - Vortex SALSA MLPA Buffer and Probemix tubes before use.

- Prepare master mix on ice. For each sample add:
    - SALSA MLPA Buffer: 1.5µL
    - Probemix: 1.5µL
- Mix by vortexing.
- Add 3µL to each sample tube.
- Place the tubes in the thermocycler and run the following program:
    - Incubation: 95°C for 1 minute.
    - Hybridization: 60°C for 16-20 hours.

3. Ligation reaction
    - Vortex Ligase Buffer A and Ligase Buffer B.
    - Prepare the Ligase-65 master mix at RT. For each sample add:
        - Milli-Q water: 25µl
        - Ligase Buffer A: 3µl
        - Ligase Buffer B: 3µl
        - SALSA Ligase-65: 1µL
    - Mix by pipetting up and down.
    - Add 32 µl to each sample tube placed on the thermocycler and mix by pipetting up and down. Run the following program:
        - Ligation: Hold at 54°C
        - Activation of Ligase-65 enzyme: 98°C for 5minutes.
        - Hold at 20°C
4. PCR reaction
    - Vortex SALSA PCR Primer Mix.
    - Prepare the Polymerase master mix at room temperature. For each sample add:
        - Milli-Q water: 7.5µL
        - SALSA PCR Primer: 2µL
        - SALSA Polymerase: 0.5µL
    - Add 10µL Polymerase master mix to each reaction. Mix by pipetting up and down.
    - Place the tubes in the thermocycler and run the following program:
        - Cycle 35 times: 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 60 seconds
        - Incubate: 72°C for 20 minutes.
        - Hold at 15°C

    At this point PCR products could be stored at 4°C for one week or at -25°C for longer periods.
5. Fragment separation on a capillary electrophoresis instrument
    - For each sample prepare the following mix:
        - Formamide: 10µL
        - GeneScan™ 500 ROX™ dye size Standard: 0,5µL

- Add 10µL of the previous mix to 2µL of the PCR product.
- Mix by vortexing and spin in a centrifuge.
- Place the samples in the thermocycler at 95°C for 5 minutes.
- Place the samples on ice for 5 minutes.
- Place the plate on the genetic analyser.

6. Data analysis

Data analysis was performed with the GeneMapper® Software v4.0 (Thermo Fisher Scientific). To normalise MLPA, first each peak height is divided by the sum of the peak values of control probes. These control probes are not supposed to contain CNV and are outside the screened region. Given values are then compared to the height of control DNA samples. This ratio balance between 0.8 and 1.15 indicate 2 gene copies or normal sample. According to the criteria established in our laboratory ratio balances under 0.8 indicate a decrease gene copy number and above 1.15 are indicative of a gene copy number increase in the sample.

7.8.2. Quantitative Multiplex Polymerase Chain Reaction of Short Fluorescent

QMPSF, as well as MLPA, relies on the simultaneous amplification of small genomic sequences under quantitative conditions, using fluorescent primers, and on the comparison of profiles generated from tested and control DNA.

Comparing to the MLPA technique, QMPSF is faster and cheaper, but it is less specific due to the absence of a ligation step. We designed a QMPSF assay for those genes that were not included in the MLPA probemix available in the laboratory (P185-C2 Intersex or P118 WT1).

We made QMPSF to check variations in copy number in the *AR* and *SRD5A2* genes. We designed 5' 6-FAM dye-labelled primers corresponding to five target regions that were simultaneously amplified. The PCR comprises the processes of denaturation, annealing and elongation, but it is stopped at exponential amplification allowing the quantification. In these cases, QMPSF was stopped at cycles 25 and 30, respectively. For the *SRD5A2* two multiplex PCRs were done because of the similar size of some amplicons (Primers sequences and amplicons size are given in Supplementary data 6).

QMPSF included control amplicons. For the detection of CNV on *SRD5A2* and *AR*, control fragments from other not DSD-related genes, such as *RET* and *HNF1B* located on chromosomes 10q11.21 and 17q12, respectively, were used. We also added 3 normal samples and a negative control containing water. All the samples were tested in duplicates and we used 100ng of DNA.

Reagents and materials

- KAPA Taq PCR Kit (Kapa Biosystems, Cat. No. KK1014), containing KAPA Taq DNA Polymerase (5U/µL), KAPA Taq Buffer A (10X) KAPA Taq Buffer B (10X) and $MgCl_2$ (25Mm).
- GeneScan™ 500 ROX™ dye size Standar (Thermo Fisher Scientific, Cat.No. 401734).
- 10X Genetic Analyser Running Buffer (Thermo Fisher Scientific, Cat. No. 402824).
- Formamide (Thermo Fisher Scientific, Cat. No. 4311320).
- dNTP Mix (5mL) (Merck KgaA, Cat. No. D7295-5ML).
- DMSO (50mL) (Merck KgaA, Cat. No. D8418-50Ml).
- POP-7™ Polymer (Thermo Fisher Scientific, Cat. No. 4352759).
- 5' 6-FAM Dye-labelled primers pool (250µM) (Integrated DNA Technologies Inc, Skokie, Il, USA).
- Milli-Q ultrapure water (Merck KgaA).
- ABI PRISM 3130xl Genetic Analyser (Thermo Fisher Scientific, Cat. No. 4359571).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Beckman Spinchron R centrifuge (Beckman Coulter, Cat. No. 358723).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).
- MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher Scientific, Cat. No. 4316813).
- MicroAmpTM Clear Adhesive Film (Thermo Fisher Scientific, Cat. No. 4306311).
- PCR-strip tubes (Sarstedt AG&Co KG, Nümbrecht, Germany, Cat. No 72.982.002).
- PCR lid strips (Sarstedt AG&Co KG, Cat. No. 65.989.002).
- GeneMapper® Software v4.0 (Thermo Fisher Scientific).

Procedure

The protocol can be divided into the following steps:

1. PCR multiplex reaction

For each amplification a mix containing Milli-Q water, Buffer A, dNTPs, FAM dye labelled primers at 25µM, Taq DNA polymerase and DNA (100ng) is prepared. DMSO is added when required. Conditions in the thermocycler are also different for each DNA amplification sequence. Conditions for amplification and primers can be found in Supplementary data 6.

2. Fragment separation on a capillary electrophoresis instrument
    - For each sample prepare the following mix:
        - Formamide: 10µL

- GeneScan™ 500 ROX™ dye size Standard: 0.5µL
- Add 10µL of the previous mix to 2µL of the PCR product.
- Mix by vortexing and spin in a centrifuge.
- Place the samples in the thermocycler at 95°C for 5 minutes.
- Place the samples on ice for 5 minutes.
- Place the plate on the genetic analyser.

3. Data analysis

Data analysis was performed with the GeneMapper® Software v4.0 (Thermo Fisher Scientific). To normalise data, first each peak height is divided by the sum of the peak values of control probes. These control probes are not supposed to contain CNV and are outside the screened region. Given values are then compared to the height of control DNA samples. This ratio balance between 0.8 and 1.15 indicate 2 gene copies or normal sample. According to the criteria established in our laboratory ratio balances under 0.8 indicate a decrease gene copy number and above 1.15 are indicative of a gene copy number increase in the sample.

7.8.3. Localization of genetic markers

A genetic marker is a fragment of DNA with a known location on a chromosome. They may be a short DNA sequence surrounding a single base pair change or a long one, like microsatellites. Microsatellites can be used for mapping locations within the genome to locate a gene or a mutation.

In this case, we used microsatellites to determine the extension of the deletions previously showed by PCR and MLPA. For this study labelled primers were designed and analysed by fluorescent PCR and subsequent electrophoresis on the ABI3130xl genetic analyser using GeneMapper® v4.0 software (Applied Biosystems) was done.

We analysed the microsatellites in two samples in which a *NR0B1* complete gene deletion was observed by PCR and MLPA (GN0101 and POL0285). The 4 microsatellites we studied were located in Xp21, specifically DX1218 and DXS8039 were placed downstream *NR0B1* while DXS1083 and DXS992 were located upstream the gene (Supplementary data 6).

Reagents and materials

- KAPA Taq PCR Kit (Kapa Biosystems, Cat. No. KK1014), containing KAPA Taq DNA Polymerase (5U/µL), KAPA Taq Buffer A (10X) KAPA Taq Buffer B (10X) and MgCl2 (25Mm).
- GeneScan™ 500 ROX™ dye size Standar (Thermo Fisher Scientific, Cat.No. 401734).

- 10X Genetic Analyser Running Buffer (Thermo Fisher Scientific, Cat. No. 402824).
- Formamide (Thermo Fisher Scientific, Cat. No. 4311320).
- dNTP Mix (5mL) (Merck KgaA, Cat. No. D7295-5ML).
- Agarose D-1 (Pronadisa, Madrid, Spain, Cat. No. 8016).
- 50X TAE Buffer (Bio-Rad, Hercules, CA, USA, Cat. No. 1610773).
- EZ Load™ 100bp Molecular Ruler (Bio-Rad, Cat. No. 1708352).
- Nucleic Acid Sample Loading Buffer 5X (Bio-Rad, Cat. No. 161-0767).
- GelRed™ (Biogen, Cambridge, MA, USA, Cat. No. BT41003).
- DMSO (50mL) (Merck KgaA, Cat. No. D8418-50ML).
- POP-7™ Polymer (Thermo Fisher Scientific, Cat. No. 4352759).
- 5' 6-FAM Dye-labelled primers pool (250µM) (Integrated DNA Technologies Inc, Skokie, Il, USA).
- Milli-Q ultrapure water (Merck KgaA).
- ABI PRISM 3130xl Genetic Analyser (Thermo Fisher Scientific, Cat. No. 4359571).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Beckman Spinchron R centrifuge (Beckman Coulter, Cat. No. 358723).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).
- Electrophoresis chamber Sub® Cell GT MINI (Bio-Rad, Hercules, CA, USA, Cat. No. 1704467).
- PowerPac™ 3000 Power Supply (Bio-Rad, Cat. No. 1655057).
- G: BOX Chemi 16 Bio Imaging system (Syngene, Cambridge, UK, Cat. No. SGBOX)
- GeneSnap Software v7.12 (Syngene).
- Sub-Cell GT UV-Transparent Gel Tray (Bio-Rad, Cat. No. 1704435).
- Fixed-Height Comb (Bio-Rad, Cat. No. 1704465).
- MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher Scientific, Cat. No. 4316813).
- MicroAmpTM Clear Adhesive Film (Thermo Fisher Scientific, Cat. No. 4306311).
- PCR-strip tubes (Sarstedt AG&Co KG, Nümbrecht, Germany, Cat. No 72.982.002).
- PCR lid strips (Sarstedt AG&Co KG, Cat. No. 65.989.002).
- GeneMapper® Software v4.0 (Thermo Fisher Scientific).

Procedure

The protocol can be divided into the following steps:

1. PCR multiplex reaction

For each amplification, mix Milli-Q water, Buffer A, dNTPs, FAM dye labelled primers at 25µM, Taq DNA polymerase and DNA (20-100ng). DMSO is added when

required. Conditions for amplification and primers can be found in Supplementary data 6.

2. Perform an agarose gel electrophoresis

PCR fragments are separated on the agarose gel and visualized when exposed to UV light. If amplification of the genetic marker happens, a specific band appears at the expected location compared to the molecular ruler.

See above for agarose gel electrophoresis preparation.

3. Fragment separation on a capillary electrophoresis instrument
    - For each sample prepare the following mix:
        - Formamide: 10µL
        - GeneScan™ 500 ROX™ dye size Standard: 0.5µL
    - Add 10µL of the previous mix to 2µL of the PCR product.
    - Mix by vortexing and spin in a centrifuge.
    - Place the samples in the thermocycler at 95°C for 5 minutes.
    - Place the samples on ice for 5 minutes.
    - Place the plate on the genetic analyser.
4. Data analysis

Data analysis was performed with the GeneMapper® Software v4.0 (Applied Biosystems). For each genetic marker, the results were determined by comparing the presence or absence of the region of interest with that observed on the DNA control samples.

7.8.4. Array-based comparative genomic hybridization

The main advantage of using aCGH is the facility to simultaneously detect aneuploidies, deletions, duplications and amplifications of any *locus* represented on the array, offering higher resolution than traditional cytogenetic methods.

This method uses slides arrayed with small fragments of DNA, known as probes, as the targets for analysis. Probes vary in size and resolution is determined by the genomic distance between DNA probes and the length of these probes. The methodology uses the sample DNA and a control DNA, differentially labelled with fluorescent dyes, which are mixed and applied to the slide to hybridize with the arrayed single-strand probes. Digital imaging systems are used to capture and quantify the relative fluorescence intensities of the labelled DNA probes hybridized to each target. The fluorescent signal intensity of the sample DNA, relative to that of the control DNA, can then be linearly plotted across each chromosome, allowing the identification of copy number changes.

We used the SurePrint G3 Human CGH Microarray 8x60K (Agilent Technologies, Santa Clara, CA, USA), containing 55.077 biological probes and 41kb overall median probe spacing. The optimal amount of sample DNA for the aCGH is 500ng.

Reagents and materials

- SureTag Complete DNA Labeling Kit (Agilent Technologies, Cat. No. 5190-4240), containing 10X restriction Enzyme Buffer, BSA (Bovine Serum Albumin), AluI, RsaI, random primer, 5X Reaction Buffer, 10X dNTPs, Cyanine 3'-dUTP and Cyanine 5'-dUTP.
- TE Buffer Solution (pH 8.0) (100mL) (Merck KGaA, Cat. No. 8890-100ML).
- Oligo aCGH/ChIP-on-chip Hybridization Kit (Agilent Technologies, Cat. No. 5188-5220), containing 10x aCGH Blocking Agent and 2xHI-RPM Hybridization Buffer.
- Human Cot-1 DNA (1mg/mL) (Agilent Technologies, Cat. No. 5190-3393).
- Agilent Oligo aCGH Wash Buffer 1 and 2 Set (Agilent Technologies, Cat. No. 5188-5226).
- Milli-Q ultrapure water (Merck KGaA).
- SurePrint G3 Human CGH Microarray 8x60K (Agilent Technologies, Cat. No. G4450A).
- Purification columns (50 u) (Agilent Technologies, Cat. No. 5190-3391).
- GeneAmp®PCR System 9700 (Thermo Fisher Scientific, Cat. No. 4339386).
- Minicentrifuge (Nippon Genetics Europe, Cat. No. NG002R).
- Hybridization oven rotator (Sheldon Manufacturing, Inc., Cornelius, OR, USA, Cat. No. SHO1).
- Hybridization Chamber (Agilent Technologies, Cat. No. G2534A).
- Hybridization Chamber Gasket Slides (Agilent Technologies, Cat No. G2534A).
- SureScan Microarray Scanner G4900DA (Agilent Technologies, Cat No. G4900DA).
- Gasket slide, 8 microarrays/slide format (Agilent Technologies Cat. No. G2534-60014).
- Slide-staining dishes (VWR, Radnor, PA, USA, Cat. No. 470175-194).
- Magnetic stir plate (Velp Scientifica, Usmate, Italy, Cat. No. F20700431).
- SureScan Microarray Scanner (Agilent Technologies).
- NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific).
- NanoDrop 1000 Spectrophotometer Software v3.8 (Thermo Fisher Scientific).
- Agilent CytoGenomics software v2.7 (https://www.agilent.com/en/download-agilent-cytogenomics-software).
- Feature Extraction software (https://www.agilent.com/en/product/mirna-microarray-platform/mirna-microarray-software/feature-extraction-software-228496).

Procedure

The protocol can be divided into the following steps:

1. Label the samples
    - For each reaction add the amount of DNA to the appropriate nuclease-free water to bring to a final volume of 13µL.
    - Add 2.5µL of Random primer to each reaction tube. Mix by pipetting up and down.
    - Place the tubes in the thermocycler and run the following program:
        - 98°C for 10 minutes.
        - Hold at 4°C.
    - Spin the samples in the minicentrifuge
    - Prepare the Labelling master mix on ice. For each sample add:
        - 5X Reaction Buffer: 5µL
        - 10X dNTPs: 2.5µL
        - Cyanine3'-dUTP (For control DNA): 1.5µL
        - Cyanine 5'-dUTP (For samples DNA): 1.5µL
        - Exo (-) Klenow: 0.5µL
    - Add 9.5µL of the Labelling master mix to each sample. Mix by pipetting up and down.
    - Place the tubes in the thermocycler and run the following program:
        - 37°C for 2 hours
        - 65°C for 10 minutes
        - Hold at 4°C

    At this point reactions can be stored at -20°C in dark for 1 month.
2. Purification of labelled gDNA
    - Centrifuge the samples for 1 minute at 6,000 x g.
    - Add 430µL TE to each reaction tube.
    - Place a column in a 2mL collection tube. Load each labelled gDNA onto a column and cover the column with a cap.
    - Spin for 10 minutes at 14,000 x g in a microfuge at room temperature. Discard the flow-through.
    - Transfer the column to a new 2mL collection tube. Spin for 1 minute at 1,000 x g in a microfuge. Spin.
    - Dry the labelled gDNA to bring the sample volume to 9.5µL.
    - Take 1.5µL of each sample and measure the DNA concentration on the NanoDrop® ND-1000 spectrophotometer.
    - Combine 8µL test sample with 8µL control sample for a total volume of 16µL.

    At this point reactions can be stored at -20°C in dark for 1 month.
3. Hybridization

- Prepare the 10X Blocking Agent (Supplementary data 7): Vortex the solution before use.
- Prepare the Hybridization master mix at room temperature. For each sample add:
    - Cot-1 DNA (1.0mg/mL): 2µL
    - 10X aCGH Blocking Agent: 4.5µL
    - 2x HI-RPM Hybridization Buffer: 22.5µL
- Add 29µL of the Hybridization master mix to each sample. Mix by pipetting up and down.
- Spin the samples in the minicentrifuge
- Place the tubes in the thermocycler and run the following program: 98°C for 2 minutes and 37°C for 30 minutes.
- Prepare de hybridization assembly as manufacturers' instructions.
- Load 40µL of hybridization sample mixture into the gasket.
- Hybridize at 65°C for 24 hours in the oven rotator.

4. Microarray wash
    - Pre-warm Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 2 at 37°C overnight.
    - Prepare and fill slide-staining dishes as follows:
        - Slide-staining No. 1: Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 1 at room temperature.
        - Slide-staining No. 2: Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 1 at room temperature on a magnetic stir plate.
        - Slide-staining No. 3: Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 2 at 37°C on a magnetic stir plate. Add a magnetic stir bar and turn on the heating element to maintain temperature.
    - Remove the hybridization chamber from the incubator.
    - Disassemble hybridization chamber in Slide staining No. 1, following manufacturers' instructions.
    - Remove the microarray slide to Slide staining No. 2 and stir at 350 rpm for 5 minutes.
    - Transfer to Slide staining No. 3 for 1 minute.
    - Slowly remove the slide rack trying to minimize droplets on the slides.

5. Microarray scanning and analysis

Put assembled slide holders in the scanner cassette and scan the slide. Extract features to process microarray image files with the Feature Extraction Software. This, locates microarray grids, rejects outliers and determines intensities and ratios. Finally, analyze the microarray image using the Agilent CytoGenomics software v2.7. The algorithm used by the software allows the data analysis by the conversion of the fluorescence emission ratio between the sample and control DNA into an image of

coloured plots along the chromosomes or loci where the log2 expression form is represented. When the sample DNA is normal, the fluorescence between both sample and control is similar, therefore the ratio between both fluorescence emissions is 1 and log2 is 0, representing the same number of copies for both DNA sample and control. However, an variants in the Cy3:Cy5 ratio indicates a loss or gain of sample DNA in that specific region. We considered the log2 values above 0.5 or below -0.5 of 3 consecutive probes as a gain or loss of a copy of a gene or region.

To achieve a high accuracy quality control metrics should be within the normal ranges, according to manufacturers.

7.8.5. Fluorescence *in situ* hybridization

Sex determination in humans is dependent upon the *SRY* gene, located on chromosome Yp11.3. Translocation to the X chromosome or to an autosome results in a mismatch between the sexual phenotype and the cytogenetic genotype.

Fluorescence *in situ* hybridization (FISH) represents a widely-used molecular technique that allows detecting relocation of the *SRY* gene away from its normal location on the Y chromosome. This procedure is based on the construction of one or several probes, designed to hybridize to those parts of the chromosome with a high degree of complementarity, and tagged with fluorophores. In our case, the results were visualized on a fluorescence microscope by using DAPI nuclear counterstain, which stains nuclei and contrast to fluorescent probes of other structures.

For the analysis, we used peripheral blood cells reared in metaphase and fixed on microscope slides. Blood samples (3.5mL) were collected on Sodium Heparin tubes.

Reagents and materials

- Vysis SRY Probe LSI SRY Spectrum Orange/Vysis CEP X Spectrum Green Probe Kit (Abbott Mol Inc., Chicago, IL, USA, Cat. No. 06N29-020), containing Vysis LSI SRY SpectrumOrange/CEP X SpectrumGreen Probes (1 x 20µL) (Cat. No. 30-191007) and Vysis LSI/WCP Hybridization Buffer (1 x 150µL) (Cat. No. 30-804824).
- DAPI II counterstain (2 x 500µL) (Abbott Mol. Inc., Cat. No. 06J50-001).
- RPMI 1640 Medium (1L) (Thermo Fisher Scientific, Cat. No. 11875085).
- Fetal bovine serum (FBS) (1L) (Thermo Fisher Scientific, Cat. No. 10106185).
- Phytohemagglutinin (10mL) (Thermo Fisher Scientific, Cat. No. 10576015).
- KaryoMAX™ Colcemid™ Solution (10mL) (Thermo Fisher Scientific, Cat. No. 15212012).
- 20X SSC (500g) (Abbott Mol. Inc., Cat. No. 02J10-032).

- NP-40 (50mL) Thermo Fisher Scientific, Cat. No. 85124).
- Formamide (Thermo Fisher Scientific, Cat. No. 4311320).
- Potassium chloride (KCl) BioXtra ≥99.0% (500g) (Merck KGaA, Cat. No. P9333).
- PBS tablets (Phosphate buffered saline) (50 tablets) (Merck KGaA, Cat. No. P4417-50TAB).
- Ficoll 1077 (100mL) (Merck KGaA, Cat. No. 10771-100ML).
- Methanol (1L) (Merck KGaA, Cat. No. 34860-1L-R).
- Acetic acid (1L) (Merck KGaA, Cat. No. A6283-1L).
- Absolute Ethanol (Thermo Fisher Scientific, Cat. No. BP2818100).
- 100% Acetone (10x 0.5mL) (Merck KGaA, Cat. No. 175862-10X0.5ML).
- 36-38% Hydrochloric acid (HCl) (100mL) (Merck KGaA, Cat. No. H1758-100ML).
- Milli-Q ultrapure water (Merck KGaA).
- Slide and coverslip (Thermo Fisher Scientific, Cat. No. 12373118 and Cat. No. 12343138).
- Coplin jars (VWR, Radnor, PA, USA, Cat. No. 470175-194).
- pH-meter BASIC 20 (Crison instruments S.A, Alella, Spain, Cat. No. Hach LPV2000.98.0002).
- Water baths (Thermo Fisher Scientific, Cat. No. TSGP02).
- Bunsen lighter (Merck KGaA, Cat. No. Z270318-1EA).
- Magnetic stirrer hotplate (Velp Scientifica, Usmate, Italy, Cat. No. F20700431).
- TP2555 Thermal Plate (inTEST Corporation, Mt. Laurel, NJ, USA).
- Minicentrifuge (Nippon Genetics Europe, Cat. No. NG002R).
- Centrifuge Sorvall™ Legend™ Micro 17R (Thermo Fisher Scientific, Cat. No. 75002440).
- Oven (Sheldon Manufacturing, Inc., Cornelius, OR USA, Cat. No. SHO1).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).
- Membrane MF-Millipore, 0.45µm (Merck KGaA, Cat. No. HAWP04700).
- Fluorescence microscope Nikon Eclipse 50i (Nikon Inc, Tokio, Japan).

Procedure for preparing blood cells

1. Prepare the 1X PBS, KCl 57mM, Carnoy solution and1% HCl solutions (See Supplementary data 7)
2. Place 3.5mL of blood in a falcon tube and add 500µL of 1X PBS.
3. Slowly add 4mL Ficoll 1077, avoiding the mix. Centrifuge at 700 x g for 30 minutes.
4. Harvest the ring with the lymphocytes without touching the Ficoll using a sterile pipette tips. Collect in a falcon tube and add 15mL of 1X PBS.
5. Centrifuge at 900 rpm for 20 minutes.
6. Discard the supernatant. Resuspend the pellet in 5mL supplemented RPMI medium (See preparation on Supplementary data 7). Place it in a culture flask.

7. Incubate the lymphocytes at 37°C for 3 days.
8. Add 50µL colcemid solution (10µg/mL) and place the flask at 37°C for 20 minutes in the culture oven.
9. Transfer the lymphocytes to a falcon tube and add 50mL 1X PBS.
10. Centrifuge at 900 rpm for 15 minutes.
11. Discard the supernatant, except 500µL for resuspending the pellet. Add 15mL KCl 75mM.
12. Incubate at 37°C for 20 minutes in the water bath.
13. Centrifuge at 900 rpm for 15 minutes.
14. Discard the supernatant, except 500µL for resuspending the pellet. Slowly, drop by drop, add 2mL Carnoy solution. Faster, add another 8mL. Tap the tube continuously for mixing.
15. Incubate for 10 minutes at room temperature.
16. Centrifuge at 900 rpm for 15 minutes.
17. Repeat steps 15 to 17 three more times. In the last step resuspend pellet in 2mL Carnoy solution.
    At this point metaphases can be stored for 2-3 days at 4°C, or for longer periods at -20°C.
18. Clean the slides as shown in Supplementary data 7.
19. Collect 500µL metaphases. With the slide over the steam of the water bath at 100°C, pour three drops of the metaphases. Drops should be poured at a distance of 60-70cm.
20. Dry the slide upon and below the sample with the Bunsen lighter. Place the slide in the thermal plate at 45°C until it completely dries.
21. Leave the preparations at 4°C before use.

Procedure for FISH

1. Prepare 20X SSC, 2X SSC/0.1% NP-40 wash solution, denaturation solutions, 0.4X SCC/0.3% NP-40 Wash solution and ethanol solutions at 70%, 85% and 100%.
2. Prepare the target and probe mixture
    - Place Coplin jar containing denaturation solution to a water bath at 73°C.
    - Immerse the slides in the denaturation solution for 5 minutes.
    - Immerse the slides in 70% ethanol for 1 minute, followed by 1 minute in ethanol 85% and 1 minute in ethanol 100% to dehydrate. Keep slides in 100% ethanol until probe mixture is ready.
    - For each target, prepare the following:
        - 7µL LSI/WCP hybridization buffer
        - 1µL probe
        - 2µL Milli-Q water

- Centrifuge tube for 3 seconds, vortex and centrifuge the tube again.
- Place the tube in the water bath at 73°C for 5 minutes.
- Add denaturation solution and 70mL of each ethanol dilutions to different Coplin jars and bring to room temperature.

3. Hybridize the probe
    - Prepare a humidified box: Place a humidified paper towel on an airtight container at 37°C.
    - Remove the slides from 100% ethanol and place them on a 45 to 50°C slide warmer for up to 2 minutes.
    - Apply 10µL of probe mixture and put the coverslip.
    - Place slides in the pre-warmed humidified box and place in the incubator at 37°C overnight.
4. Washing and detection
    - Pour 70mL of 0.4X SCC/0.3% NP-40 solution into a Coplin jar and place in a 73°C water bath at least for 30 minutes before use.
    - Pour 70mL of 2X SCC/0.1% NP-40 into a Coplin jar and bring to room temperature.
    - Remove the coverslip from the slide and immerse the slide in the 0.4X SCC/0.3% NP-40 solution. Agitate slide for 3 seconds and immerse for 2 minutes.
    - Immerse the slide in the 2X SCC/0.1% NP-40 solution. Agitate slide for 3 seconds and immerse for 1 minute.
    - Dry slide in darkness.
    - Apply 10µL of DAPI II counterstain and put coverslip.
5. Visualize results on the fluorescence microscope.

The optimal visualization of the probe allows viewing the orange and green fluorophores. For male specimens with a 46, XY karyotype is one orange signal on the Y chromosome and one green signal on the X chromosome. The signal pattern of one green signal on each X chromosomes with lack of orange signal is seen in 46, XX karyotype specimens.

7.9. BIOINFORMATIC ANALYSIS

The sequencing data were analysed using the Ion Reporter software (Thermo Fisher Scientific), which comprises a group of bioinformatics tools that simplifies the study. Base calling, read filtering, alignment to the human genome and variant filtering were done with Ion Reporter Suites. The trimmed reads were aligned using BWA-MEM (Burrows-Wheeler Aligner Memory) algorithm followed by Genome Analysis ToolKit (GATK) for further base quality score analysis and variant filtering. Data generated on

the Ion PGM system were automatically uploaded from the Torrent Browser to our Ion Reporter software for read mapping, annotation and reporting of variants.

Data analysis was automated with a predefined workflow in which annotation sources about the biological meaning of variants for priorization were selected, as well as the alignment to the human genome (hg19), the analysis of a single sample and germline variant detection was set.

Two workflows were applied to each sample. First workflow was done for variant calling of single-nucleotide variants (SNV) and small insertions or deletions (indels) and a second one was used for copy number variations (CNV). Then, filter chains were created based on Ion Reporter Software variant data types.

7.9.1. Variant filtering for point mutations and small indels

To identify the DSD causative variant in each patient we filtered as follows:

- Variants with insufficient sequence coverage, depth and a Phred-like quality score below 20 were excluded. For evaluating the coverage and depth the variants were visualized using the Integrative Genomics Viewer (IGV) (http://www.broadinstitute.org/igv/). In those cases in which the amplicon was not covered or the depth was low (<20x fold) validation by Sanger sequencing was done. Phred-like score measures the quality of the identification of the bases generated by DNA sequencing or the confidence in the assignment of each base call by the sequencer. Phred-like score >20 means a 99% accurate assignment of each base, with a 1% chance of error. Quality metrics of each run and library was generated and evaluated.
- Variants were filtered to include those with a p-value <0.001 and a Minor Allele Frequency (MAF) <0.05 in the annotation settings of the Ion Reporter software, namely dbSNP, ClinVar and 5000 exomes databases. Moreover, the allele frequency was further checked for each ethnic group in 1000 genomes browser (http://browser.1000genomes.org/), Exome Aggregation Consortium (ExAC) Consortium (http://exac.broadinstitute.org/), Genome Aggregation Database (GnomAD) (https://gnomad.broadinstitute.org/) and Exome Sequencing Project (ESP) Variant Server (http://evs.gs.washington.edu/EVS/) and variants were discounted if they were common. Thus, we filtered to include only rare variants with MAF≤0.001. Moreover, we identified repeated variants within our cohort that may result from amplification or sequencing errors. These, were also discarded if found in more than 4 samples in the same run or in the total cohort, as well as if they are found in different phenotypes. Then, gene changes in the untranslated or intronic regions were discarded, with the exception of previously recorded changes or splice site variants. Finally, a

genotype/phenotype correlation was done, where the DSD variant is consider for further curation if it fits with the clinical presentation of the patient.

7.9.2. Variant filtering for CNV

The detection of CNV is dependent on several parameters. Control samples need to be included in the study as CNV will be reported based on their copy number relative to these samples. Previously studied samples will be used to make a baseline and thus, it is desirable to select those samples with not known CNV in any region covered by the panel. The karyotype of the patient is also important because copy number on X and Y chromosomes will be determined relative to its presence.

We included 4 samples (GN0038, GN0109, GN0141 and RE0045) to make the baseline and used the Confident Germline CNV filter added to filter chains in Ion Reporter Software. This filter returns CNV with high confidence levels, between 10 and 10,000,000. Moreover, we chose a MAPD (Median of the absolute value of all pairwise differences) score >0.4 as the cut-off.

7.9.3. Categorization of the variants

Variants were categorized based on the recommendations of the American College of Medical Genetics (ACMG) for variant classification and reporting (229) into 5 categories: Pathogenic, likely pathogenic, VUS (Variant of Unknown Significance), likely benign and benign. We used the following rules:

- Variants previously reported to cause disease in HGMD, OMIM, ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) or PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) were noted as reported, and were considered pathogenic if described in one or more unrelated individuals. They were reported as likely pathogenic if the same variant had been associated to a different DSD-phenotypes or if no functional studies had been done.
- Variants with MAF ≤0.1% that disrupt gene function, such as nonsense, frameshift, splice sites, initiation codons and exon deletion, were classified as pathogenic when the mechanism of pathogenicity is consistent with the inheritance pattern, previous DSD-associated variants have been described in the gene and the variant in a splice site leads to exon skipping or shortening.
- Novel missense variants with MAF ≤0.1%, with evidence for pathogenicity according to *in silico* programs and that demonstrated to be *de novo* in patient with the disease and no family history were classified as pathogenic.
- A novel missense variant with MAF≤ 0.1% involving an amino acid or nucleotide change known to be pathogenic, but changes to a different amino acid or nucleotide were classified as likely pathogenic. If the mechanism of

pathogenicity is well known to be caused by altered protein function it was classified as pathogenic.

- Novel missense variants with MAF ≤0.1% in known DSD genes and evidence for pathogenicity according to *in silico* programs were considered as likely pathogenic.
- Novel single variants with MAF ≤0.1% were considered VUS when contradictory evidence of pathogenicity according to prediction software.
- Variants previously reported, but with contradictory evidence in bibliography were considered VUS.
- Variants with an allele frequency >0.1% and synonymous variants that had not been previously reported to have a deleterious effect were classified as benign.

In this work, we only present those variants classified as pathogenic, likely pathogenic and VUS, as well as likely benign polymorphism previously related to DSD.

7.9.4. Variant effect prediction software tools

The most direct method to identify causal variations in an individual is by comparison with known mutations and disease-associated genes, like those referred in databases such as the HGMD (230), OMIM (231), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), UCSC (University of California Santa Cruz Genomics Institute) (https://genome.ucsc.edu/), DGV (Database of Genomic Variants) (http://dgv.tcag.ca/dgv/app/home) and dbVar (Database of human genomic structural Variation) (https://omictools.com/dbvar-tool). This is of utility in well studied diseases and can be used to find novel and known mutations in previously identified genes.

The remaining variants that have not been previously described can be prioritized. Multiple algorithms have been developed for predicting deleteriousness based on different information of the variant including its sequence homology, protein structure and evolutionary conservation. Although there are many, we have used the following prediction software tools.

7.9.4.1. Interpretation of single-nucleotide variants and small indels.

- PROVEAN (http://provean.jcvi.org/index.php)

PROVEAN (Protein Variation Effect Analyzer) is an informatics tool developed to predict whether a protein sequence variation affects protein function. It introduces a delta score based on the reference and variant versions of a protein query sequence with respect to sequence homologs collected from the NCBI NR (National Centre for Biotechnology Information Non-Redundant) protein database through BLAST (Basic Local Alignment Search Tool).

The delta score is computed for each supporting sequence and then averaged within and across clusters to generate the final score. The protein variant is predicted to have a deleterious effect if this score is equal or below the default threshold, set at -2.5.

- SIFT (http://sift.jcvi.org/)

SIFT (Sorting Intolerant From Tolerant) predicts whether an amino acid substitution affects protein function. It is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences, collected through PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool) and calculates the probability of the amino acid being tolerated based on observed alignment.

- Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/)

PolyPhen-2 (Polymorphism Phenotyping v2) is a tool for the prediction of the possible impact of an amino acid substitution on the structure and function of a human protein using features that comprise the sequence, phylogenetic and structural information.

The user can choose between HumDiv and HumVar models. HumDiv dataset was compiled from the damaging alleles with known effects on the molecular function causing human Mendelian diseases, present in the UniProtKB (UniProt Knowledgebase) database. HumVar consists of human disease-causing mutations together with common nsSNPs (Non-synonymous SNPs) without involvement in disease and therefore treated as non-damaging. Both models were considered in this work.

For a variation, PolyPhen-2 calculates the probability that this mutation is damaging and is evaluated qualitatively, as benign, possibly damaging, or probably damaging.

- MutationTaster (http://www.mutationtaster.org/)

MutationTaster integrates information from different biomedical databases and uses established analysis tools. Analyses comprise evolutionary conservation, splice-site changes, loss of protein features and changes that might affect the amount of mRNA (Messenger RNA). It employs Bayes classifier to predict the disease-potential of an alteration and calculates a probability value to be either a disease mutation or a harmless polymorphism. For this prediction, the tool studies the frequencies of known disease mutations or polymorphisms from HGMD and the 1000 genome project.

We have used the Mutation Taster for the comparison of the proteins for non-synonymous changes among different species.

- SNPs and GO (http://snps.biofold.org/snps-and-go/snps-and-go.html)

It is a GO (Gene Ontology) integrated predictor that uses a SVM (Support Vector Machine) classifier that takes in input protein sequence, profile and functional information to predict if a given variation can be classified as disease-related or neutral.

- MutPred (http://mutpred1.mutdb.org/)

MutPred (Mutation Prediction) is a web application tool developed to classify an amino acid substitution as disease-associated or neutral in human. Moreover, it predicts the molecular cause of disease of the substitution based upon the gain or loss of 14 different structural or functional properties, for example, gain of helical propensity.

- PANTHER (http://www.pantherdb.org/)

The PANTHER (Protein Analysis Through Evolutionary Relationships) system classifies protein according to molecular function, biochemical process, pathway and family or groups of evolutionarily related proteins.

- VarSome (https://varsome.com/)

This engine is used as an annotation tool, search instrument for genomic variants and also as a platform that enables the sharing of information (232). The variant of interest is mapped to a specific genomic location, identifying equal variants, the variant type and the coding effect. Variant pathogenicity is given using an automatic classifier that evaluates the submitted variants according to the American College of Medical Genetics (ACMG) guidelines (229).

7.9.4.2. Interpretation of variants located on the splicing sites

Alternative splicing is a process during gene expression in which particular exons of a gene are included or not in the final mRNA from that gene. This process results in a single gene coding for multiple proteins, that contain differences in their amino acid sequence and therefore, in their biological functions.

We used the Neural Network Splice Site Prediction (http://www.fruitfly.org/seq_tools/splice.html) and NetGene2 Server (http://www.cbs.dtu.dk/services/NetGene2/) to deduce the effect of the variants located in the boundaries of the exons. These programs compare the structure of the

donor and acceptor sites in the wild type (wt) to the newly altered gene and stablish if these structures have suffered any modification.

7.9.4.3. Interpretation of genomic structural variations

The analysis of structural variations was firstly performed with the CytoGenomics software v2.7, which applies automated filters based on several databases, such as the Database of Genomic Variants (DGV) and polymorphisms of Agilent Reference DNA.

The DGV (http://dgv.tcag.ca/dgv/app/home?ref=) provides structural variations identified in healthy control samples. It is continuously updated and aiming to correlate with phenotypic data.

Structural variations were also compared in dbVar, the NCBI's database of human genomic structural variation, ClinVar, which aggregates information about genomic variation and its relation to health and UCSC, a collection of organism assemblies and annotations.

7.10. *IN VITRO* FUNCTIONAL AND EXPRESSION STUDIES

Likely pathogenic or variants of unknown significance were selected and functionally characterized in different cell systems. In order to test the transcriptional activity of identified variants, several experiments were carried out including the site-directed mutagenesis, cloning and purification and the functional characterization of the variant.

The *in vitro* functional studies were developed in the Paediatric Endocrinology, Diabetology and Metabolism Department of BioMedical Research in the University Hospital Inselspital in Bern (Switzerland) under the guidance of Prof. Dr. med. Christa Flück.

7.10.1. Used vectors for the *in Vitro* functional studies

7.10.1.1. Expression vectors

To assess the impact of the identified variants, we used previously constructed mammalian expression vectors by cloning the gene cDNA into the corresponding vectors. 3 different expression vectors were used (Figure 15):

- pCMV: Derived from a high-copy-number pUC-based plasmid. Mammalian expression is driven by the human cytomegalovirus (CMV) promoter to promote constitutive expression of cloned inserts in a wide variety of cell lines. Selection is made possible in bacteria by the ampicillin-resistance gene under control of the β-lactamase promoter.
- pSG5: This ampicillin resistant vector was constructed by combining pKCR2 and pBS vectors. The SV40 early promoter and polyadenylation signal promotes expression in vivo.
- pcDNA3: It has a CMV enhancer-promoter for high-level expression and ampicillin resistance gene, as well as a pUC origin for selection and maintenance in *E. coli.*

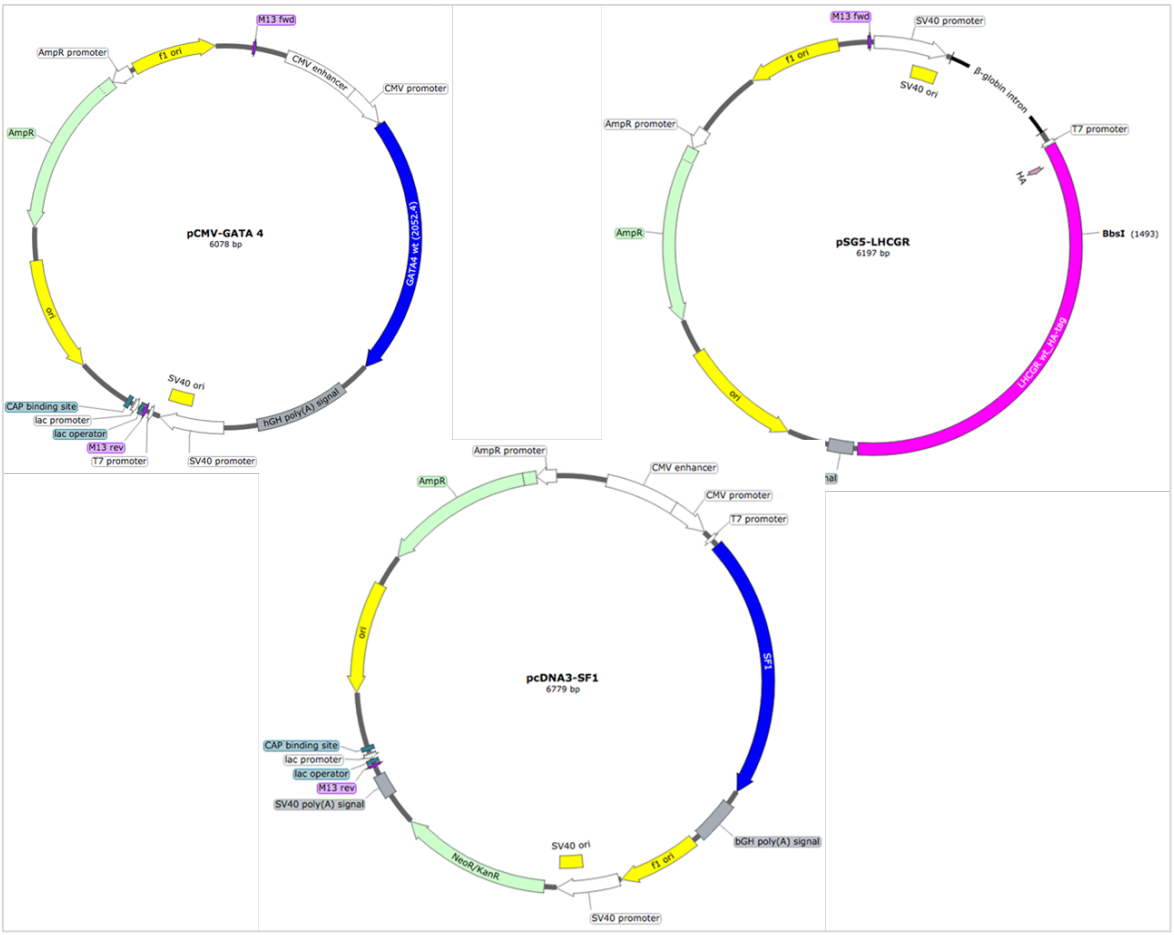

Figure 15. Image of the three expression vectors pCMV, pSG5 and pcDNA3 where cDNA from GATA4, LHCGR and NR5A1 was inserted.

Human wild type (wt) GATA4 cDNA was inserted into the mammalian pCMV expression vector. For the LHCGR expression vector we used the HA-tagged (Human influenza Hemagglutinin) wt LHCGR_pSG5 kindly gifted by Dr. Adolfo Rivero-Müller (Institute for Biomedicine, University of Turku, Turku, Finland) and has been previously

described (233). Whereas, wt cDNA of NR5A1 in a pcDNA3 vector with an HA-tag added at 5‘ by PCR strategy was available from previous work (14).

GATA4, LHCGR and NR5A1 mutants were generated by site-directed mutagenesis using specific primers and the QuickChange protocol by Stratagene (Agilent technologies Inc., Santa Clara, CA, USA).

7.10.1.2. Promoter reporter vectors

Mammalian expression vectors were tested on different responsive promoters. These reporter vectors included a gene described as being regulated by the genes of interest (*GATA4* or *NR5A1*) linked to a firefly-luciferase (luc) and inserted in a plasmid. The following reporter vectors were used in this work:

- -227CYP17A1_ Δluc
- -152CYP11A1_pGL3
- -301HSD3B2_pGL3
- CRE-luc

The pGL3 promoter vector is an ampicillin resistant vector which contains an SV40 promoter upstream of the luciferase gene for monitoring transcriptional activity in transfected eukaryotic cells. *CYP11A1* and *HSD3B2* genes were inserted in this pGL3 vector. *CYP17A1* was inserted into the reporter plasmid pMG3 luciferase. The CYP17A1, CYP11A1 and HSD3B2 luciferase reporter vectors were used in either the co-transfection of GATA4 or NR5A1 expression vectors.

On the other side, CRE-luc reporter was used in the co-transfection of either the wild type LHCGR expression vector, the empty vector (pSG5) or the mutant LHCGR vectors. CRE-luc reporter contains the firefly luciferase gene under the control of cAMP (Cyclic adenosine monophosphate) response element (CRE), located upstream of a promoter. Elevation of the intracellular cAMP level activates cAMP response element binding protein (CREB) to bind CRE and induces the expression of luciferase.

The construction and cloning of the promoter luciferase reporter vectors has been previously described and were available from earlier work (14, 24, 117).

7.10.1.3. Renilla luciferase vectors

To normalize the transfection efficiency, we added a pRL vector that contains wt Renilla luciferase (Rluc). This provides constitutive expression of Renilla luciferase and then, if used in combination with a firefly luciferase vector to cotransfect

mammalian cells, provides an internal control value to which expression of the experimental firefly luciferase reporter gene can be normalized.

pRL vectors contain the cDNA encoding Renilla luciferase cloned from the *Renilla reniformis*. We used the pRL vectors with the HSV-thymidine kinase promoter (pRL-TK) and the CMV promoter region (pRL-CMV) (Figure 16). Both were isolated from an *E. coli* K host strain and are ampicillin resistant. pRL-TK was used to normalize the transfection efficiency of the LHCGR and NR5A1 promoter transactivation assays while pRL-CMV was used for the GATA4 functional studies. These vectors were available from previous works (14).

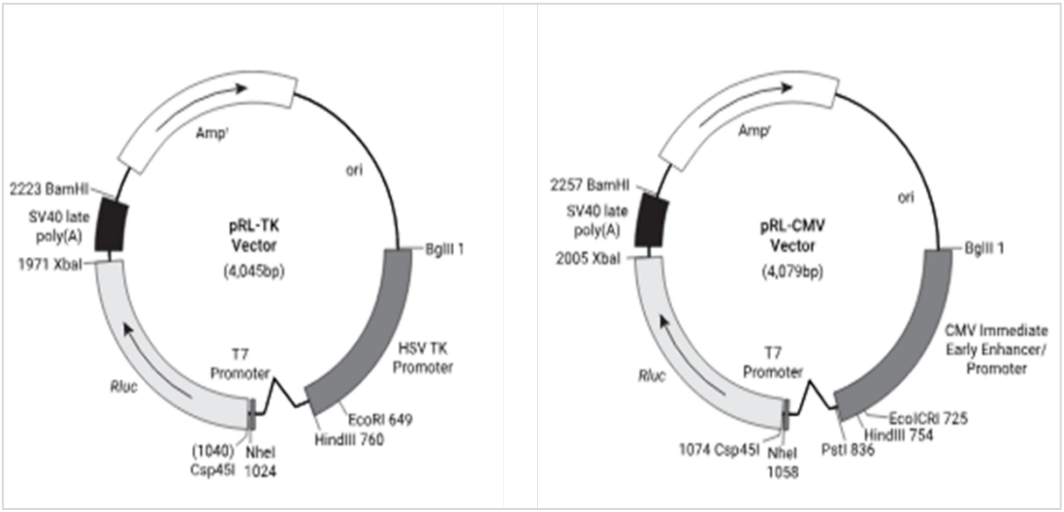

Figure 16. pRL-TK and pRL-CMV vectors used as Rluc control reporter vectors. Modified from Promega (https://www.promega.es/products/reporter-assays-and-transfection).

7.10.2. From Site-directed mutagenesis to plasmid purification

The procedure requires the synthesis of a DNA primer containing the desired mutation, complementary to the template DNA, so it can hybridize with the DNA in the gene of interest. Then, the single-strand primer is extended using a DNA polymerase, which copies the rest of the gene. During PCR, the mutation is incorporated into the amplicon, replacing the original sequence. The template DNA is removed using a methylation-dependent endonuclease, such as DpnI, and the mutated gene is introduced to a host cell (*E. coli*) as a vector and cloned. Plasmids are isolated (Miniprep) from the resulting colonies and both mutants and wild type (wt) were checked by DNA sequencing to ensure that they contain the desired mutation or the wt DNA. These clones were used for all downstream procedures. They were again amplified in *E. coli* and purified using a maxiprep kit. The purified plasmids were then used for cell transfection.

7.10.2.1. Primers design and site-directed mutagenesis

Site-directed mutagenesis is an *in vitro* method used to make specific and intentional changes to the DNA sequence of a gene. Primers containing the change of interest were designed using the Agilent tool (www.genomics.agilent.com/primerdesignprogram) and ordered to Microsynth AG Company (https://www.microsynth.ch/home-ch.html) (Balgach, Switzerland). We designed primers (Supplementary data 6) for the following sequence changes:

- LHCGR: c.757C and c.1660T variants
- GATA4: c.677T variant
- NR5A1: c.88A, c.902A and c.71T variants.

Mutant expression vectors (GATA4: c.677T; LHCGR: c.757C and c.1660T; NR5A1: c.88A, c.902A and c.71T) were generated by PCR-based site-directed mutagenesis (Figure 17) using specific primers and the QuickChange protocol by Stratagene (Agilent Technologies Inc.) using the wild type expression vector as template.

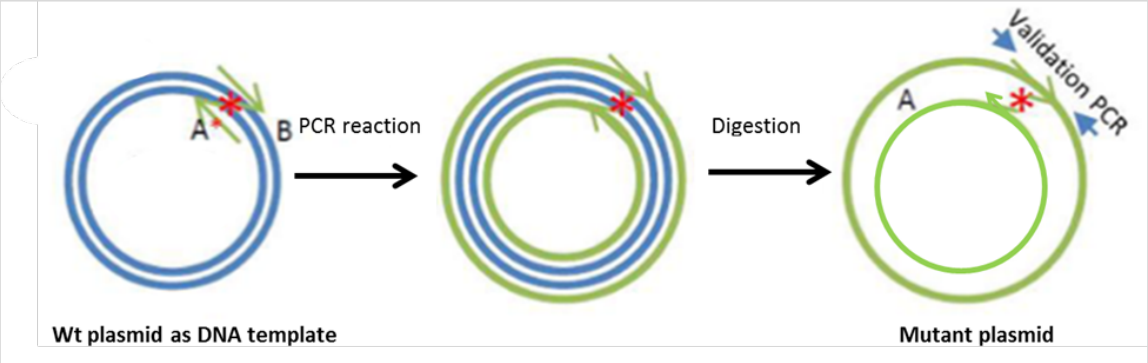

Figure 17 Diagram of the site-directed mutagenesis method. Modified from Addgene (https://blog.addgene.org/).

Reagents and material

- QuikChange II Site-Directed Mutagenesis Kit (30 Rxn) (Agilent Technologies, Inc, Cat. No 200524) containing Pfu Ultra High-Fidelity DNA polymerase (2.5 U/µll) (80 U), 10× reaction buffer (500µl), DpnI restriction enzyme (10 U/µl) (300 U) and dNTP mix (30µl).
- Primers for each mutant (20µM) (Microsynth AG).
- Plasmid (*GATA4*-CMV5 plasmid 128,5ng/µL; *LHCGR*-pSG5 plasmid 415,4ng/µL; *NR5A1*-pcDNA3 plasmid 206,8ng/µL), obtained from previous works.
- Bidistilled water.
- T100™ Thermal Cycler (Bio Rad, Cat. No. 1861096).

Procedure for site-directed mutagenesis

1. Mutant Strand Synthesis Reaction
    - Prepare the sample reaction by mixing 5µL 10X Buffer, 250ng plasmid, 1µL forward primer (20µM), 1µL forward primer (20µM), 1µL dNTPs (10mM), 1µL Pfu DNA polymerase (2,5U/µL) and bidistilled water up to 50µL.
    - Place in the T100™ Thermal Cycler and perform as follows:
        - 95°C for 30 seconds.
        - Cycle 12 times: 95°C for 30 seconds, 55°C for 60 seconds and 68°C for 8 minutes.
        - Hold at 4°C.
2. DpnI digestion of the amplification product
    - Add 1µL DpnI restriction enzyme to the sample tube.
      Include also a positive and a negative control. Prepare negative control with 10X Buffer, plasmid, DpnI and bidistilled water up to 50µL and positive control with 2µL plasmid and bidistilled water up to 50µL.
    - Mix by pipetting up and down.
    - Incubate the reaction at 37°C for 1 hour.
    - Proceed to transformation of *E. coli* cells or store the product at -20°C.

7.10.2.2. Plasmid transformation and purification

Transformation is the process by which foreign DNA is introduced into a host cell. Transformation of bacteria with plasmids is widely used due to the bacterial capacity of storage and replication but also because plasmids carry both the bacterial origin of replication and an antibiotic resistance gene, which is a selective marker in bacteria.

Some bacterial strains have been created to be more easily transformed, as DH5α which is the most frequently used *E. coli* strain for routine cloning applications. In order to make bacteria capable of internalizing the exogenous DNA, they must be made competent. Cells are made competent by treating them with divalent cations, such as calcium chloride ($CaCl_2$) or magnesium chloride ($MgCl_2$), which make the bacterial wall permeable to DNA, facilitating genetic transformation.

The method for transformation of a DNA construct into a cell is chemical and calcium chloride transformation is the most efficient technique. The addition of calcium chloride to the cell suspension permits the binding of plasmid DNA to bacterial cell membrane, although both are negatively charged, and when heat is provided, plasmid DNA passes into the cell. Once plasmid DNA is inside the cell, it will be replicated as bacteria are grown. Finally, purification of the plasmid DNA is required to

isolate and purify plasmid DNA from genomic DNA. The purified plasmids will be used for cell transfection.

We followed the next steps to get the purified plasmid DNA:

- Transformation of DH5α cells
- Growth of bacterial culture (small-scale growing)
- Plasmid DNA purification (Miniprep)
- Quantification of the recombinant DNA
- Sequencing to ensure that the plasmid DNA has inserted the change
- Growth of bacterial culture (large-scale growing)
- Plasmid DNA purification (Maxiprep)
- Quantification of the recombinant DNA

7.10.2.2.1. Transformation of E. coli competent cells

We used the transformation storage solution (TSS) buffer method to make DH5α cells competent. Competence of cells is induced by the addition of low concentrations of Mg^{2+}, which alters the permeability of the membrane, and DMSO, acting as a preserving agent. Polyethylene Glycol (PEG) shields the negative charges on the DNA and host cell membrane, reducing the repulsion between them. For the transformation, we prepared a KCM buffer with both $CaCl_2$ and $MgCl_2$ divalent cations for altering the permeability of the bacterial membrane

Reagents and materials

- DH5α cells (From previous works).
- LB (Luria-Bertani) broth (Luria/Miller) (1kg), (Carl Roth Gmbh + Co., Karlsruhe, Germany, Cat. No. X968.1), containing Tryptone 10g/L, Yeast extract 5g/L, Sodium chloride (NaCl) 10g/L.
- LB Agar (Luria/Miller) (1kg) (Carl Roth Gmbh + Co., Cat. No. X969.1), containing Tryptone 10g/L, Yeast extract 5g/L, NaCl 10g/L, Agar 15g/L.
- Ampicillin, disodium salt (10g) (Carl Roth Gmbh + Co., Cat. No HP62.1).
- TSS Buffer (See preparation on Supplementary data 7).
- 5X KCM Buffer (See preparation on Supplementary data 7).
- Bidistilled water.
- Orbital shaker (Thermo Fisher Scientific, Cat. No. SHKE420HP).
- Incubator (Eppendorf AG, Cat. No CO17321001).

Procedure for preparing *E. coli* competent cells

1. Prepare LB medium and LB agar plates (Supplementary data 7)
2. Inoculate the cells in 10mL LB medium and grow overnight at 37°C.

3. Dilute 0.5mL of the culture into 50mL of fresh LB medium in a 200mL conical flask. Grow at 37°C the diluted culture to an OD600 of 0.2–0.5.
4. Cold the TSS buffer and the microcentrifuge tubes on ice.
5. Split the culture into two 50mL falcon tubes and incubate on ice for 10 minutes.
6. Centrifuge cells for 10 minutes at 3000 rpm at 4°C.
7. Discard the supernatant and resuspend the pellet in chilled TSS buffer. The volume of the TSS to use is 10% of the culture volume that was discarded.
8. Resuspend the culture by vortexing.
9. Aliquot 200µl into the chilled microcentrifuge tubes. Store at -80°C until use.

Procedure for transformation of *E. coli* competent cells

1. Prepare LB medium and LB agar plates (Supplementary data 7).
2. Pre-warm LB medium and LB plates at 37°C.
3. Defrost DH5α competent cells on ice.
4. Prepare DNA mixture by mixing 58µL bidistilled water, 20µL 5X KCM Buffer (Supplementary data 7) and 2µL plasmid DNA.
5. Chill DNA mixture on ice for 10 minutes.
6. Add 100µL DH5α competent cells to DNA mixture and place on ice for 20 minutes.
7. Add 1mL pre-warmed LB medium and incubate at 37°C with shacking for 1 hour.
8. Centrifuge for 4 minutes at 200-300 rpm at room temperature. Discard the supernatant, except 300µL.
9. Spread cells with 300µL LB into an LB plate.
10. Incubate the plate at 37°C for 24 hours. Then store at 4°C until use.

7.10.2.2.2. Growth of bacterial cultures

LB (Luria-Bertani) broth has been widely used for the growth of bacteria in molecular applications. These bacterial media are nutrient-rich formulations which provide peptides, peptones, vitamins, and trace elements, that permit the fast and easy growth of many species, including *E. coli*. Different formulations have been created to vary in the amount of sodium chloride and, therefore, in the osmotic conditions for the particular bacterial strain and culture conditions.

We used the Luria-Miller formulation with a high salt content, ideal for propagation of *E. coli* and culturing cells for plasmid preparation. Only those bacterial colonies that have incorporated the plasmid and thus the resistance to the antibiotic will grow on the LB agar plate. Various colonies of recombinant and wt plasmids were selected from the LB plates and were grown to get numerous copies of plasmid DNA.

Reagents and materials

- LB broth (Luria/Miller) (1kg), (Carl Roth Gmbh + Co., Cat. No. X968.1), containing Tryptone 10g/L, Yeast extract 5g/L, Sodium chloride (NaCl) 10g/L.
- Ampicillin, disodium salt (10g) (Carl Roth Gmbh + Co., Cat. No HP62.1).
- Magnesium sulfate ($MgSO_4$) (100g) (Carl Roth Gmbh + Co., Cat. No. 0682.1).
- Autoclave (JP Selecta group, Abrera, Spain, Cat. No. 4002136).
- Bidistilled water.
- Orbital shaker (Thermo Fisher Scientific, Cat. No. SHKE420HP).
- Magnetic stirrer hotplate (Thermo Fisher Scientific, Cat. No. 13889336).

Procedure for small-scale growing

1. Prepare LB medium (Supplementary data 7).
2. Add 12mL LB medium to a 15mL tube.
3. Add 12µL ampicillin (100mg/mL) to the tube and mix.
4. Distribute 3mL of the LB medium into 4 tubes of 15mL.
5. Add 6µL $MgSO_4$ (1M) to each tube.
6. Collect one single colony and mix it with the culture. Repeat it for the four tubes.
7. Incubate the cell culture at 37°C with constant shaking (250 rpm) for 12-16 hours.

   Store the bacterial cultures at 4°C until use.

Procedure for large-scale growing

1. Prepare 500mL LB medium. Mix in a magnetic mixer for 1 hour.
2. Distribute 100mL to 4 flask and autoclave on liquid cycle at 121°C.
3. Cool to 60°C for 15 minutes.
4. Dilute ampicillin (100mg/mL) to 1/1000. Add 100µL to each flask.
5. Inoculate 200µL of the bacterial growth (obtained at the small-scale growing step). Cover with foil.
6. Incubate at 37°C for 12-16 hours with constant shaking at 230-240 rpm.

7.10.2.2.3. Plasmid DNA purification

We used the NucleoSpin® Plasmid QuickPure kit (Macherey-Nagel GmbH & Co.KG, Düren, Germany) which permits to achieve up to 15µg of plasmid DNA and the GenElute™ HP Plasmid Maxiprep Kit (Merck KGaA) to isolate large-scale pure plasmid DNA.

Plasmid DNA miniprep purification was performed to verify that the mutation had been inserted in the plasmid. This confirmation was done by sequencing.

Reagents and materials

- NucleoSpin® Plasmid QuickPure kit (50 preps) (Macherey-Nagel GmbH & Co.KG Cat. No. 740615.50), containing Resuspension Buffer A1 (15mL), Lysis Buffer A2 (15mL), Neutralization Buffer A3 (20mL), Wash Buffer AQ (6mL), Elution Buffer AE (13mL), RNase A (lyophilized) (6mg), NucleoSpin® Plasmid QuickPure Columns (50), Collection Tubes (2mL) (50).
- GenElute™ HP Plasmid Maxiprep Kit (Merck KGaA, Cat. No. NA0310) containing, Column Preparation Solution (225mL), RNase A Solution (2.5mL), Resuspension Solution (375mL), Lysis Solution (375mL), Neutralization Solution (375mL), Binding Solution (280mL), Wash Solution 1 (375mL), Wash Solution 2 (75mL), Elution Buffer (115mL), GenElute HP Maxiprep Filter Syringe (25 units), GenElute HP Maxiprep Binding Column (25 units), Collection Tubes 50mL (50 tubes).
- Absolute ethanol (Merck KGaA, Cat. No. 1009831000).
- Centrifuge Sorvall ST 8 (Thermo Fisher Scientific, Cat. No. 75007202).
- Vortex mixer (Heidolph Instruments GmbH & Co.KG, Cat. No. 541-10000-00).
- Heating-block (Thermo Fisher Scientific, Cat. No. 88880028).

Procedure for miniprep

1. Cultivate and harvest bacterial cells
    - Add 2mL of the saturated *E. coli* LB culture to a tube and centrifuge at 13000 rpm for 30 seconds at room temperature. Discard the supernatant and remove as much liquid as possible.
2. Cell lysis and clarification of lysate
    - Add 250µL Buffer A1. Resuspend the cell pellet completely by vortexing.
    - Add 250µL Buffer A2. Mix by inverting the tube 6–8 times.
    - Incubate at room temperature for 5 minutes.
    - Add 300µL Buffer A3. Mix by inverting the tube 6–8 times.
    - Centrifuge for 5 minutes at 13000 rpm at room temperature.
3. Bind DNA
    - Place a NucleoSpin® Plasmid QuickPure Column in a 2mL Collection Tube and pipet a maximum of 750µL of the supernatant onto the column.
    - Centrifuge for 1 minute at 13000 rpm. Discard flow-through and place the NucleoSpin® Plasmid QuickPure Column back into the collection tube.
    - Repeat this step to load the remaining lysate.
4. Wash silica membrane
    - Add 450µL Buffer AQ and centrifuge for 3 minutes at 13000 rpm at room temperature. Discard the collection tube and the flow-through.
5. Elute DNA

- Place the NucleoSpin® Plasmid QuickPure Column in a 1.5mL microcentrifuge tube
- Add 50µL Buffer AE and incubate for 1 minute at room temperature.
- Centrifuge for 1 minute at 13000 rpm.
- Pipet the 50µL of plasmid DNA into a new 2mL tube.
  Store at 4°C until use.

Procedure for maxiprep

1. Harvest Cells
   - Transfer the bacterial culture to proper centrifuge tubes and centrifuge at 4600 rpm for 20 minutes at 4°C.
   - Discard the supernatant.
2. Resuspend and lyse cells
   - Add 12mL Resuspension/RNase A Solution to the bacterial pellet. Resuspend by pipetting up and down. Place to 50mL tubes.
   - Add 12mL Lysis Solution and mix by inverting 6 to 8 times. Let the mixture sit for 5 minutes until it becomes clear
3. Prepare Filter Syringe
   - Prepare a filter syringe by removing the plunger and placing the barrel in a rack so that the syringe barrel is upright.
4. Neutralize and add binding solution
   - Add 12mL Neutralization Solution to the mixture and invert 4 to 6 times.
   - Add 9mL Binding Solution and invert 1 to 2 times.
   - Immediately pour into the barrel of the filter syringe.
   - Allow the lysate to sit for 5minutes.
5. Prepare Binding Column
   - Place a GenElute HP Maxiprep Binding Column onto the vacuum manifold and apply vacuum.
   - Add 12mL of Column Preparation Solution to the column and allow it to pass through.
   - Centrifuge for 2 minutes at 3000 x g at 4°C.
6. Washing
   - Add 12mL of Wash Solution 1 to the column.
   - Centrifuge for 2 minutes at 3000 x g at 4°C. Discard the eluate.
   - Add 12mL of Wash Solution 2 to the column.
   - Centrifuge for 5 minutes at 3000 x g at 4°C.
7. Elute Plasmid DNA
   - Transfer the binding column to a clean 50mL collection tube.
   - Add 3mL of Elution Solution or molecular biology reagent water to the column.

- Centrifuge for 5 minutes at 3000 x g at 4°C. Discard the column. Store the plasmid DNA at 4°C until use.

7.10.2.2.4. Plasmid DNA quantification

Plasmid DNA quantification was performed after miniprep and maxiprep. We used a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific).

7.10.2.2.5. Plasmid DNA sequencing

Microsynth AG Company (https://www.microsynth.ch/home-ch.html) performed the Sanger sequencing of the plasmid DNA to ensure that the site-directed mutagenesis worked properly and the variant of interest was introduced in the plasmid.

7.10.3. Functional characterization of the variants of interest

7.10.3.1. Cell lines and culture

Human placental JEG3 cells (CLS, 300222) and non-steroidogenic human embryonic kidney HEK293 cells (ATCC CRL-1573) (http://www.lgcstandards-atcc.org/) (Figure 18) were cultured and used for functional assays.

JEG 3 cells are clonally derived lines isolated from human choriocarcionoma (Woods strain of the Erwin-Turner tumor) by Kohler *et al* (234). These cells release human chorionic gonadotrophin (hCG) and somatomammotrophin and progesterone to media and are able to transform steroid precursors to estrone and estradiol. On the contrary, HEK293 cells are derived from human embryonic kidney (235).

JEG3 cells were cultured in MEM (Minimum Essential Medium), supplemented with 10% Fetal Bovine serum (Thermo Fisher Scientific), 1% L-glutamin serum (Thermo Fisher Scientific) and 1% Penicillin/streptomycin serum (Thermo Fisher Scientific). HEK293 cells were cultured in DMEM (Dulbecco's Modified Eagle Medium), supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin and 1% sodium pyruvate (Thermo Fisher Scientific) supplemented media.

Reagents and materials

- JEG3 cells (CLS Cell Lines Service GmbH, Eppelheim, Switzerland, Cat. No. 300222).
- HEK293 Cells (American Type Culture Collection, Manassas, VA, USA, Cat. No. 1573).

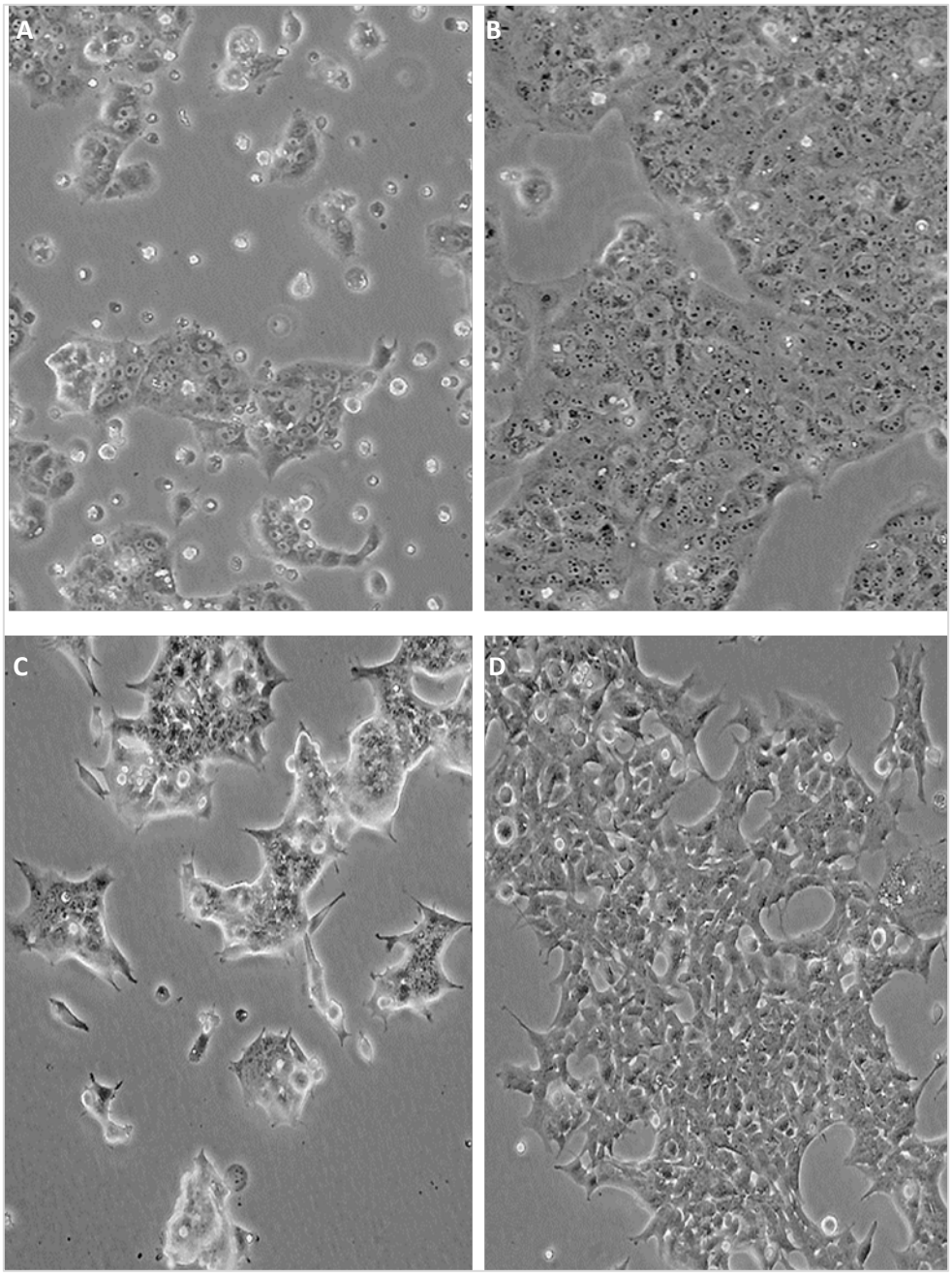

Figure 18. JEG3 cell lines cultured at low (A) and high density (B). Human embryonic kidney HEK293 cells cultured at low (C) and high density (D).

- MEM medium (500mL) (Thermo Fisher Scientific, Cat. No. 21090-022).
- DMEM medium (500mL) (Thermo Fisher Scientific, Cat. No. 41965039).
- Fetal bovine serum (1L) (Thermo Fisher Scientific, Cat. No. 10106185).
- Penicillin-Streptomycin (10,000 U/mL) (100mL) (Thermo Fisher Scientific, Cat. No. 15140-122).
- L-Glutamin (100mL) (Thermo Fisher Scientific, Cat. No. 25030081).
- Sodium pyruvate ($C_3H_3NaO_3$) (100mM) (100mL) (Thermo Fisher Scientific, Cat. No. 11360070).
- Trypsin-EDTA (0.05%) (100mL) (Thermo Fisher Scientific, Cat. No. 25300-054).

- Cell culture hood (Class II cabinet, Bio II Advance, Azbil Telstar, Terrassa, Spain).
- Incubator (Eppendorf AG, Cat. No CO17321001, Hamburg, Germany).
- Water bath (JP Selecta group, Abrera, Spain, Cat. No. 032012).
- Centrifuge Sorvall ST 8 (Thermo Fisher Scientific, Cat. No. 75007202).
- Vacuum aspiration system (Integra Biosciences™, Hudson, NH, USA, Cat. No. 158320).
- Microscope (Carl Zeiss AG, Oberkochen, Germany).

Procedure for thawing frozen cells

1. Prepare supplemented media (Supplementary data 7).
2. Remove the cryovial containing the frozen cells from storage and place into a 37°C water bath.
3. Pre-warm media at 37°C.
4. Use a small crystal pipet to transfer the vial content to a flask and dilute with 14mL of media.
5. Incubate the culture at 37°C in the incubator at 5% CO_2 for 4 days.

Procedure for passaging cells

1. Pre-warm supplemented media and trypsin-EDTA at 37°C.
2. Remove and discard the cell culture media from the culture vessel with a crystal pipet and vacuum.
3. Add 2.5mL trypsin-EDTA to the vessel and incubate at 37°C for 5 minutes.
4. Prepare 2 new culture vessels by adding 14mL media.
5. Observe the cells under the microscope for detachment.
6. Add 7.5mL media to the vessel and pipet up and down.
7. Pipet the appropriate volume into new cell culture vessels so that cell suspension is diluted.
8. Incubate at 37°C in the incubator at 5% CO_2 until cells are ready to be subcultured.

7.10.3.2. Promoter transactivation experiments

Transfection is the delivery of genetic material into eukaryotic cells to study the regulation of the genes, protein expression and function. It can be accomplished using various methods and reagents. These methods include physical (i.e. electroporation), chemical (i.e. liposome-mediated) or viral-based (i.e. adeno-associated virus) delivery systems. Chemical transfection methods are the most widely used methods.

For functional studies, cells were cultured on 12 or 24-well plates and transfected with wild type (wt) or mutant vectors and a promoter reporter using either a calcium phosphate transfection protocol (Thermo Fisher Scientific) or Lipofectamine®

2000 Transfection Reagent (Thermo Fisher Scientific). 48h after transfection, cells were washed and lysed, before luciferase activity was measured with the Dual-Luciferase® Reporter (DLR™) Assay system (Promega AG, Wallisellen, Switzerland) on a Veritas microplate Luminometer reader (Turner Bio systems Luminometer and Software by Promega). Specific Firefly luciferase readings were standardized against Renilla control readings and results were expressed as relative luciferase units.

We used calcium phosphate for the cotransfection of the vectors containing wt or mutant GATA4 and LHCGR and promoter reporters. When mixing DNA with calcium chloride to a buffered saline solution, a fine precipitate is formed and added to the surface of the cells. The cell will take up the precipitate by endocytosis. On the other hand, lipofection uses a cationic lipid to form a vesicle with the DNA and easily merge with the cell membrane. This liposome transfection was used for the transactivation studies of the wt and alterations found in the *NR5A1* gene.

Reagents and materials

- Calcium phosphate transfection Kit (Cat. No. K2780-01, Thermo Fisher Scientific), containing Tissue Culture Sterile Water (2 × 12mL), 2X Hepes Buffered Saline (HBS) (2 × 12mL) and 2 M Calcium phosphate ($CaCl_{2)}$ (3 × 1mL).
- Lipofectamine® 2000 Transfection Reagent (1.5mL) (Cat. No. 11668019, Thermo Fisher Scientific).
- MEM medium (500mL) (Thermo Fisher Scientific, Cat. No. 21090-022).
- DMEM medium (500mL) (Thermo Fisher Scientific, Cat. No. 41965039).
- Opti-MEM™ I Reduced Serum Medium (500mL) (Thermo Fisher Scientific, Cat. No. 31985047).
- Fetal bovine serum (1L) (Thermo Fisher Scientific, Cat. No. 10106185).
- Penicillin-Streptomycin (10,000U/mL) (100mL) (Thermo Fisher Scientific, Cat. No. 15140-122).
- L-Glutamin (100mL) (Thermo Fisher Scientific, Cat. No. 25030081).
- Sodium pyruvate (100mM) (100mL) (Thermo Fisher Scientific, Cat. No. 11360070).
- Trypsin-EDTA (0.05%) (100mL) (Thermo Fisher Scientific, Cat. No. 25300-054).
- Trypan Blue (50mL) (Merck, Cat. No. 93595).
- Chorionic Gonadotropin, human (1mg) (BioVision, Inc., Milpitas, CA, USA, Cat. No. 4778-1000).
- Cell culture hood (Class II cabinet, Bio II Advance, Azbil Telstar, Terrassa, Spain).
- Incubator (Eppendorf AG, Cat. No CO17321001, Hamburg, Germany).
- Water bath (JP Selecta group, Abrera, Spain, Cat. No. 032012).
- Automated cell counter (Cat. No. 145-0011, Biorad).
- Vacuum aspiration system (Integra Biosciences™, Hudson, NH, USA, Cat. No. 158320).

Procedure for calcium phosphate transfection

- 24 hours before transfection, trypsinize cells and culture to be 50% confluent.
- Prepare transfection mixture for each variant. To a tube labeled A, add the corresponding volume of plasmid DNA (500ng/well), empty vector (750ng/µL), promoter luciferase reporter vector (750 ng/well), pRL Renilla Luciferase Control Reporter Vector (25ng/well), 15µL 2 M $CaCl_2$ and sterile water up to 150µL.
- To a tube labelled B add 150µL 2X HBS.
- Drop by drop add tube A to tube B.
- Incubate at RT for 30 minutes.
- Add 100µL to each well as it corresponds.
- Incubate for 5 hours at 37°C.
- Pre-warm medium at 37°C.
- Remove medium and discard it. Add 1mL new medium to each well.
- Incubate for 48 hours at 37°C.

Procedure for lipofection

- 24 hours before transfection, trypsinize cells and culture to be 80% confluent.
- Prepare DNA/ Opti-MEM™ solution for each variant. To a labeled tube, add the corresponding volume of plasmid DNA (125ng/well), empty vector (125ng/µL), promoter luciferase reporter vector (500ng/well), pRL Renilla Luciferase Control Reporter Vector (25ng/well), and 50µL Opti-MEM™.
- Prepare lipofectamin/ Opti-MEM™ solution for each well by mixing 5µL lipofectamin to 45µL Opti-MEM™.
- Add 110µL lipofectamin/ Opti-MEM™ solution to each DNA/ Opti-MEM™ solution.
- Incubate the DNA-lipid complex for 5 minutes at room temperature.
- Trypsinize and culture cells in 1mL medium in 24-well plates.
- Add 100µL DNA-lipid complex per well as it corresponds.
- Incubate for 48 hours at 37°C.

7.10.3.2.1. hCG stimulation on LHCGR promoter transactivation experiments

Subsequent experiments were performed to assess the cAMP produced by the HEK293 cells under both basal and human Chorionic Gonadotropin-stimulated conditions after the transfection with the different *LHCGR* variants.

To determine the cAMP reporter response (CREB) on HEK293 cells transfected with *LHCGR* wt and mutant, we added increasing concentrations of human Chorionic

Gonadotropin 48 hours after transfection had started. We prepared the following concentrations: 0ng/mL, 3ng/mL, 30ng/mL, 300ng/mL, 1000ng/mL and 3000ng/mL and added 5µL to each corresponding well. We incubated overnight and measured luciferase.

7.10.3.2.2. Luciferase assay

The luciferase reporter assay (Figure 19) is commonly used as a tool to study gene expression at the transcriptional level. Firefly luciferase enzyme is able to emit light via a chemical reaction in which luciferin is converted to oxyluciferin by the luciferase enzyme. Some of the energy released by this reaction is in the form of light. To perform the reporter assay, the regulatory region of the gene of interest has been cloned upstream of the luciferase gene in one expression vector. After transfection and cell lysate has been collected, luciferase activity is measured by adding luciferin and other necessary cofactors using a luminometer. Since the gene of interest is fused to the luciferase reporter gene, the luciferase activity is directly correlated with its activity.

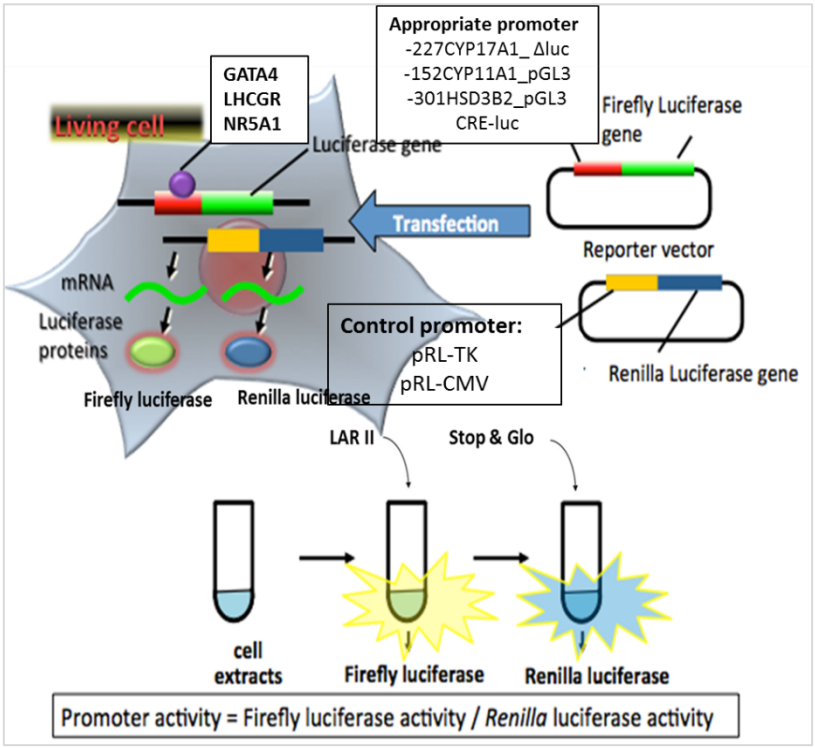

Figure 19. Picture of the principles of a dual luciferase reporter assay. Modified from Ohmiya Y, Applications of bioluminescence (http://photobiology.info/Ohmiya.html).

Reagents and materials

- Dual-Luciferase® Reporter (DLR™) Assay system (Promega AG, Wallisellen, Switzerland, Cat. No (E1910), containing Luciferase Assay Substrate (1 vial),

Passive Lysis Buffer (PLB), 5X (30mL), Luciferase Assay Buffer II (10mL), Stop & Glo® Substrate (200µl) and Stop & Glo® Buffer (10mL).

- Rocking Platform Shaker (VWR, Radnor, PA, USA, Cat. No. 10860-662).
- 96-well plate (Thermo Fisher Scientific, Cat. No. 136101).
- Veritas microplate Luminometer reader (Promega AG, Cat. No. E4861).
- Veritas™ Software (Promega AG).

Procedure for Dual luciferase reporter assay

1. Prepare Passive Lysis Buffer 1X from PLB 5X. Place on ice until use.
2. Remove media from all the wells.
3. Add 150µL PLB 1X to the wells.
4. Place the culture plate on a rocking platform for 15 minutes at room temperature.
   At this point plate could be stored at -20°C.
5. Prepare Stop & Glo solution for each well by mixing 1µL Stop & Glo substrate 50X with 49µL Stop & Glo Buffer. Place on ice until use.
6. Prepare LAR II solution for each well by transferring 50µL of the Luciferase Assay Buffer II to a 15mL tube.
7. Pipette up and down the lysate to obtain a homogeneous suspension and add 50µL to a 96-well plate for luminometer.
   Store the rest of the lysate at -20°C.
8. Read the plate on the Veritas microplate luminometer reader following manufacturer's instructions.

7.10.3.2.3. Protein expression study

For protein detection and quantification western blot technology was used. From a complex mixture of proteins extracted from transfected cells the specific protein will be identified through a separation by size, transference to a solid support and marking a target protein using primary and secondary antibodies to visualize.

HEK293 cells were transfected with either wt or mutant NR5A1 expression vector, which carried a HA-Tag. Cells were lysed and western blot was performed using an antibody against HA (Human influenza hemagglutinin)-Tag (Merck KGaA, Cat. No. H3663-100UL). Expression of β-actin protein was used as a control. Steroidogenic Factor-1 (SF1) expression bands should be visualized at 53KDa and the positive control at 42KDa.

Reagents and materials

- Protease Inhibitor Cocktail (Merck, Cat. No. 11836153001) (See preparation on Supplementary data 7).
- EDTA (Fluka Analytical, St. Gallen, Switzerland, Cat. No. 3610).
- Tris ($C_4H_{11}NO_3$) (Merck, Cat. No. 8382).
- Triton X-100 ($C_{14}H_{22}O(C_2H_4O)_9$) (Merck, Cat. No. T8787-100ML).
- DC™ protein assay Kit II (Bio-Rad Cat. No. 5000112), containing DC™ Protein Assay Reagent A, DC™ Protein Assay Reagent B and DC™ Protein Assay Reagent S.
- Bovine Serum Albumin (1g) (Roth AG, Zürich, Switzerland, Cat. No.80763).
- Tween 20 ($C_{58}H_{114}O_{26}$) (50mL) (Merck, Cat. NoP9416).
- Milk (350g) (Rapilait).
- β-Mercaptoethanol (C_2H_6OS) (100mL) (Merck, M6250-100ML).
- Sodium chloride (NaCl) (1Kg) (Merck, Cat. No. S9888) (mw: 58.4g/mol).
- Potassium chloride (KCl) (Merck, Cat. No. P9333).
- Disodium hydrogen phosphate (Na_2HPO_4) (Merck, Cat. No. NIST2186II).
- Potassium dihydrogen phosphate (KH_2PO_4) (Merck, Cat. No. NIST200B).
- Tris-MOPS-SDS Running Buffer Powder (5u) (GenScript, Cat. No. Cat. No. M00138).
- Transfer Buffer Powder (10u) (GenScript, Cat. No. Cat. No. M00139).
- Electrophoresis marker (Bio-Rad, Cat. No. 161-0374).
- Methanol (CH_3OH) (1L) (Merck, Cat. No. 34860-1L-R).
- Monoclonal Anti-HA antibody produced in mouse (1.0mg/mL) (Merck, Cat. No. H3663-100UL).
- IRDye® 800CW Goat anti-Mouse IgG (Immunoglobulin G) (H + L) (1.0mg/mL), (LI-COR Biotechnology GmbH, Bad Homburg, Germany, Cat. No. P/N 926-32210).
- IRDye® 680RD Donkey anti-Mouse IgG (H + L) (1.0mg/mL), (LI-COR Biotechnology GmbH, Cat. No. P/N 925-68072).
- Anti-β-Actin antibody, mouse monoclonal (1.0mg/mL) (Merck, Cat. No. A1978-100UL).
- 1X Odyssey Blocking Buffer (500mL) (LI-COR Biotechnology GmbH, Cat. No. P/N 927-40000).
- Tris-buffered saline with Tween® 20 (TBS-T) (10 tablets) (Merck, Cat. No. 91414-10TAB).
- GenScript ExpressPlus™ PAGE Gels (20u) (GenScript, Piscataway, NJ, USA, Cat. No. M41212).
- Centrifuge Sorvall ST 8 (Thermo Fisher Scientific, Cat. No. 75007202).
- ELISA Microplate Reader (Molecular Devices LLC, San Jose, CA, USA, Cat. No. 0200-2013).

- 96-well plate (Thermo Fisher Scientific, Cat. No. 136101).
- Gel tank (Bio-Rad).
- Heating-block (Thermo Fisher Scientific, Cat. No. 88880028).
- Cassette Opener (GenScript, Cat. No. L00674).
- Blot Absorbent Filter Paper (Bio-Rad).
- Blotting system fiber pad (Merck, EP1504-8EA).
- PVDF Western Blotting Membranes (Merck, Cat. No. 3010040001).
- Rocking Platform Shaker (VWR, Radnor, PA, USA, Cat. No. 10860-662).
- LI-COR Odyssey Imaging system (LI-COR Biotechnology GmbH).
- SoftMax® Software (Molecular Devices LLC).

Procedure for Western Blot

1. Measurement of protein concentration
    - Prepare 1X Lysis Buffer for each well by mixing 15µL 10X Protein Inhibiting Cocktail with 135µL Protein Lysis Buffer.
    - Remove media from all the wells.
    - Add 150µL 1X Protein Lysis Buffer to the wells.
    - Homogenize the lysate with a syringe and transfer to 1.5mL tubes.
    - Centrifuge at 13200 rpm for 10 minutes at 4°C.
    - Prepare standards by diluting Lysis Buffer in BSA to obtain concentrations that range from 5.6mg/mL to 0.0875mg/mL.
    - Prepare master mix for each well by mixing 0.5µL DC™ Protein Assay Reagent S with 24.5µL DC™ Protein Assay Reagent A.
    - Transfer 5µL of the samples, blank and standards to the 96-well plate.
    - Add 25µL master mix to the wells.
    - Add 200µL DC™ Protein Assay Reagent B to the wells.
    - Incubate plate for 20 minutes at room temperature.
    - Read by Softmax at an absorbance of 650nm and calculate protein concentration.
2. Protein denaturation
    - Prepare the protein denaturation mix to have 50µg of total protein by mixing protein with 5X β-mercaptoethanol and bidistilled water up to 50µL.
    - Incubate at 95°C for 10 minutes and spin down
3. Gel electrophoresis
    - Prepare 1X MOPS and Transfer buffer (Supplementary data 7).
    - Remove the comb of the already prepared gel and place in the gel tank.
    - Fill the gel tank with 1X MOPS buffer
    - Load 24µL of the sample and 5µL of electrophoresis marker.
    - Run gel at 140V for 45 minutes.

- Open the cassette and remove the gel.

4. Wet transfer
    - In a tray, pre-wet for 10 minutes 4 filter papers and 2 fiber pads.
    - Pre-wet PVDF membrane in methanol for 2 minutes.
    - Prepare the sandwich on the black panel as follows: one fiber pad, two filter papers, SDS gel, PVDF membrane, two filter papers and one fiber pad.
    - Fill the gel tank with Transfer buffer.
    - Run at 100V for 90 minutes.
5. Blocking and incubating with antibodies
    - Prepare Blocking buffer, Washing buffer and primary antibody dilution buffer (Supplementary data 7).
    - Incubate membrane with 10mL Blocking buffer on a shaker for 2 hours.
    - Add 2µL primary antibody, Anti HA produced in mouse, to 4mL primary antibody dilution buffer
    - Immerse the membrane in the primary antibody dilution buffer and incubate overnight at 4°C on a shaker.
    - Remove the buffer and wash with washing buffer in the shaker for 5 minutes.
    - Repeat previous step 3 more times.
    - Dilute 1µL secondary antibody, Goat anti-mouse, in 10mL Blocking buffer.
    - Immerse the membrane in the buffer and incubate for 1 hour at RT on the shaker.
    - Remove the buffer and wash with washing buffer in the shaker for 5 minutes.
    - Repeat previous step 3 more times.
    - Incubate membrane with B-actin antibody diluted in 5%BSA for 2 hours in a shaker.
    - Remove the buffer and wash with 1X TBS-T in the shaker for 5 minutes.
    - Repeat previous step three more times.
    - Dilute 1µL secondary antibody, Donkey anti-mouse, in 10mL 0.5X Odyssey Blocking Buffer.
    - Immerse the membrane in the buffer and incubate for 2 hour at room temperature on the shaker.
    - Wash as previously described.
6. Visualize membrane in the LI-COR Odyssey Imaging system at an absorbance of 680 and 800nm.

7.10.4. Statistical analysis

Obtained data from the repeated luciferase assays were summarized giving the mean±SEM (Standard error of the mean). Data were statistically analysed using Student's t test (Microsoft Excel, Microsoft Corporation, Redmond, WA, USA). Results were considered statistically significant when $P<0.05$.

RESULTS

Hereafter, I will present the results from the clinical and molecular characterization of our DSD cohort. Firstly, a general description of the population is made. Subsequent to the validation of the targeted gene panel, I present the genetic results obtained either by a gene candidate approach or next generation sequencing (NGS). Finally, results from the *in vitro* studies are shown.

8. GENERAL CHARACTERISTICS OF THE DSD COHORT

A total of 125 independent patients with a DSD diagnosis were analysed in this study. We have classified the cohort of patients according to the consensus statement on management of intersex disorders published in 2006 (172). The various aetiologies of DSD are shown in Supplementary data 2.

Out of the 125 patients that were included, 99 had 46,XY DSD and 24 had 46,XX DSD. We also identified a sex chromosome DSD in 2 patients (Supplementary data 2). The age of presentation in our series varied from newborn to 41 years and represented individuals from 14 populations, which have been grouped into European Caucasian (Spain, Switzerland and Rumania) (110 cases, 88.0%), Northern African (Morocco, Algeria, Tunisia) (3 cases, 2.4%), Sub-Saharan African (Mauritania, Sierra Leone, Madagascar, Equatorial Guinea) (8, 6.4%), Asian (Syria, China, India) (3, 2.4%) and American (Mexico) (1, 0.8%).

It is important to note that persons diagnosed with congenital adrenal hyperplasia (CAH) due to steroidogenesis defects, such as 21-hydroxylase or 11-hydroxylase deficiency were not included in this study. We only examined patients with non-CAH primary adrenal insufficiency (PAI).

Family members from these 125 individuals were analysed once a genetic variant explaining the phenotype of the index case was identified. This comprised 97 family members that were studied. However, samples of first degree family members of 35 probands were not available and segregation studies could not be made.

8.1. 46,XY DSD PATIENTS

Based on chromosomal karyotyping and clinical presentation, 99 individuals were classified as having 46,XY DSD in which external genitalia were not clearly male or female (See Figure 20, page 142). The most common cause at clinical presentation was the disorders of testicular development, diagnosed in 47 patients. Testicular dysgenesis was suspected in 43 individuals and among the 17 patients with complete gonadal dysgenesis, female external and internal genitalia, primary amenorrhea and

lack of secondary sex characteristics was mainly reported. Out of them, three presented a clitoral hypertrophy, although imaging revealed uterus or vagina. Twenty-two patients with partial gonadal dysgenesis presented with a wide range of genital ambiguity. Four patients were referred without a classified gonadal dysgenesis. On the other side, three patients had gonadal regression with male genitalia and non-localized gonads after laparoscopy, and one case was reported as an ovotesticular DSD.

On the other side, disorders of androgen synthesis or action were referred in 32 patients. Six individuals had a suspected androgen biosynthesis defect. Of these, five were assigned as males and were born with severe hypospadias, normal testes and in two cases curved penis. Four patients out of these were diagnosed based on hCG-stimulated testosterone to dihydrotestosterone (DHT) ratio. On the other side, one patient was reared as female and was virilised at puberty.

Androgen insensitivity syndrome (AIS) was suspected in 21 individuals. Among the ten patients with complete AIS (CAIS), four had female genitalia and primary amenorrhea during adolescence. Inguinal hernias or palpable gonads were reported in 3 newborns. Another three cases were suspected to be CAIS after amniocentesis had revealed a disagreeing 46,XY karyotype in females. All the infants were born with female external genitalia and inguinal gonads. Four PAIS patients presented with gynecomastia or different degrees of genital undervirilization, such as micropenis with hypospadias in some cases, and hypertrophic erectile organ. Other seven cases were not classified as complete or partial due to the absence of clinical information.

Furthermore, we had 5 male patients presenting with precocious puberty, adrenarche or bilateral gynecomastia during childhood that led to a suspected LH receptor defect.

Among the cases that were classified as other 46,XY DSD, two males were reported to have isolated hypospadias and descended scrotal testis while seven patients were diagnosed with a non-CAH primary adrenal insufficiency and were initially referred to clinician with varying degrees of symptoms such as respiratory distress, vomiting and asthenia, hypoaldosteronism, salt wasting and hyponatremia. The age of presentation of these ranged between newborn and 3 years. The phenotype was not determined in one case.

Additionally, 7 individuals were stated as 46,XY DSD but couldn’t be classified. They presented mainly with undervirilized phenotypes, such as cryptorchidism, micropenis or non-isolated hypospadias.

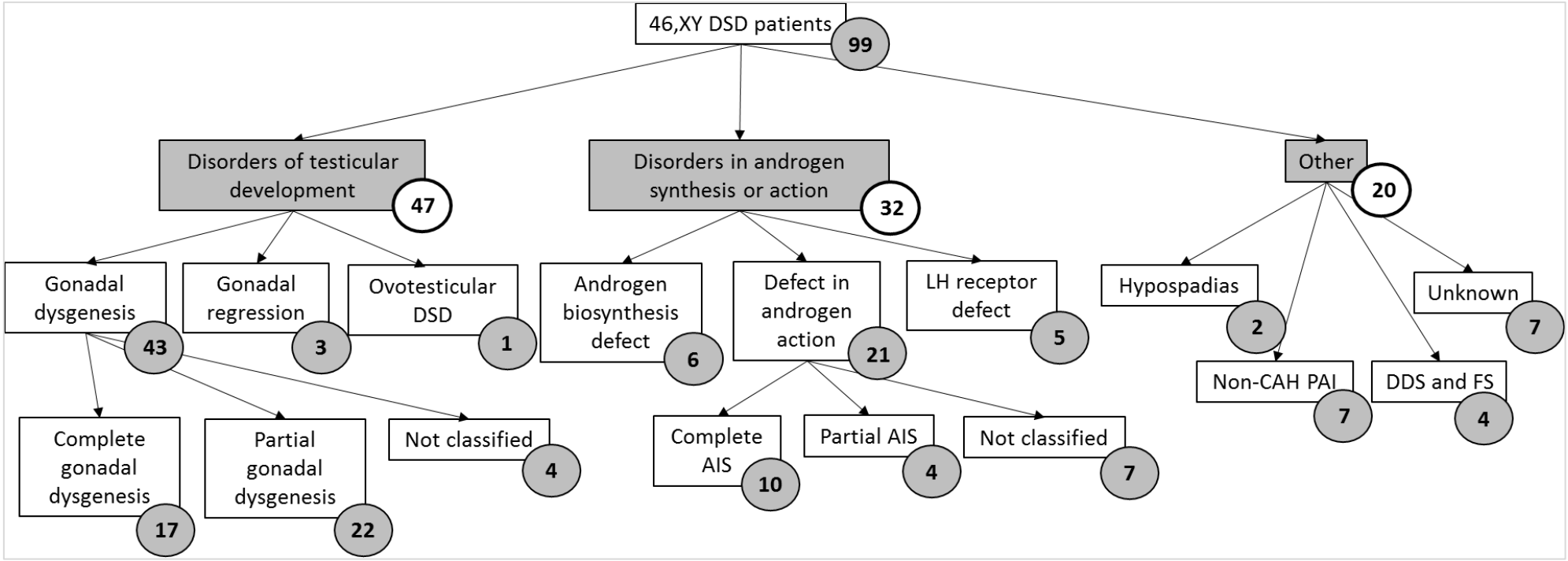

Figure 20. Classification of the 46,XY DSD patients based on the clinical diagnosis. Circles indicate the number of cases.
AIS, androgen insensitivity syndrome; CAH, congenital adrenal hyperplasia; DDS, Denys-Drash syndrome; FS, Frasier syndrome; PAI, primary adrenal insufficiency

8.2 46,XX DSD PATIENTS

We also studied 24 patients with 46,XX DSD (Figure 21). Among the patients with disorders of ovarian development, we tested 8 cases with suspected gonadal dysgenesis. These, largely presented female genitalia, primary amenorrhea, lack of secondary sexual characteristics or hypergonadotropic hypogonadism. One patient had bilateral inguinal hernias and neonatal ovaries were shown on probing. Three cases with ovotesticular (OT) DSD were included in the cohort. They were reared as males and presented with different degrees of genital ambiguity. Laparoscopy confirmed the presence of ovotestes in the infants. Moreover, six patients had testicular (T) DSD and denoted mainly cryptorchidism, as well as gynecomastia and azoospermia in some cases when diagnosed in adulthood. Finally, two out of three cases with a suspected Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome had primary amenorrhea at puberty, while one subject was diagnosed at childhood. Imaging revealed the absence of uterus, vagina and unilateral kidney agenesia in all the cases.

The disorder could not be defined in 4 cases. These, were referred with different phenotypes such as complete pubertal development and secondary amenorrhea, male external and internal genitalia or primary amenorrhea and delayed puberty in which maternal virilisation during pregnancy had occurred.

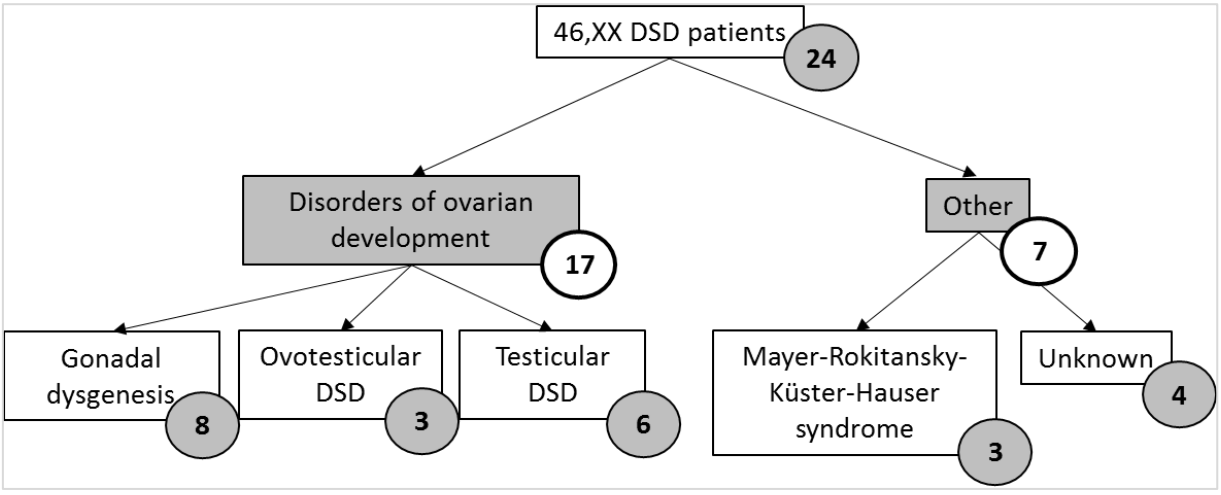

Figure 21. Classification of the 46,XX DSD patients based on the clinical diagnosis. Circles indicate the number of cases.

None of the 46,XX DSD patients was referred due to a disorder of androgen excess.

8.3 SEX CHROMOSOME DSD

Two patients had mixed gonadal dysgenesis. In the newborn case, labioscrotal fusion and inguinal hernia were recorded and genitoplasty was required. The second patient referred primary amenorrhea as a young adult. On Magnetic Resonance Imaging (MRI), a hypoplastic uterus was shown, but gonads were not visualized in any case. Karyotype was 45,X0/46,XY in the two subjects.

9. ANALYSIS OF THE TARGETED GENE PANEL

At the beginning direct Sanger sequencing was used for the molecular diagnosis of DSD patients. In 2015 we developed a customized gene panel and therefore, panel validation was required as a first step.

9.1. PANEL VALIDATION

9.1.1. Evaluation of quality metrics in each run and library

To validate the performance of the targeted panel sequencing we assessed the quality metrics given in each run (Table 7). Considering the 11 runs, the mean number of reads was 1123.2 Mbp and the read length was 242.1 bp. We examined the AQ20 score, which corresponds to the quality level at which the error rate is 1% or less, and found that the total number of bases and the mean segment length were 1023.9 Mbp and 232.6 bp, respectively, which means that 90.7% and 96% were high quality base reads and segments. The mean longest alignment was 487.5pb at AQ20.

Total number of reads varied between 532 Mega and 1.9 Giga base pairs. The difference in the number of reads among the runs is given due to the use of Ion 318™ Chips instead of Ion 316™ Chips since run 7. The greater capacity of the Ion 318™ Chip increased the total number of reads and the mean depth of all the libraries. The percentage of aligned bases to the reference sequence was 99% in all the runs, except in run 7, and accuracy was between 99.3% and 99.6%.

Table 7. Quality metrics given in the 11 runs performed for the sequencing of the DSD cohort.

Run	Number of bases	Mean read length (bp)	Aligned bases (%)	Accuracy (%)	Metrics at the quality level (AQ20)			Metrics of the libraries in each run			
					Number of bases (% bp)	Mean read length (% bp)	Longest alignment (% bp)	Mean depth	Uniformity (%)	On target (%)	Mapped reads
1	532 M	229	99	99.3	485 M (91.1)[1]	217 (94.7)[1]	470	101.4-524.6	86.1-91.1	94.5-96	112109-561952
2	691 M	223	99	99.3	606 M (87.7)[1]	211 (94.6)[1]	460	118.5-431.4	77.3-90.7	92.4-96.2	133276-491692
3	711 M	217	99	99.4	628 M (88.3)[1]	208 (95.8)[1]	460	148.8-354.8	87.4-90.4	94.7-95.9	166327-407373
4	871 M	241	99	99.5	786 M (90.2)[1]	232 (96.3)[1]	488	147.0-345.6	75.1-90.9	91.9-94.6	170072-355568
5	978 M	243	99	99.5	889 M (90.9)[1]	234 (96.3)[1]	510	124.0-413.4	63.97-89.6	93.1-95.2	148919-420999
6	812 M	238	99	99.4	729 M (89.8)[1]	231 (97.0)[1]	506	160.1-302.1	87.6-89.9	81.8-94.4	176079-345955
7	1.9 G	265	100	99.6	1.75 G (92.1[1])	258 (97.3)[1]	552	1200	89.49	92.4	1164105
8	1.26 G	231	99	99.4	1.14 G (90.5)[1]	220 (95.2)[1]	429	196.2-349.2	87.2-90.2	93.8-95.9	211364-380333
9	1.48 G	259	99	99.5	1.37 G (92.6)[1]	250 (96.5)[1]	491	162.8-575.7	86.2-89.6	93.9-96.3	153650-548956
10	1.64 G	264	99	99.5	1.52 G (92.7)[1]	256 (96.9)[1]	525	325.4-485.7	88.4-90.0	95.0-95.7	306575-456078
11	1.48 G	253	99	99.4	1.36 G (91.9)[1]	242 (95.6)[1]	472	370.3	89.8	96	359366
Mean	1123.2 M	242.1	99.1	99.4	1.02 M (90.7)[1]	232.6 (96.0)[1]	487.5	288.3	89.5	94.2	761735.5

bp, base pair; G, giga base pair; M, mega base pair; [1]Percentage of number of bases or mean read length at AQ20 quality level.

The mean depth analysis of the libraries showed that the occasions a given nucleotide in the genome was read varied between 101.4 and 1200. In addition, on average 94.2% of the targeted regions were covered and the distribution of the reads along targeted regions, or uniformity, was 89.5%. A single DSD library was included in a chip along with other disease related libraries in run 7. As the other samples created smaller libraries compared to DSD panel, this sample was overexpressed in the chip and the total number of bases and mean depth was higher in this run. In run 11, a single DSD library was also included but without overexpression of the sample, then quality metrics were in the normal range.

In supplementary data 10, run results from the genetically diagnosed patients and positive controls are shown as an example of the values.

9.1.2. Analytical validation of the targeted gene panel

We validated the gene sequencing panel by comparing test results against a gold standard that establishes the true status of the subject. This validation evaluated the ability of our panel to detect a variant in our DSD-cohort. For panel validation purposes, we made measures of sensitivity, specificity plus false positive rate, and finally, repeatability. Because the performance varies by the type of variant, sensitivity and false positives were calculated separately for single variants and small indels or CNV (236).

A selection of positive samples with mutation-specific assays for *SRY, HSD17B3, AR, NR5A1, WT1* and *NR0B1* during routine work, were included in the validation process. As shown in Table 8, positive controls involved different types of variants and were included in all the runs, except in runs 7 and 11 were only one DSD sample was included in the sequencing chip together with samples studied for other diseases. Moreover, we included 4 samples of healthy individuals, named true negatives. The next table (Table 8) summarizes the analytical values obtained in the panel validation.

Table 8. Summary of the analytical validation values obtained for the targeted gene panel.

	Single-nucleotide Variants (%)	**Copy number variations (%)**
Sensitivity	81.8	33.3
Specificity	100	
False positives	13.3	81.2
Repeatability	89.9	

9.1.2.1. Sensitivity

Sensitivity is defined as the ability to detect a positive result compared to a reference method. Sensitivity was calculated and compared with results obtained by Sanger sequencing for point mutations or Multiplex Ligation-dependent Probe Amplification (MLPA) for copy number variations (CNV).

For point mutation testing we determined 10 single-nucleotide variants (SNVs) and a 4-nucleotide deletion. Among the SNV, 7 were missense changes and 3 were deletions or insertions of 1 nucleotide. All the changes detected in positive controls were confirmed by gene sequencing panel, except c.182delC in *WT1* gene (Table 9), showing test sensitivity up to 81.8% (Table 8). The sensitivity yield was also tested for chromosome rearrangements. We compared deletions found by MLPA in samples POL0285 and GN0041, which was duplicated, with the results determined by NGS. Regarding POL0285, the workflow for CNV reported one variant which was the *NR0B1* gene deletion, while for GN0041 individual a single deletion was shown in one of the samples but did not correspond to the deletion of the *AR* exon 2. Then, sensitivity was 33.3% for the determination of CNV (Table 8).

Table 9. Sensitivity test done in positive controls.

Case	Run	Gene and variant	Zygosity	Detected variant
GN0012	Run 8	*NR5A1,* c.437G>C; p.Gly146Ala	Het	Yes
GN0038	Run 2	*HSD17B3,* c.845C>T; p.Pro282Leu	Het	Yes
GN0038	Run 5			Yes
GN0041	Run 3	*AR,* c.(1616+1_1617-1)_(1768+1_1767-1)del; p.(Arg539_Asp305del)	Hemi	Yes
GN0041	Run 6			No
GN0042	Run 4	*NR5A1,* c.614_615insC; p.Gln206ThrfsTer20	Het	Yes
GN0109	Run 3	*NR5A1,* c.910_913delGAGC; p.Glu304CysfsTer26	Het	Yes
GN0111	Run 9	*NR5A1, c*.902G>A; p.Cys301Tyr	Het	Yes
GN0119	Run 10	*NR5A1,* c.437G>C; p.Gly146Ala	Hom	Yes
GN0123	Run 10	*NR5A1,* c.71A>T; p.His24Leu	Het	Yes
GN0141	Run 1	*SRY,* c.391C>T; p.Pro131Ser	Hemi	Yes
OT0327	Run 4	*WT1,* c.182delC; p.Pro61ArgfsTer28	Het	No
OT0327	Run 5			No
POL0285	Run 6	*NR0B1, g*.(?_29978097)_(30361290_?)del	Hemi	Yes
Sane controls	Run 1	-	-	-

Hemi, hemizygous; Het, heterozygous; Hom, homozygous.

9.1.2.2. Specificity and false positives

Specificity is defined as the capacity of the method to correctly recognize the cases without the disease under study. It is the proportion of true negatives that are correctly identified by the test. In our 4-sample validation, we did not identify any DSD-related variant. Then, specificity was 100% (Table 8, page 146).

However, due to the limited number of healthy subjects that were included, we further analysed the false positive rate for the DSD-cohort. A false positive is the error in which a test result indicates the presence of a condition, when in reality it is not present. Among the 98 point mutations found by NGS that were then tested by Sanger sequencing, 13 were not confirmed, so the false positive rate was 13.3% (Table 8).

Validation of false positive rate for CNV was done considering samples (GN0034, GN0046 and OT0567) and positive controls (GN0041 and POL0285) in which a big deletion or insertion had been detected by NGS and had been confirmed by another method. Number of detected variants with the CNV workflow is shown in Table 10. The false positive rate for the detection of CNVs was 81.2% (Table 8).

Due to the low sensitivity and high false positive rate found in the CNV validation (Table 8), the results obtained in the analysis of these types of mutations were poorly reliable. In those cases in which a CNV was suspected a Multiplex Ligation-dependent Probe Amplification (MLPA), Quantitative Multiplex Polymerase chain reaction of Short Fluorescent Fragments (QMPSF) or array-based Comparative Genomic Hybridization (aCGH) was performed.

Table 10. Rate of false positives determined for copy number variations by NGS.

Case	Copy Number Variations		False positive rate
	Detected by NGS	False positive	
GN0034	1	0	
GN0041	0	0	
GN0041	1	1	
GN0046	1	1	
OT0567	12	11	
POL0285	1	0	
TOTAL	16	13	81.2%

9.1.2.3. Repeatability

This term refers to the ability to replicate the findings, calculated as the percentage of detected variants in a sample included in two separate runs. We did not separate the measure of repeatability by type of variant because the ion reporter does not distinguish between copy number variations (CNV) and single-nucleotide variants (SNV) or small indels when using the workflow for CNV detection without applying the confident germline filter.

We considered variants with a coverage depth ≥20x in three different samples and determined a mean repeatability of 89.9% (84.7% in GN0038 sample; 90.1% in OT0327 and 94.4% in GN0041) (Table 11).

Table 11. Detected variants (≥20x) in the three different samples used for the calculation of repeatability.

Case	Variants		Variants among the different runs		Repeatability (%)	Mean repeatability (%)
	1st run	2nd run	Repeated	Not repeated		
GN0038	230	228	210	38	84.7	
GN0041	250	243	241	13	94.9	89.9
OT0327	245	253	236	26	90.1	

9.1.3. Evaluation of the amplicons

After analysing the coverage data files, we observed a highly variable performance of the 874 amplicons in the panel. Besides the not covered regions *in silico* (Supplementary data 9), we evaluated the coverage of the amplicons in each sample using IGV (Integrative Genomics Viewer) and found a small number of them to which no reads were mapped in most of the studied samples (Supplementary data 11). We analysed the location of these amplicons and only those located in exons and exons boundaries were further studied via PCR-based Sanger sequencing (Supplementary data 12). In total, the surface that could not be read was 9.3%, then the design covered 90.7% of the amplicons. Those amplicons with a base coverage <20x were also sequenced.

9.2. VARIANT CALLING AND CHARACTERISTICS OF OBSERVED VARIANTS BY NGS

Before filtering, 27,444 point variants or small indels were observed in the entire population in the genes included in the panel (Supplementary data 13). Once the filtering was done (p-value <0.001 and MAF <0.05), 4,688 variants were still annotated. Of these, 36% were observed recurrently in the cohort and were understood as common population variants or as sequencing artefacts. In total, we found 34.5 observations on average for each patient. Remaining variants were curated based on the ACMG (American College of Medical Genetics) guidelines (229).

Regarding the type of variant, almost 58% (57.9%) were characterized as SNV (Single-Nucleotide Variants) or MNV (Multiple Nucleotide Variant) after filtering process. Only 10% of the protein changing variants were missense. Small indels were also found, mainly frameshift deletions (23.5%). Among the curated variants 29 were missense (82.8%), 3 were frameshift deletions (8.5%) and 3 variants were determined as nonsense, frameshift insertion and inframe deletion.

In the analysis of the CNVs, 29,885 variants were called and then reduced to 457 CNV after the application of the confident germline CNV filter. Although the number of chromosome rearrangements detected by the customized panel was significant, the panel validation showed low sensitivity and high false positive rate for the CNV. Big deletions and insertions were only trustworthy when they were not repeated among the samples and could be visually confirmed using IGV (Integrative Genomics Viewer) and the coverage data files. The 4 samples with CNV that were confirmed by aCGH, MLPA or FISH techniques, included the whole and partial deletions of *LHCGR, SRD5A2* and *WT1* genes, as well as the translocation of *SRY*. Only 3 of the CNV had been called by the Confident Germline CNV filter. As done previously, variants were annotated and classified following the ACMG criteria.

10. MOLECULAR CHARACTERIZATION OF THE DSD COHORT

Firstly, genetic analysis was done in patients presenting a sex development anomaly by studying the genes or regions that are most often associated with the disease by a candidate gene approach. In a second phase and after the implementation of next generation sequencing (NGS) in our laboratory, new patients and those with negative results in classic DSD-related genes were studied by a targeted panel.

10.1. SINGLE-GENE TESTING

As previously mentioned, gene by gene sequencing was done regularly for the genetic diagnosis of DSD patients. In 2015, NGS was implemented in our laboratory and single-gene tests, such as Sanger sequencing, MLPA or aCGH, were done depending on the judgement on the most likely genetic cause. As shown in Table 12, a total of 127 DSD-gene tests were requested for 81 patients, mostly the PCR-based sequencing of *SRY* and *AR* gene (34 tests, each), followed by the sequencing of the coding exons of *NR5A1* gene (21 tests).

From the aforementioned requests, only five Sanger sequencing tests were done after the targeted gene panel had been validated. Two of these genetic orders corresponded to the sequencing of the *NR0B1* gene in POL0285 and POL0301 cases, presenting with adrenal insufficiency. In the remaining cases *LHCGR* and *AR* genes were studied due to the suspicion of LH receptor defect (GN0164) or AIS (GN0202, GN0177). Due to syndromic features associated with DSD an aCGH was performed in case GN0159.

Finally, a plausible molecular finding was done in 36 patients by applying a candidate gene approach (36 out of 81, 44.4%) (Table 12). Among these, 13 patients had clinical features for an androgen insensitivity syndrome (Table 13, page 154) and sequencing of the *AR* gene was the prior genetic test in most cases. *SRY* gene missense variants or its translocation was found in 7 cases with 46,XX testicular DSD or 46,XY gonadal dysgenesis. The two coding exons of the *NR0B1* gene were studied in the 7 cases with an adrenal insufficiency diagnosis and the *WT1* gene in the two individuals with a suspicion of Frasier syndrome. Two activating alterations were found in the luteinizing hormone receptor gene (*LHCGR*) that lead to precocious puberty in patients GN0068 and GN0088 and a deficiency in androgen biosynthesis defect due to a change in *HSD17B3* in case GN0038. Moreover, variants in *NR5A1* were determined in 4 cases with different phenotypes.

Table 12. Used approach to analyse each subject and previously studied genes.

Case	Previous studies	Approach	Case	Previous studies	Approach	Case	Previous studies	Approach
GN0001	*FSHR* and *WNT4*	NGS	GN0038[2]	__	CGA	GN0096	__	NGS
GN0003	*SRY*	NGS	GN0039	*AR, SRD5A2* and *SRY*	NGS	GN0100	*AR, NR5A1* and MLPA[4]	NGS
GN0004	-	CGA	GN0041[1]	-	CGA	GN0101	-	CGA
GN0007	-	CGA	GN0042[1]	*AR* and *HSD17B3*	CGA/NGS	GN0103	*FSHR* and *LHCGR*	NGS
GN0009	-	CGA	GN0043	*SRY*	NGS	GN0104	*SRY*	NGS
GN0011	*NR5A1* and *SRY*	NGS	GN0046	*AR*	NGS	GN0108	*AR, NR5A1 and SRD5A2*	NGS
GN0012[1]	-	NGS	GN0050	*NR5A1* and *SRY*	NGS	GN0109[2]	*AR and SRY*	CGA
GN0013	*NR5A1* and *SRY*	NGS	GN0051	*SRY* and MLPA[4]	NGS	GN0111[1]	*HSD17B3 and SRY*	CGA
GN0014	*SRY*	NGS	GN0054	-	CGA	GN0112	-	CGA
GN0017	-	NGS	GN0055	-	CGA	GN0114	*AR, NR5A1, SRY* and MLPA[4]	NGS
GN0018	*SRY*	CGA	GN0056	*SRD5A2 and WT1*	NGS	GN0118	*SRY*	NGS
GN0020	*SRY*	NGS	GN0059	*SRY*	NGS	GN0119[1]	-	NGS
GN0023	-	CGA	GN0064	*SRY*	NGS	GN0122	-	NGS
GN0024	*SRY*	CGA	GN0066	-	NGS	GN0123[1]	*AR*	CGA
GN0025	-	NGS	GN0068	-	CGA	GN0124	*AR, NR5A1 and SRD5A2*	NGS
GN0026	*SRY*	NGS	GN0070	-	NGS	GN0125	-	CGA
GN0027	*SRY*	NGS	GN0075	-	NGS	GN0132	-	CGA
GN0028	*AR, FMR1* and *SRD5A2*	NGS	GN0076	-	CGA	GN0133	-	CGA
GN0029	-	NGS	GN0078	-	CGA	GN0138	*NR5A1*	NGS
GN0031	*AR*	NGS	GN0080	-	CGA	GN0139	-	CGA
GN0033	*SRD5A2*	NGS	GN0084	*SRY* and *WNT4*	NGS	GN0141[1]	-	CGA
GN0034	*AR, HSD17B3, NR5A1* and *SRD5A2*	NGS	GN0088	-	CGA	GN0142	*AR, HSD17B3, NR5A1, SRD5A2* and MLPA[4]	NGS

Table 12. Used approach to analyse each subject and previously studied genes (Continuation).

Case	Previous studies	Approach	Case	Previous studies	Approach	Case	Previous studies	Approach
GN0035	-	CGA	GN0090	MLPA[4]	NGS	GN0144	-	NGS
GN0037	-	CGA	GN0091	-	CGA	GN0145	*AR*	NGS
GN0146	-	CGA	GN0169	-	NGS	GN0194	-	NGS
GN0147	*AR*	NGS	GN0171	-	NGS	GN0195	-	NGS
GN0148	*AR, NR5A1 and SRY*	NGS	GN0173	-	NGS	GN0196	-	NGS
GN0150	*AR*	NGS	GN0174	-	NGS	GN0198	-	NGS
GN0151	*SRY*	NGS	GN0175	-	NGS	GN0199	-	NGS
GN0152	*AR*	NGS	GN0176	-	NGS	GN0200	-	NGS
GN0153	-	CGA	GN0177	-	CGA	GN0201	-	NGS
GN0154	*SRY and MLPA4*	NGS	GN0178	-	NGS	GN0202	*AR*	NGS
GN0155	*AR, NR5A1 and MLPA4*	NGS	GN0179	-	NGS	GN0203	-	NGS
GN0156	*LHCGR*	NGS	GN0182	-	NGS	GN0204	-	NGS
GN0157	-	NGS	GN0183	-	NGS	GN0205	-	NGS
GN0158	*SRY and MLPA4*	NGS	GN0185	-	NGS	GN0207	-	NGS
GN0159	-	CGA	GN0186	-	NGS	OT0567	-	NGS
GN0160	*SRY*	NGS	GN0187	-	NGS	POL0274	-	CGA
GN0162	-	NGS	GN0189	-	NGS	POL0285	-	CGA
GN0163	-	NGS	GN0190	-	NGS	POL0301	-	CGA
GN0164	*LHCGR*	NGS	GN0191	-	NGS	RE0045	*AR, NR5A1, RET, SRY, WT1 and MLPA*[5]	NGS
GN0167	-	NGS	GN0192	-	NGS			

CGA, candidate gene approach; NGS, next generation sequencing. [1] Sample used as a positive control; [2]Relative of the sample used as positive control; [3]Sample used as positive control, but is not a DSD patient; [4] SALSA MLPA probemix P185-C2 Intersex; [5] SALSA MLPA P118 WT1 probemix; -, not don

Table 13. Genes analysed according to the suspected clinical diagnosis in the single-gene approach.

Suspected diagnosis	Altered gene	Number of cases
Androgen insensitivity syndrome	*AR*	13
46,XY DSD gonadal dysgenesis		
46,XY Ovotesticular DSD		
46,XX Testicular DSD or 46,XX DSD	*SRY*	7
46,XY DSD complete gonadal dysgenesis		
Primary adrenal insufficiency	*NR0B1*	7
Frasier syndrome	*WT1*	2
46,XY DSD LH Receptor defect	*LHCGR*	2
46,XY DSD Androgen biosynthesis defect	*HSD17B3*	1
46,XY DSD gonadal dysgenesis	*NR5A1*	4
Androgen insensitivity syndrome		

10.2. GENERAL CHARACTERISTICS OF THE GENETIC FINDINGS

In our cohort 81 rare variants were identified in DSD-associated genes either by the candidate gene approach or NGS (Table 14). Of these, 20 were repeated more than once, thus we determined 61 single variants in 20 clinically relevant DSD genes.

Variants in the *AR* gene were the most common. Among the 15 different changes (24.6%) found in 16 46,XY DSD individuals, all but one were classified as pathogenic, as they had been described in a DSD phenotype or were null variants.

Secondly, we determined 7 pathogenic variants in the *NR0B1* gene (11.3%). All but 2 had been previously associated to adrenal insufficiency, including missense, frameshift causing alterations and big deletions.

NR5A1 gene had the third highest number of variants called. Twenty-one patients had a *NR5A1* gene change, although only seven were unique (11.3%). Among these, 3 were missense and one was a frameshift that had not been described before.

We found five variants in *WT1* (8.2%), a gene described in several syndromes in which gonadal dysgenesis is associated. We found two known mutations in intron 9 in two cases with Frasier syndrome and a partial deletion (from exon 7 to 10) in a male infant with cryptorchidism and Wilms' tumour. Another two missense changes were discovered in *WT1* gene.

Table 14. Number of patients, variants and type of variants in each gene.

Gene	Number of patients	Number of different variants	Type of variant
AR	16	15	11 missense,2 frameshift, 1 deletion and 1 nonsense
NR0B1	7	7	2 deletions, 2 frameshifts, 2 nonsense and 1 missense
NR5A1	21	7	5 missense and 2 frameshift
WT1	5	5	2 missense, 2 intronic and 1 deletion
SRY	8	3	6 gene translocations, 1 nonsense and 1 missense
MAP3K1	1	1	1 missense
HSD17B3	1	1	1 missense
CYP17A1	1	1	1 missense
SRD5A2	3	3	2 missense and 1 deletion
STAR	2	2	2 missense
LHCGR	6	6	4 missense, 1 deletion and 1 nonsense
WWOX	2	2	2 missense
ESR1	1	1	1 missense
ESR2	1	1	1 missense
AMH	1	1	1 missense
DMRT2	1	1	1 missense
GATA4	1	1	1 missense
HSD17B4	1	1	1 frameshift
MAMLD1	1	1	1 missense
ZFPM2	1	1	1 missense

On the other side, *SRY* translocations were observed in 6 patients. Otherwise nonsense and an unreported missense variant were detected in two more cases harbouring a *SRY* defect. Of interest, a single mutation in the *MAP3K1* gene, causing gonadal dysgenesis, was detected by NGS.

Mutations in genes causing an androgen biosynthesis defect, such as *HSD17B3, CYP17A1* and *SRD5A2*, were observed in 4 patients. All the subjects harboured homozygous pathogenic variants, except one patient with two variants in compound heterozygosis in the *SRD5A2* gene. We also identified two unique not described variants in *STAR* gene.

A total of 6 changes were observed in the *LHCGR* gene (9.8%). These, were mainly disease-causing known mutation (4 out of 6), divided into 3 single-nucleotide

variants and a complete gene deletion. We also identified 2 LH receptor novel variants in equal number of patients.

We identified another 10 novel candidate variants in 10 patients. Two *WWOX* missense changes detected in our study were either classified as likely pathogenic and VUS, due to the results given by the prediction software. Interestingly, we identified 2 single variants in *ESR1* and *ESR2* genes, all unreported. As *ESR2* variants have been described in 46,XX DSD cases, we considered the missense change as probably deleterious. The remaining missense changes in *AMH, DMRT2, GATA4, HSD17B4, MAMLD1* and *ZFPM2* genes were classified as VUS and were only found in one patient each.

10.3. CLASSIFICATION OF GENETIC VARIANTS

Variants were evaluated and classified as previously reported in methodology.

As shown in Table 15, a total of 81 variants, including 61 unique changes, found in 72 individuals are reported. Repeated changes are *SRY* gene translocations in 46,XX DSD individuals, or alterations located in *AR* (c.2323C>T;p.Arg775Cys) and *NR5A1* (c.437G>C;p.Gly146Ala) genes. Results given by the prediction software are shown in Table 16 (Page 161).

We classified 54 variants out of 81 as pathogenic or likely pathogenic (66.7%) (See Figure 34, page 195). These, were mainly previously reported in association with a DSD phenotype (40 out of 54, 74.1%) and were mostly located in the *AR* and *NR0B1* genes, or were a Y-chromosome translocation confirmed by the detection of *SRY* material. The remaining 14 novel variants classified as pathogenic or likely pathogenic were identified in *AR, NR5A1*, *NR0B1 and LHCGR* most commonly, but also in *ESR2, SRY* and *WWOX* genes (1 variant called each).

Moreover, 12 variants of unknown significance were identified in 11 individuals (12 out of 81, 14.8%). Of these 11 individuals, one harboured two variants of unknown significance and in 3 patients a deleterious variant in an additional gene was found.

Regarding the p.Gly146Ala polymorphism in the *NR5A1* gene, it was classified as likely benign and was determined in 15 cases. In 4 individuals (GN0147, GN0156, GN0157 and GN0194), further sequence variations were detected in more than one gene.

Table 15. Classification of the genetic variants identified in this work.

Case	Gene	Variants	Classification	Location	Zygosity	Previously reported*	Diagnostic approach
GN0004	*SRY*	46,XX (SRY+)	Pa	-	-	Fechner PY (1993)	CGA
GN0007	*SRY*	c.289C>T; p.Gln97Ter	Pa	Exon 1	Hemi	Bilbao (1996)	CGA
GN0023	*SRY*	46,XX (SRY+)	Pa	-	-	Fechner PY (1993)	CGA
GN0054	*SRY*	46,XX (SRY+)	Pa	-	-	Fechner PY (1993)	CGA
GN0133	*SRY*	46,XX (SRY+)	Pa	-	-	Fechner PY (1993)	CGA
GN0141	*SRY*	c.391C>T; p.Pro131Ser	LP	Exon 1	Hemi	No	CGA
GN0159	*SRY*	46,XX.ish der(X)t(X;Y)(p22.3;p11.3)(SRY)	Pa	-	-	Yen (1991)	CGA
GN0187	*SRY*	46,XX.ish der(X)t(X;Y)(p22.3;p11.3)(SRY)	Pa	-	-	Yen (1991)	NGS
GN0012	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0028	*NR5A1*	c.88T>A; p.Cys30Ser	Pa	Exon 2	Het	No	NGS
	STAR	c.361C>T; p.Arg121Trp	VUS	Exon 4	Het	No	
GN0042	*NR5A1*	c.614_615insC; p.Gln206ThrfsTer20	Pa	Exon 4	Het	Camats (2012)	CGA/NGS
	AMH	c.428C>T; p.Thr143Ile	VUS	Exon 2	Het	No	
GN0051	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0070	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0075	*NR5A1*	c.250C>T; p.Arg84Cys	Pa	Exon 4	Het	Reuter (2007)	NGS
GN0090	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Hom	Wada (2006)	NGS
GN0096	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Hom	Wada (2006)	NGS
GN0109	*NR5A1*	c.910_913delGAGC; p.Glu304CysfsTer26	Pa	Exon 5	Het	No	CGA
GN0111	*NR5A1*	c.902G>A; p.Cys301Tyr	LP	Exon 5	Het	No	CGA
GN0118	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0119	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Hom	Wada (2006)	NGS
GN0123	*NR5A1*	c.71A>T; p.His24Leu	Pa	Exon 2	Het	No	CGA
GN0158	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Hom	Wada (2006)	NGS

Table 15. Classification of the genetic variants identified in this work (Continuation).

Case	Gene	Variants	Classification	Location	Zygosity	Previously reported*	Diagnostic approach
GN0163	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0182	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0199	*NR5A1*	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	NGS
GN0078	*NR0B1*	c.913C>T; p.Gln305Ter	Pa	Exon 1	Hemi	Rodriguez-Estévez (2015)	CGA
GN0091	*NR0B1*	c.291delC; p.Glu98ArgfsTer166	Pa	Exon 1	Hemi	Guoying (2012)	CGA
GN0101	*NR0B1*	g.(?_30327014)_(30361290_?)del	Pa	-	Hemi	Guo (1995)	CGA
GN0153	*NR0B1*	c.528C>G; p.Tyr176Ter	Pa	Exon 1	Hemi	No	CGA
POL0274	*NR0B1*	c.871T>A; p.Trp291Arg	Pa	Exon 1	Hemi	Yeste (2009)	CGA
POL0285	*NR0B1*	g.(?_29978097)_(30361290_?)del	Pa	-	Hemi	Guo (1995)	CGA
POL0301	*NR0B1*	c.712_713delAC; p.Thr238LeufsTer60	Pa	Exon 1	Hemi	No	CGA
GN0009	*WT1*	c.1447+5G>A	Pa	Intron 9	Het	Bruening (1992)	CGA
GN0132	*WT1*	c.1447+4C>T	Pa	Intron 9	Het	Barbaux (1997)	CGA
GN0150	*WT1*	c.223G>A;p.Glu75Lys	VUS	Exon 1	Het	No	NGS
GN0156	*WT1*	c.545T>A; p.Met182Lys	VUS	Exon 2	Het	No	NGS
	NR5A1	c.437G>C; p.Gly146Ala	LB	Exon 4	Hom	Wada (2006)	
OT0567	*WT1*	c.(1099-?_1551+?)del; p.(Asp367?_Leu517?)del	Pa	-	Het	Finken (2015)	NGS
GN0011	*MAP3K1*	c.2291T>G; p.Leu764Arg	LP	Exon 13	Het	Granados (2017)	NGS
GN0020	*WWOX*	c.1096C>G; p.Pro366Ala	VUS	Exon 9	Het	No	NGS
GN0203	*WWOX*	c.184G>A; p.Gly62Arg	LP	Exon 3	Het	No	NGS
GN0198	*ESR1*	c.1781C>T;p.Thr594Met	VUS	Exon 9	Het	No	NGS
	HSD17B4	c.524delC; p.Ala175GlufsTer26	VUS	Exon 8	Het	No	
GN0207	*ESR2*	c.661A>G; p.Arg221Gly	LP	Exon 5	Het	No	NGS
GN0142	*DMRT2*	c.1607C>T; p.Ser536Leu	VUS	Exon 4	Het	No	NGS
GN0155	*ZFPM2*	c.3077C>T, p.Ala1026Val	VUS	Exon 8	Het	No	NGS
GN0154	*MAMLD1*	c.2009C>T; p.Thr670Ile	VUS	Exon 5	Het	No	NGS

Table 15. Classification of the genetic variants identified in this work (Continuation).

Case	Gene	Variants	Classification	Location	Zygosity	Previously reported*	Diagnostic approach
GN0018	*AR*	c.2323C>T; p.Arg775Cys	Pa	Exon 6	Hemi	Brown (1990)	CGA
GN0024	*AR*	c.2086G>A; p.Asp696Asn	Pa	Exon 4	Hemi	Ris-Stalpers (1991)	CGA
GN0035	*AR*	c.2522G>A; Arg841His	Pa	Exon 7	Hemi	McPhaul (1992)	CGA
GN0037	*AR*	c.2178C>G; p.Phe726Leu	Pa	Exon 5	Hemi	Quigley (1995)	CGA
GN0041	*AR*	c.(1616+1_1617-1)_(1768+1_1767-1)del; p.(Arg539_Asp305del)	Pa	Exon 2	Hemi	Quigley (1992)	CGA
GN0055	*AR*	c.2710G>A; p.Val904Met	Pa	Exon 8	Hemi	McPhaul (1992)	CGA
GN0076	*AR*	c.2566C>T; p.Arg856Cys	Pa	Exon 7	Hemi	McPhaul (1992)	CGA
GN0080	*AR*	c.298insC; p.His100ProfsTer3	Pa	Exon 1	Hemi	No	CGA
GN0112	*AR*	c.827delC;p.Pro276HisfsTer20	Pa	Exon 1	Hemi	No	CGA
GN0125	*AR*	c.865G>T; p.Glu289Ter	Pa	Exon 1	Hemi	Holterhus (2003)	CGA
GN0139	*AR*	c.2642T>G; p.Leu881Arg	LP	Exon 8	Hemi	No	CGA
GN0146	*AR*	c.2473C>A; p.Gln825Lys	Pa	Exon 7	Hemi	Hellmann (2012)	CGA
GN0164	*AR*	c.2270A>G; p.Asn757Ser	Pa	Exon 5	Hemi	Hiort (1996)	NGS
GN0177	*AR*	c.1301C>T; p.Ser434Phe	Pa	Exon 1	Hemi	Holterhus (2005)	CGA
GN0189	*AR*	c.2567G>A; p.Arg856His	Pa	Exon 7	Hemi	Batch (1992)	NGS
GN0194	*AR*	c.2323C>T; p.Arg775Cys	Pa	Exon 6	Hemi	Brown (1990)	NGS
	NR5A1	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	
GN0046	*SRD5A2*	c.377A>G; p.Gln126Arg	Pa	Exon 2	Het	Thigpen (1992)	NGS
	SRD5A2	c. (-1+1_1-1)_(281+1_280-1)del; p.(Met1_Arg94del).	Pa	Exon 1	Het	Fenichel (2013)	
GN0186	*SRD5A2*	c.271T>G; Tyr91Asp	Pa	Exon 1	Hom	Wilson (1993)	NGS
GN0034	*LHCGR*	arr [hg19] 2p16.3(48,905,663-48,983,208)x0	Pa	-	Hom	Richard (2011)	NGS
GN0068	*LHCGR*	c.1713G>T; p.Met571Ile	Pa	Exon 11	Het	Kremer (1993)	CGA
GN0088	*LHCGR*	c.1193T>C; p.Met398Thr	Pa	Exon 11	Het	Kraaij (1995)	CGA
GN0147	*LHCGR*	c.757T>C; p.Ser253Pro	Pa	Exon 9	Hom	No	NGS
	NR5A1	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	

Table 15. Classification of the genetic variants identified in this work (Continuation).

Case	Gene	Variants	Classification	Location	Zygosity	Previously reported*	Diagnostic approach
GN0157	*LHCGR*	c.568C>A; p.Gln190Lys	LP	Exon 7	Het	No	NGS
	NR5A1	c.437G>C; p.Gly146Ala	LB	Exon 4	Het	Wada (2006)	
GN0171	*LHCGR*	c.1660C>T; p.Arg554Ter	Pa	Exon 11	Het	Latronico (1996)	NGS
	GATA4	c.677C>T; p.Pro226Leu	VUS	Exon 3	Het	Martinez de LaPiscina (2018)	
RE0045	*CYP17A1*	c.1246C>T; p.Arg416Cys	Pa	Exon 8	Hom	Takeda (2001)	NGS
GN0038	*HSD17B3*	c.845C>T; p.Pro282Leu	Pa	Exon 11	Hom	Andersson (1996)	CGA
GN0185	*STAR*	c.50T>G; p.Met17Arg	VUS	Exon 1	Het	No	NGS

CGA, candidate gene approach; Hemi, hemizygosis; Het, heterozygosis; Homo, homozygosis; LP, likely pathogenic; NGS, next generation sequencing; Pa, pathogenic; VUS, variant of unknown significance.*Previously reported in a DSD phenotype; -, none.

Table 16. Results obtained by the different prediction software.

Case	Gene and variants	Provean	SIFT	Polyphen HumDiv/HumVar	Mutation Taster	SNPs&go	MutPred	Panther	VarSome
GN0020	*WWOX*; p.Pro366Ala	N(-2.36)	T(0.328)	B (0.009)/B(0.004)	DC	N(8;0.102)	<0.50	PsD(324)	VUS
GN0028	*NR5A1;* p.Cys30Ser	Del(-8.22)	Dam(0.000)	PrD(0.988)/PrD(0.988)	DC	D(9;0.965)	[1](0.03),[2]at S32 (0.04)	PrD(911)	LP
	STAR; p.Arg121Trp	Del(-3.98)	Dam(0.009)	PrD (1.0)/PrD(0.995)	DC	N(4;0.309)	[3](0.04)	PsD(324)	VUS
GN0042	*AMH*; p.Thr143Ile	Del(-2.88)	Dam(0.035)	B (0.411)/B(0.187)	P	N(5;0.231)	<0.50	PrB(103)	VUS
GN0080	*AR*; p.His100ProfsTer3	N/A	N/A	N/A	DC	N/A	N/A	N/A	VUS
GN0111	*NR5A1*; p.Cys301Tyr	Del(-6.67)	Dam(0.0)	PrD(1.0)/PrD(1.0)	DC	Dam(6;0.795)	[4](3.7e^{-03}),[5](0.02)	PrD(797)	LP
GN0123	*NR5A1*; p.His24Leu	Del(-10.13)	Dam(0.0)	PrD(0.997)/PrD(0.993)	DC	D(6;0.787)	[6](0.02),[7]at Y25(0.03),[4](0.02),[8](0.04),[9](0.03)	N/A	LP
GN0139	*AR;* p.Leu881Arg	Del(-4.48)	Dam(0.001)	PrD 1.0)/PrD(0.897)	DC	N(1;0.457)	[10](0.27),[9](0.11)	PrD(455)	LP
GN0141	*SRY;* p.Pro131Ser	Del(-7.12)	Dam(0.001)	PrD(1.0)/PrD(1.0)	DC	D(8;0.902)	<0.50	PrD(912)	LP
GN0142	*DMRT2;* p.Ser536Leu	N(-1.69)	Dam(0.002)	PsD(0.614)/B(0.038)	DC	N(8;0.124)	<0.50	PsD(324)	VUS
GN0147	*LHCGR;* p.Ser253Pro	N(-2.10)	Dam(0.004)	PrD(0.999)/PrD(0.986)	DC	D (9; 0.969)	[9](8.9e^{-04})	PsD(361)	VUS
GN0150	*WT1*; p.Glu75Lys	N(-1.58)	Dam(0.0)	PrD(0.980)/PsD(0.856)	DC	N(1 0.427)	<0.50	Dam(750)	VUS
GN0154	*MAMLD1;* p.Thr670Ile	N(-1.54)	Dam(0.042)	N/A	P	N/A	<0.50	N/A	LB
GN0155	*ZFPM2*; p.Ala1026Val	N(0.03)	T(0.603)	B(0.0)/B(0.0)	DC	N(10;0.253)	[11](0.00),[5](0.01),[12](0.03)	N/A	VUS
GN0156	*WT1*; p.Met182Lys	Del(-2.78)	Dam(0.0)	B(0.140)/B(0.143)	DC	D(3;0.632)	<0.50	PrD(455)	VUS
GN0157	*LHCGR;* p.Gln190Lys	N(-1.22)	T(0.162)	B(0.282)/B(0.170)	DC	D(8;0.898)	<0.50	PsD(220)	VUS
GN0171	*GATA4;* p.Pro226Leu	Del(-9.67)	Dam(0.0)	PrD(1.0)/PrD(1.0)	DC	D(10;0.992)	[9](8.2e^{-03})	PrD(1368)	VUS
GN0185	*STAR;* p.Met17Arg	N(-1.74)	Dam(0.003)	B(0.091)/B(0.037)	DC	D(6;0.778)	[4](3.3e^{-03}),[13](8.3e^{-04}) [8](0.02),[7]at Y14 (0.02),[14]at K21 (0.02)	N/A	VUS

Table 16. Results obtained by the different prediction software (Continuation)

Case	Gene and variants	Provean	SIFT	Polyphen HumDiv/HumVar	Mutation Taster	SNPs&go	MutPred	Panther	VarSome
GN0198	*HSD17B4; p.Ala175GlufsTer26*	N/A	N/A	N/A	DC	N/A	N/A		Pa
	ESR1; p.Thr594Met	N(-1.23)	Dam(0.001)	PrD(1.0)/PrD(0.976)	P	N(7;0.168)	<0.50	N/A	VUS
GN0203	*WWOX;* p.Gly62Arg	Del(-5.67)	Dam(0.001)	PrD(1.0)/PrD(0.995)	DC	D(1;0.544)	[6](0.04),[15]at Q65 (0.02),[16]at Y61 (0.02),[4](0.04),[9](0.01).	PsD(1037)	VUS
GN0207	*ESR2;* p.Arg221Gly	Del(-5.73)	Dam(0.001)	B(0.085)/B(0.086)	DC	D(3;0.627)	<0.50	PrD(750)	VUS
POL0301	*NR0B1*; p.Thr238LeufsTer60	N/A	N/A	N/A	DC	N/A	N/A	N/A	Pa

Score given by each prediction program is given in brackets. Provean, the variant is predicted to be deleterious when the score is equal or below the default score threshold (-2.5). SIFT, ranges from 0 to1 and the substitution is predicted to be tolerated if >0.05. Polyphen, divided into HumDiv and HumVar, the score values nearer one are more confidently predicted to be deleterious. SNPs and GO, if disease probability is >0.5 the variant is predicted as disease. Mutpred, a g score of 0.50 suggest pathogenicity and then, a loss/gain of structural or functional property. Only predictions with a g score >0.50 are shown. (1) Gain of disorder; (2) Loss of catalytic residue; (3) Altered stability; (4) Altered disordered interface; (5) Loss of helix; (6) Loss of strand; (7) Loss of allosteric site; (8) Altered metal binding; (9) Altered transmembrane protein; (10) Gain of strand; (11) Gain of sheet; (12) Gain of loop; (13) Altered DNA binding; (14) Loss of acetylation; (15) Loss of Pyrrolidone carboxylic acid; (16) Loss of sulfation. Panther, A longer preservation time indicates a greater functional impact. B, benign; Class, classification; D, disease; Dam, damaging; DC, disease causing; Del; deleterious; LB, likely benign; LP, likely pathogenic; N, neutral; N/A, not available; ND, not determined; P, polymorphism; Pa, pathogenic; PrB, probably benign; PrD, probably damaging; PsD, possibly damaging; Pt, preservation time; RI, reliability index; T, tolerated; VUS, variant of unknown significance. Previously reported gene variations classified as pathogenic or likely benign were not analysed *in silico*.

10.4. CLINICAL AND MOLECULAR DESCRIPTION OF THE GENETICALLY POSITIVE PATIENTS

I have grouped the patients presenting a genetic finding according to the main action of the gene during development. Firstly, I present those cases having a genetic variant in a gene related to gonadal determination and development, such as *SRY*, *NR5A1,* etc. Then, I describe patients with a disorder of sex development due to a genetic change in a gene associated to genital or sex differentiation (See 10.4.2, page 182).

10.4.1. Findings in genes related to gonadal development

10.4.1.1. *SRY* variants

A total of 8 patients had an alteration in the *SRY* gene. We detected the presence of the *SRY* gene in six 46,XX testicular DSD individuals, five were detected by either PCR or NGS and one by aCGH. PCR-based sequencing revealed another two single-nucleotide variants (SNV) in two patients with 46,XY DSD.

10.4.1.1.1. Patients with a *SRY* gene translocation studied by PCR or NGS

The majority of these patients presented with unilateral cryptorchidism or small testes at diagnosis, gynecomastia and azoospermia (GN0004, GN0133 and GN0187). All the patients had otherwise a normal male phenotype (Supplementary data 3). Individual GN0023 was referred at age 2 with ambiguous genitalia due to the presence of gonads in inguinal canal that descended after hCG stimulation. On the contrary, patient GN0054 presented at 41 years old with elevated FSH levels (FSH: 31.7U/L; LH: 5.37U/L) and absence of body hair. Echography was normal. Karyotype of the patients was 46,XX.

The presence of the SRY gene was studied by PCR in all these samples, except in GN0187, which was analysed by NGS. To validate the genetic findings a FISH analysis was also performed to determine the situation of the *SRY* gene. As usual, it was translocated to the distal tip of the short arm of the X chromosome [46,XX.ish der(X)t(X;Y)(p22.3;p11.3)(SRY)].

10.4.1.1.2. Patients with a *SRY* gene translocation studied by aCGH

GN0159 male patient was born with hypospadias and unilateral cryptorchidism. At amniocentesis karyotype was 46,XX and at birth, FISH testing revealed the

translocation of the *SRY* gene (46,XX.ish der (X)t(x,y)(p22.3;p11.3)(SRY+)). 3 months later, penis was normal and 2mL testes were present in scrotum. The boy had developmental delay, microcephaly and partial agenesis of the corpus callosum. Then, Smith Lemli Opitz syndrome was suspected. At 9 years of age, he presented at Tanner stage I, augmented penis, right testis in scrotum and left in inguinal canal. Hormonal analysis indicated low testosterone with normal gonadotropins values. He was referred to dermatology service due to café au lait spots when he was 11 years old. Surgery for hypospadias was also performed. One year later, testicular scanner showed a right testis of 22.7x7.3mm in size and left testis located in inguinal canal (24x7.5mm). At 13 years of age, left orchidopexy was performed and no pubertal signs were observed.

We made an aCGH to discard a possible deletion in the short arm of chromosome 6 (6p) related to the phenotype of the patient. However, we observed an 8.12Mb deletion at Xp22.2-22.33, from position 2,709,027 to 10,830,236 (Figure 22). This deletion involved 19 genes, including *MID1* (OMIM 300552) which has been related to Opitz G/BBB syndrome (OMIM 300000), a congenital malformation syndrome characterized by hypertelorism, hypospadias, cleft lip, laryngotracheoesophageal abnormalities, imperforate anus, developmental delay, and cardiac defects (237). Moreover, the imbalance has been described in some of the most common CNV databases, such as ISCA (International Standards for Cytogenomic Arrays) consortium and DECIPHER (Database of genomic variation and phenotype in humans using ensembl resources), as structural variation with pathogenic clinical significance. Some of the deleted genes encode proteins that could be responsible for the phenotype, even the hypospadias. On the other hand, we showed a 430.7Kb Yp11.2-p11.3 region translocation, from 2,654,967 to 3,085,681, comprising *SRY, RPS4Y1, ZFY* and *TGIF2LY* genes (Figure 22). The presence of the *SRY* sequence might explain the DSD in the GN0159 patient. This *SRY* translocation was confirmed by FISH technique (Figure 23, page 166).

Both alterations suggested a Xp22; Yq11 chromosome translocations. These chromosomal abnormalities may be due to a balanced translocation in one of the parents' genomes, therefore we performed a FISH analysis to clarify the origin of the CNV. Both parents presented normal hybridization patterns (Figure 23, page 166).

10.4.1.1.3. Patients with single-nucleotide variants in *SRY*

A missense and a nonsense variant were also identified in 3 individuals of two different families by sequencing the *SRY* gene. Index patient GN0007 was a 16-year-old female with a 46,XY karyotype who had not reached puberty. A small uterus was present, and two streak gonads were removed by laparotomy. Histology identified

bilateral gonadoblastomas. The sister of this index case was an apparently normal female with a 46,XY karyotype. Echocardiography revealed a small uterus. Supplementary data 3 shows the levels of gonadotropins and gonadal steroids in the index case. We found a hemizygous change of cytosine to thymine at location 291 of the *SRY* gene in both sisters. This mutation changes codon 97 from glutamine to a stop codon (c.289C>T; p.Gln97Ter) and results in a truncated protein. The father presented the variants as a mosaicism, while other male relatives had the wt genotype. This mutation has been previously reported by our group (238).

Figure 22. aCGH detection of the Xp22; Yq11 translocation in the GN0159 patient. A, Ideogram of chromosome X. The area outlined by the dotted box is enlarged on the right. Log2 ratio plot of the patient relative to the control. Red dots represent probes with a log2 ratio below -0.5 and the blue dots greater than 0.5. The red area is the loss region, the 8.12Mb deletion at Xp22.2-22.33. B, Ideogram of chromosome Y. The blue area revealed a gain (430.7kb) at Yp11.2-11.31.

Patient GN0141 had her gonads removed at puberty. Surgery showed streak gonads with tumoral aspect and gonadoblastoma, as well as remaining of Müllerian ducts. Treatment with oestrogens was prescribed and 10 years later, hormonal test revealed elevated gonadotrophins with high oestrogens levels. We found a C to T change in hemizygosis at position 391 (c.391C>T; p.Pro131Ser). The variant was predicted as potentially pathogenic since the wt amino acid and surrounding region is highly conserved (Supplementary data 14). Furthermore, another mutation has been described in the same codon associated to gonadal dysgenesis (239).

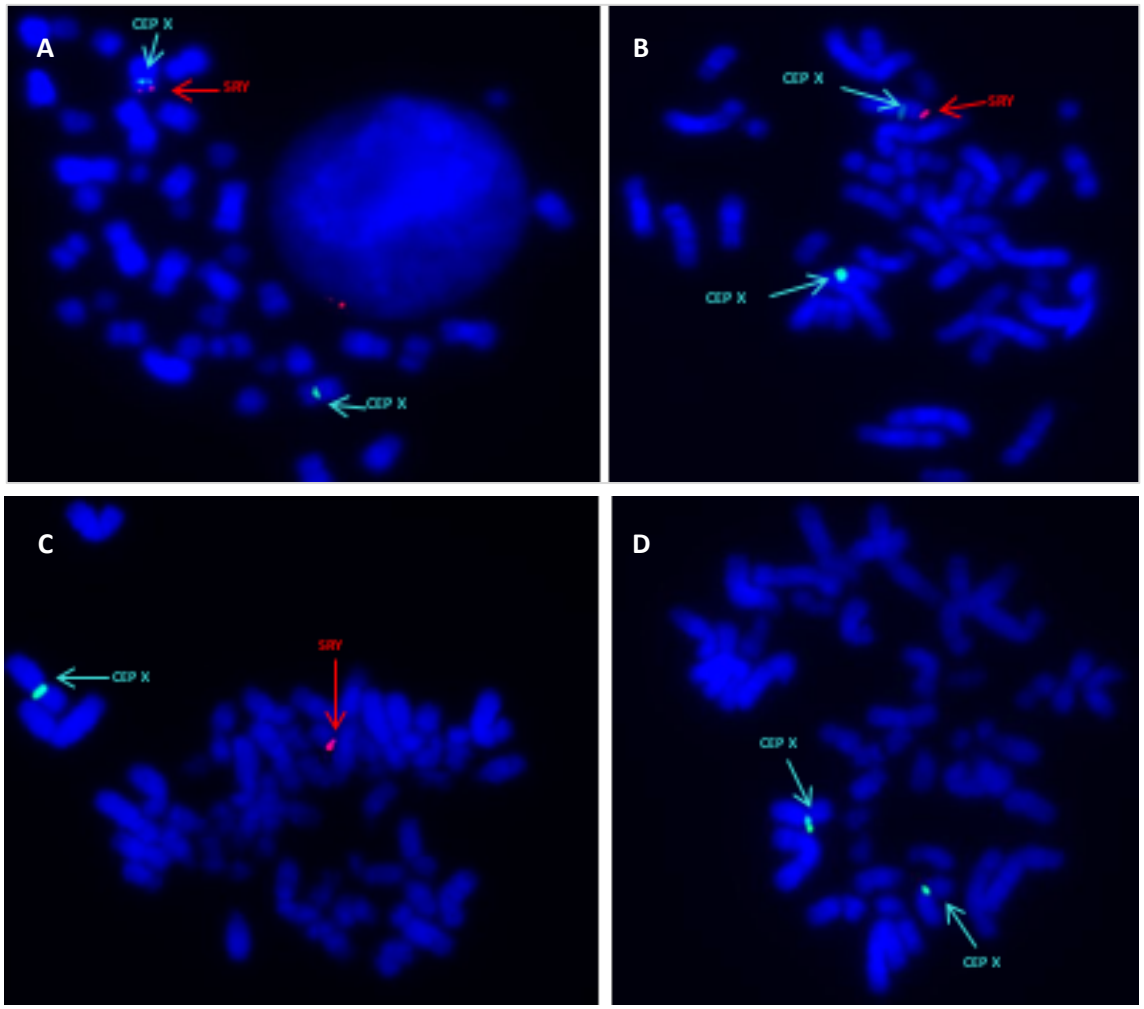

Figure 23. FISH confirms the Xp22:Yq11 translocation. A and B, representative metaphase spreads of the patient. The SRY specific probe (labelled in red) hybridizes to the Xp region of one of the X chromosomes (labelled in green, CEP X). Metaphase spreads of normal hybridization patterns of the patient's father (C) and mother (D).

10.4.1.2. *NR5A1* variants

We found 6 variants in the *NR5A1* gene in six cases with a 46,XY karyotype. The subjects presented mainly, with ambiguous genitalia but also primary amenorrhea, bilateral inguinal masses and virilisation at puberty. Four of the six gene variants were point mutations that led to a missense change and 2 were frameshift (1 insertion and 1 deletion).All were located throughout exons 2 to 5 and four of them were novel. All the identified variants were in heterozygosis. Three variants were found in other family members (cases GN0075, GN0109 and GN0111), whereas the c.88T>A in individual GN0028 appeared to be *de novo*. Complete family studies could not be done in the other patients (GN0042 and GN0123).

10.4.1.2.1. Patients with missense variants in the *NR5A1* gene

Case GN0028 was a newborn noted to have curved micropenis with scrotal hypospadias and bilateral cryptorchidism, corrected by surgery later. At 10 years of age, stimulation test with hCG resulted in a normal rise of testosterone (187.4ng/dL). At 12 years, his penis was 7.6cm in size, testes of 5mL were observed and presented a pubarche and axilarche stage V. Magnetic Resonance Imaging (MRI) showed testes of 20x15x15mm and 14x15x35mm corresponding to right and left testis respectively. A small cyst was also observed in the right epididymis. At 14 years, biochemical analysis revealed elevated serum levels of gonadotropins (LH 11.4U/L, FSH 35.9U/L) and testosterone (412.5ng/dL) with low levels of AMH (<0.1ng/mL). At recent follow-up at 15 years of age, penis was in the normal limit. He still presents high levels of LH and FSH with a testosterone of 518ng/dL. MRI demonstrated small cysts in both epididymis of 6.4mm and 2.7mm. Asperger syndrome was suspected at the age of 13 years old. Family history is remarkable for the following: patient's mother has unilateral renal agenesis, as well as brother, with also vas deferens and left epididymis agenesis. We found the novel C to A change in the exon 2 (c.88T>A), which is predicted to result in the p.Cys30Ser variants. Both parents were studied and did not carry the variant, thus this *NR5A1* variant is a *de novo* change in the patient. Brother was also negative for the variant.

Index case of GN0075 family presented micropenis, scrotal hypospadias, undescended testis and bifid scrotum at birth (Supplementary data 2). Abdominal ultrasound (US) revealed no anomalies. Biochemically, LH and FSH were elevated (2.5 and 3.7U/L, respectively) with a testosterone of 6750ng/dl during the first month. At 3 months, four doses of testosterone were prescribed (50mg/dose) with a good response of the penis. At 5 years of age, physical examination showed a penis of 2.5cm in size, right testis of 2ml and testicular left hydrocele. At age 6, replacement testosterone treatment was started. Penis was normal (3cm) and pubic hair increase was noted. Interestingly, mother presented with premature menopause, maternal uncle with scrotal hypospadias at birth and aunt with menstrual disorders. A heterozygous C to T transition located at c.250 of the coding *NR5A1* sequence predicted to replace an arginine with cysteine at position p.84 (c.250C>T; p.Arg84Cys) was found. The same heterozygous variant was also found in the mother, brother, maternal uncle, aunt and grandfather (Figure 24). The variant, located in exon 4, was described firstly by Reuter *et al* in a Japanese individual with dysgenetic testes and normal adrenal function (240). Since then, it has been reported in several cases with different phenotypes (241).

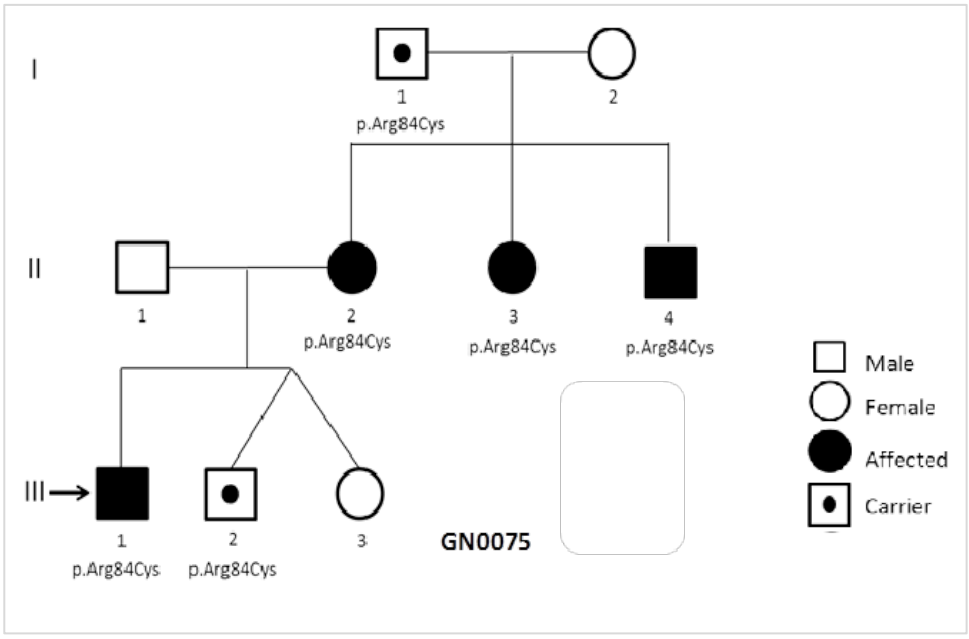

Figure 24. Pedigree and genotype information of the GN0075 family. Index case is indicated with an arrow.

Patient GN0123 was noted at birth to have curved penis buried in pubis, perineal hypospadias, bifid scrotum and palpable gonads. At 2 years of age, biopsy showed testicular tissue without variants and laparoscopy didn't show Müllerian rests. During the following years, hypospadias was corrected and at 10 years of age penis was 2.6cm, and both testicles were correctly located. Stimulation test with hCG resulted in a normal rise of testosterone. He started treatment for micropenis with testosterone (250 mg), which increased phallic size to 3.5 cm and began pubarche. At 14 years of age, the boy had a fallus of 5.5 cm and testes of 2-3cc. Biochemically, gonadotropins were elevated (LH: 15U/L; FSH 55U/L) at a testosterone of 1.8ng/dL after HCG stimulation test. ACTH, cortisol and 17-hydroxypreogesterone were in the normal range. Intramuscular testosterone treatment was prescribed again for 6 months. At the age of 15 years, penis was 7cm and testis 2cc. After two months of the suspension of the treatment LHRH stimulation test was done and gonadotropins and testosterone resulted in a normal rise (Supplementary data 3). The subsequent years, a second biopsy revealed the absence of germinal cells and Leydig cells hypoplasia, orchidopexy was performed and testosterone treatment was restored. Currently, he presents with a penis of 3-4 cm and testis of 2cc. LH and FSH were elevated in the presence of a low testosterone (500ng/dL). Apart from the development of his external genitalia, he is obese. Direct analysis by Sanger sequencing revealed the novel c.71A>T; p.His24Leu change in exon 2. Samples from the parents were not available.

Patient GN0111 presented with primary amenorrhea and obesity at 14 years of age. She had a stenotic and enlarged vagina, which was possibly a vaginal pouch. Ultrasound didn't identify female internal genitalia. MRI showed a rudimentary uterus

and vagina, but ovaries were not ascertained, although streak gonads couldn't be discarded. Biochemical analysis at this time showed increased levels of gonadotropins with a slight high testosterone (36.60ng/dL). Stimulation test with hCG resulted in a minor rise of testosterone (47.70 and 46.0ng/dL for the 3rd and 5th dose of hCG, respectively). Laparoscopy at the age of 15 years revealed a rudimentary uterus and streak gonads. Gonadectomy and biopsy showed compatible gonads with non-functioning testicular parenchyma and normal Fallopian tube. No interesting family history was reported. Analysis with the customized panel showed that the patient harboured the novel c.902G>A; p.Cys301Tyr change in heterozygosis. Curiously, the healthy mother of the patient carried the change in *NR5A1* as a mosaicism, while the father and brother were negative.

In addition to *in silico* tests and comparison of the residues against diverse species (Supplementary data 14), *in vitro* functional studies were performed to study the transactivation activity of the mutants (See 10.5, page 200). Then, we classified the p.Cys30Ser and p.His24Leu as pathogenic, while p.Cys301Tyr was categorized as likely pathogenic.

10.4.1.2.2. Patients with frameshift variants

Case GN0042 was a female referred at age 11 because of clitoromegaly, bilateral inguinal masses and growth increase. She presented with beginning of pubarche and facial hair growth. Pelvic ultrasound (US) showed a vaginal pouch and absence of uterus, as well as palpable gonads in inguinal canal (2 and 3ml). Biochemical analysis was remarkable with elevated gonadotropins at a testosterone of 250ng/dL. Gonadectomy and subsequent histology revealed 2 testes of 2.4 and 2.5cm with germinal hypoplasia. Oestrogen oral therapy was initiated for the next 9 years and then was replaced with patches. This resulted in good breast development and sexual hair although mammary regression was observed when the dose was decreased. During the following years reconstructive surgery was performed. Abdominal US showed a normal bladder and a small structure (26mm) above the urethra, indicating a possible prepuberal hypoplasic uterus. No ovaries were perceived. At 27 years of age the patient was well developed. Hormonal values were in the normal, age-and sex-appropriate range. The c.614_615insC variant in exon 4 leads to an amino acid change and to a stop codon 20 residues afterwards (p.Gln206ThrfsTer20). This had been previously reported in a 46,XY female with ambiguous genitalia (14). Only the mother of the proband was studied and brought no variants.

The second frameshift variant was identified in GN0109. This female was seen at the age of 14 because of virilization during puberty. Patient referred no ambiguous

genitalia at birth and pubarche and axilarche at the age of 8 years. At examination, female external phenotype, erectile hypertrophic organ with perineal urethra and absence of inguinal masses was denoted. Pelvic Magnetic resonance imaging showed a rudimentary uterus, and computed tomography (CT) scan revealed uterus and atrophic gonads in left inguinal canal. Analysis resulted in elevated LH and FSH levels (37.8 and 113U/L, respectively) with high levels of testosterone (193ng/dL). At 15 years, orchidopexy was performed and biopsy showed muscle tissue with a Sertoli-cell phenotypic testicle and epididymis. After surgery patient began treatment with oestrogens. The novel c.910_913delGAGC; p.Glu304CysfsTer26 variant was found in *NR5A1* and was classified as pathogenic due to its null effect. The mother of the proband presented the change in heterozygosis and curiously, her sister had hirsutism.

10.4.1.2.3. Identification of the p.Gly146Ala polymorphism in our cohort

Since the first description of the rs1110061 SNP by WuQiang *et al* in 2003 in which no dominant negative effect was proven (242), the c.437G>C; p.Gly146Ala variant in *NR5A1* gene has been reported in several DSD-related phenotypes, such as hypogonadotropic hypogonadism, micropenis, cryptorchidism, POI and even in diabetes, highlighting its poor role in the development of sex disorders (14) (243). Moreover, the allele frequency differs among the populations. Regarding the 1000 Genomes Project Phase 3 in combined population, the frequency of the minor allele G is 0.33, while for ExAC, ESP and gnomAD databases this frequency is lower (0.11, 0.24 and 0.23, respectively). Studies in East Asia indicated a MAF of the G allele similar to the general population (0.323), however, this frequency decreases in Spanish (MAF G allele: 0.05). On the contrary, studies in Africa have revealed that the most common allele is the guanine (0.831), specifically the frequency of the G in people from Sierra Leone is 0.894. The results given in this work emphasize the categorization of this change in the *NR5A1* gene as a benign polymorphism, due to the high frequency, wide range of phenotypes and origin of the patients in which it has been identified.

We have identified the c.437G>C; p.Gly146Ala polymorphism, located in the exon 4 of the *NR5A1* gene in 15 individuals out of 125 (12%). Out of these patients, 9 were examined due to ambiguous genitalia and were recorded as phenotypic males with micropenis and scrotal hypospadias in some cases. In GN0051, GN0070 and GN0163 probands, ultrasound detected testes located either in inguinal canal or in the scrotum. On the contrary, a laparoscopy was performed to patient GN0090 and bilateral ovotestes were reported. Karyotype was 46,XY in all the probands, except in GN0090 and GN0051. Patients GN0096 and GN0158 were both referred at age 3 and a 46,XX DSD was suspected. Female normal karyotype was set and histology revealed testicular and ovarian tissue. The two African infants (GN0118 and GN0119) had

curved penis with scrotal hypospadias and cryptorchidism at examination. GN0182 had unilateral cryptorchidism, as well as bifid scrotum at birth.

On the other hand, patient GN0012 presented as a newborn with micropenis (10mm) and bilateral anorchia. Stimulation test with hCG resulted in no response of testosterone, indicating an absence of testicular tissue, which was confirmed later. Testicular prostheses were placed at age 7 and increasing doses of testosterone were prescribed. Biochemically, he had markedly elevated gonadotropins (LH: 42.3U/L; FSH: 85U/L) in the presence of low testosterone (55ng/dL). At 15 years of age, he presented a mature penis and scrotum, pubarche IV and facial and body hair. During the following years, he was well virilised and hormones were in the normal, age-and sex-appropriate range. Finally, case GN0199 was referred because of micropenis and fused labia minora. At examination, non-palpable gonads and a tight vagina were noticed. Ultrasound (US) revealed a normal uterus and the absence of gonads. Karyotype was 46,XY. This male was diagnosed with a complete gonadal dysgenesis.

Patients GN0147, GN0156, GN0157 and GN0194 are clinically characterized elsewhere because other causative gene variants have been found along with this *NR5A1* gene change. All of them were identified by DSD targeted gene panel sequencing, except cases GN0012 and GN0119 in which the variant was found in heterozygosis and homozygosis by a candidate gene approach. These two samples were included in the panel to search for any other additional disease-causing gene variant, however none was found.

10.4.1.3. *NR0B1* variants

A total of 7 male patients (GN0078, GN0153, GN0091, POL0301, POL0274, POL0285, and GN0101) were evaluated with suspicion of non-congenital adrenal hyperplasia (non-CAH) primary adrenal insufficiency (PAI). The molecular analysis of the *NR0B1* gene was performed by PCR amplification of the 2 coding exons and subsequent sequencing. Among them, 2 new gene sequence changes were found in exon 1 (c.528C>G and c.712_713delAC). Patients GN0101 and POL0285 showed no amplification of the gene and therefore a MLPA was performed.

10.4.1.3.1. Sanger sequencing of the gene

Case GN0078 was diagnosed at 45 days of life due to his brother's neonatal death and the familial background of unexplained deaths of eight male infants in the first months of life. Despite the absence of glucocorticoid insufficiency (cortisol: 17.8µg/dL) or hyperpigmentation the clinical diagnosis led to PAI. A scan of the adrenal glands showed they were hypoplastic (0.3mL) and patient began with hydrocortisone

and fludrocortisone treatment. At 3 years, he presented with high testosterone levels (170ng/dL), and later hypogonadotropic hypogonadism. GN0153 was admitted at hospital at 11 days of life due to a hyponatremic dehydration and hyperkalaemia with compatible PAI. Treatment was given until cortisol, 17-hydroxyprogesterone and androgens values were normalized, when hydrocortisone administration was stopped. 8 months later he was controlled at hospital due to hypoaldosteronism. No interesting family background was reported. We identified 2 different point mutations (c.913C>T; p.Glu305Ter and c.528C>G; p.Tyr176Ter) in GN0078 and GN0153 cases, which predict a truncated protein product and thus the generation of a non-functional protein. Moreover, the c.528C>G; p.Tyr176Ter change in exon 1 of the *NR0B1* gene has been listed before (244), but not in a *NR0B1*-related disease. As the variant is not present in population databases (ExAC) and causes a null effect, we classified it pathogenic, as well as the c.914CT;p.Glu305Ter variants (245). Regarding family studies, a cousin of the GN0078 patient presenting with adrenal insufficiency harboured the same variants. First relatives of GN0153 family were studied and only the mother and grandmother of the patient were carriers of the variants.

Patient GN0091 manifested a salt-wasting crisis and hypoglycaemia at 24 hours of life. Treatment with hydrocortisone and fludrocortisone was prescribed. At 8 months, testes and penis measured 1mL and 3.6 cm, respectively. Analytical levels are shown on Supplementary data 3. At imaging probes no adrenal glands were observed. Remarkably, mother of the patient referred abdominal pain, weakness, hyperpigmentation and dysfunctional uterine haemorrhages. Her biochemical levels were normal (Na: 131mEq/L; K: 5.8mEq/L; glucose: 68mg/dL and cortisol: 6.1µg/dL). In both proband and mother, the c.291delC deletion was found, in hemizygosis and heterozygosis, respectively, which is predicted to induce a frameshift variants (p.Glu98ArgfsTer166), as described (246).

Another small deletion was detected in the index case of family POL0301. This 2-year-old boy was examined because of an acute hyponatremia after presenting polyuria, polydipsia, asthenia and constipation for two weeks. At examination, testes of 0.5mL and normal penis were noticed. Treatment with fludrocortisone was prescribed. During hospitalization, cortisol deficiency was reported (Supplementary data 3). The patient started with the substitutive therapy. Genetic studies revealed the novel c.712_713delAC; p.Thr238LeufsTer60 variants in exon 1 of the *NR0B1* gene, which has been classified as pathogenic. The mother of the patient is a carrier of the mutation.

Patient POL0274 was diagnosed in Mexico of adrenal insufficiency at the age of one month and was under treatment with hydrocortisone and fluorohydrocortisone. Later, this Latin-American boy was referred in Spain for the study of macrogenitosomia

and pubic hair development (9 months). He presented with hyperpigmented genitals, hypertrophic penis (5cm), testes of 2mL and pubic hair. Basal hormone analysis showed low ACTH levels (<10pg/mL) with elevated testosterone (235.3ng/dL) under treatment. Ultrasound and computed tomography (CT) proved the absence of the left gland. Biochemical analyses were repeated after treatment stopped and showed cortisol levels under 0.1µg/dl at baseline and after ACTH stimulation test, while an increase in 17-hydroxiprogesterone (17OHP) was noticed (0.62ng/mL to 1.22ng/mL). At the age of 15 months, penis was 6cm and hydrocortisone doses were progressively increased with subsequent normalization of ACTH and testosterone values, while cessation of genital development. At follow-up (3 years), testosterone levels have been maintained at prepubertal levels, with no progression in penile length. The c.871T>A; p.Trp291Arg mutation was found. Functional studies performed by Lehmann *et al* demonstrated the pathogenicity of the protein change (247). Mother of the patient carried the variant in heterozygosis.

Clinical and genetic characteristics of patients GN0078, GN0091, GN0101, POL0285 and POL0274 have been published by our group (245, 248).

10.4.1.3.2. MLPA analysis

We found a complete deletion of *NR0B1* gene in two different cases. Case POL0285 was diagnosed with adrenal insufficiency at birth and treatment with hydrocortisone and fludrocortisone was established. There is no clinical data at this time due to loss of follow-up care. At 45 years, he had an acute adrenal insufficiency and hypogonadotropic hypogonadism. He presented testes of 2mL, prepubertal penis (3cm) and scarce body hair. There was a family background, as his uncle died at age 2 years. Patient GN0101 was studied due to the severe hypoglycaemia (20mg/dL) he presented at 3 years of age. He also had salt wasting. He received gonadotropin treatment at pubertal age to induce puberty and at 17 years, testes were 2-3mL and penis measured 8cm.

As no product was obtained after PCR-based amplification, a MLPA using the SALSA MLPA P185-C1 Intersex Kit was performed and showed a complete deletion of the 2 exons that are part of the *NR0B1* gene. As shown in Figure 25 both patients lack the presence of probes in exon 1 (148 and 234 nucleotide), exon 2 (310 nucleotide) and at 17.8kb upstream NR0B1 (454 nucleotide). Moreover, patient GN0101 had a bigger deletion as probe at 126.3kb upstream NR0B1 (481 nucleotide) was not present. The complete deletion of the gene was first described by Guo *et al* in 1995 (249).

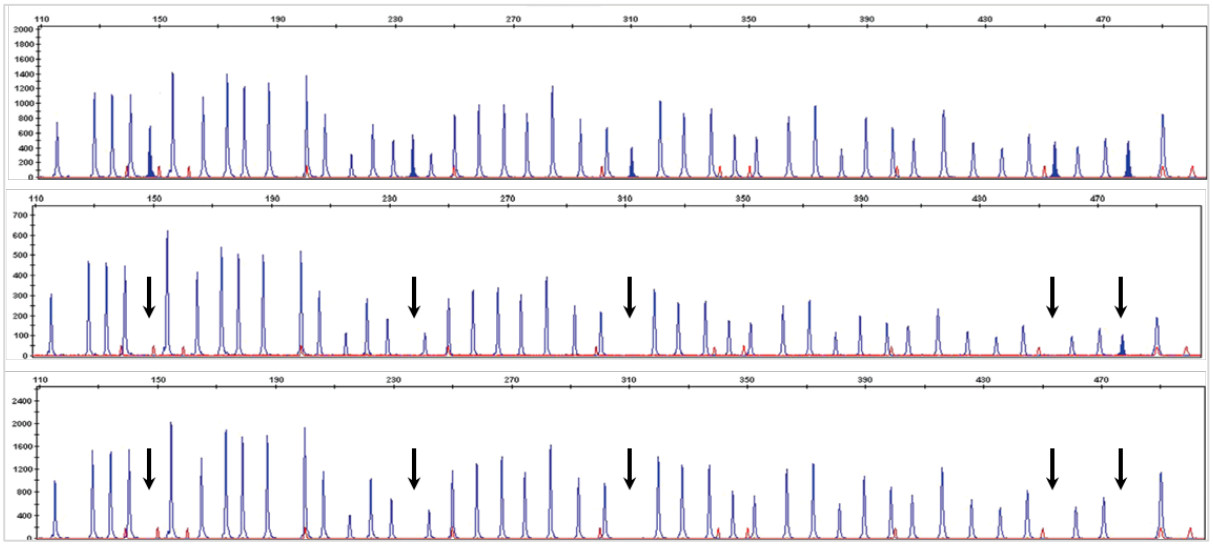

Figure 25. MLPA analysis data in patient POL0285 (middle panel) and GN0101 (lower panel) against a control (upper panel). Arrows indicate the probes for *NR0B1* gene.

Then, we used fluorescent markers to determine the extension of both deletions. As shown in the Figure 26 both patients lack the DXS1083 microsatellite, located at chromosome X: 30,361,221-30,361,359 position. Case POL0285 does not present the DSX8039 marker too, which is placed at chromosome X: 29,977,924-29,978,270. Combined with the MLPA assay, we could verify that POL0285 presented a contiguous gene deletion affecting *NR0B1* and *MAGEB1*genes (Chromosome X: 30,261,847-30,270,155). Moreover, in this case the deletion upstream *NR0B1* was up to 17.8kb, whereas deletion in GN0101 did not affect any other gene. Samples of the parents were not available for studying.

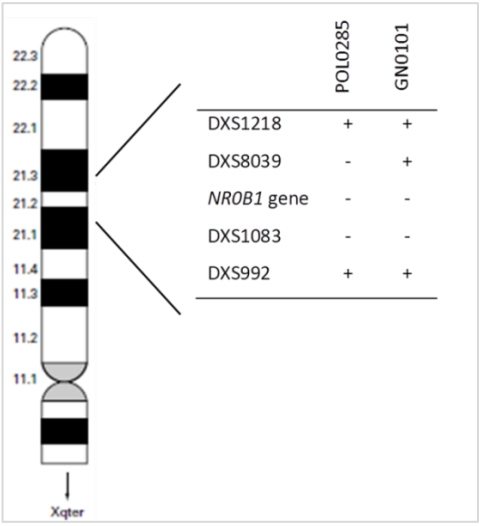

Figure 26. Analysis of genetic markers nearby *NR0B1* gene at Xp21.2. The presence or absence of the loci is represented by a plus or minus, respectively. Modified from Muroya K *et al*, 1999 (250).

10.4.1.4. *WT1* variants

We identified five different variants in *WT1* gene. Of these, 2 occurred in isolated DSD with Androgen Insensitivity Sundrome (AIS) and precocious puberty. The remaining three happened in WT1-associated syndromes.

10.4.1.4.1. *WT1* gene variants in isolated DSD

The GN0150 child presented normal female external genitalia and bilateral inguinal hernia when she was 3 years old. Surgery and subsequent histology revealed testicular tissue. Karyotype was 46,XY. At that moment uterus was not observed by ultrasound. At recent follow-up when she was 10 years old, imaging probe was repeated and a small uterus with normal kidneys was shown. Biochemical and hormonal studies revealed more or less normal values, compared to male age-controls, only elevated FSH was remarkable. *AR* was first studied and was negative, then we included the patient in the targeted gene panel analysis. We found the novel c. 223G>A; p.Glu75Lys sequence change in heterozygosis, as well as in her healthy mother.

Case GN0156 was studied because an increase in genital development and pubic hair was noted. At examination (age 7 years) this Indian patient presented pubertal genitalia with 10mL testes and a penis of 8cm. At MRI, a pituitary adenoma of 5x3mm was observed. Clinical analysis at this time indicated elevated basal testosterone (406ng/dL) and LH after LHRH stimulating test (25.03U/L). Although a precocious puberty was suspected, the genetic study of the *LHCGR* gene was negative. The novel c.545T>A; p.Met182Lys variant in the *WT1* gene in heterozygous state was detected. As the child was adopted no family history could be recorded.

Both novel variants were classified as of unknown significance due to the evaluation of disease-causing potential (Table 16, page 161) and multiple genome alignment (Supplementary data 14). Quality values of the NGS run are shown on Supplementary data 10.

10.4.1.4.2. *WT1*-associated syndromes

Two patients were diagnosed with a Frasier syndrome. Case GN0009 was admitted when she was 12 years old because of severe headaches, vomiting and fatigue for 15 days. Blood pressure was 180/120mmHg and a bilateral papilledema was detected by fundoscopy. Imaging probes and subsequent biopsy evidenced a nephronophthisis. She began treatment and later, haemodialysis due to renal insufficiency (13y). A renal transplant was performed but kidney rejection took place

after a few days. Since then, she has been treated with several drugs and receiving haemodialysis. At 17 years of age, her pubertal development was delayed with a bone age of 10 years. Ultrasound revealed a normal uterus (72mm) but no gonads were seen. Karyotype was 46,XY and biochemically, gonadotropins were elevated while oestradiol level was undetectable. At this time growth hormone (GH) treatment was prescribed because she was 137.4cm tall. One year later it was stopped since growth response was poor (1.5cm). At the age of 24 bilateral streak gonads were removed and a microscopic gonadoblastoma was reported. The second patient (GN0132) is an 18-years-old female with gonadal dysgenesis, renal insufficiency and gonadoblastoma. One year later, laparotomy was done and gonadal remnants were removed. Caryotype was 46,XY and Müllerian remnants including uterus were present. She started oestrogenic treatment until she was 33 years old. She had a second renal transplant.

Case GN0009 was a heterozygous carrier of the c.1447+5G>A change, while GN0132 carried, also in heterozygosis the c.1447+4C>T change. Both mutations are located in intron 9 and produce an alternative splicing of the protein. Only mother of GN0009 was studied and did not harbour the mutation. Both alterations have been previously associated with Frasier syndrome (251) (17).

Patient OT0567, from Tunisia, presented bilateral cryptorchidism at birth. When this male was 2 years old he was diagnosed with a bilateral Wilms' tumour due to a hypertensive crisis. He had no metastasis and began chemotherapy. After treatment, progression of masses was observed and was treated with chemotherapy. Then, surgery was suggested. Few months later he was referred to a Spanish hospital with hypertension and Denys-Drash syndrome was suspected due to proteinuria in nephrotic range. In order to discard a *WT1* gene mutation a DNA sample was sent to our laboratory before surgery. We found a partial *WT1* gene deletion in one allele, from exon 7 to 10 by NGS. We validated this CNV using the commercial MLPA containing probes for *WT1* and *AMER1* genes. As shown in Figure 27, the peak size of the probes in exon 7 (probe at 239 nucleotide), 8 (153 nucleotide), 9 (166 nucleotide) and exon 10 (185 nucleotide) are reduced, which indicates the deletion of exons 7 to 10 in one allele of patient OT0567. The parents of the patient were normal. Partial and entire deletion of the *WT1* gene has been reported before and related to isolated Wilms' tumour (252), as well as to 46,XY gonadal dysgenesis, Wilms' tumour and gonadoblastoma (253). During surgery to remove bilateral tumours, testes were found in abdominal cavity and one was visually noted to be hypoplastic. Gonads were not removed at the moment but will be in the near future.

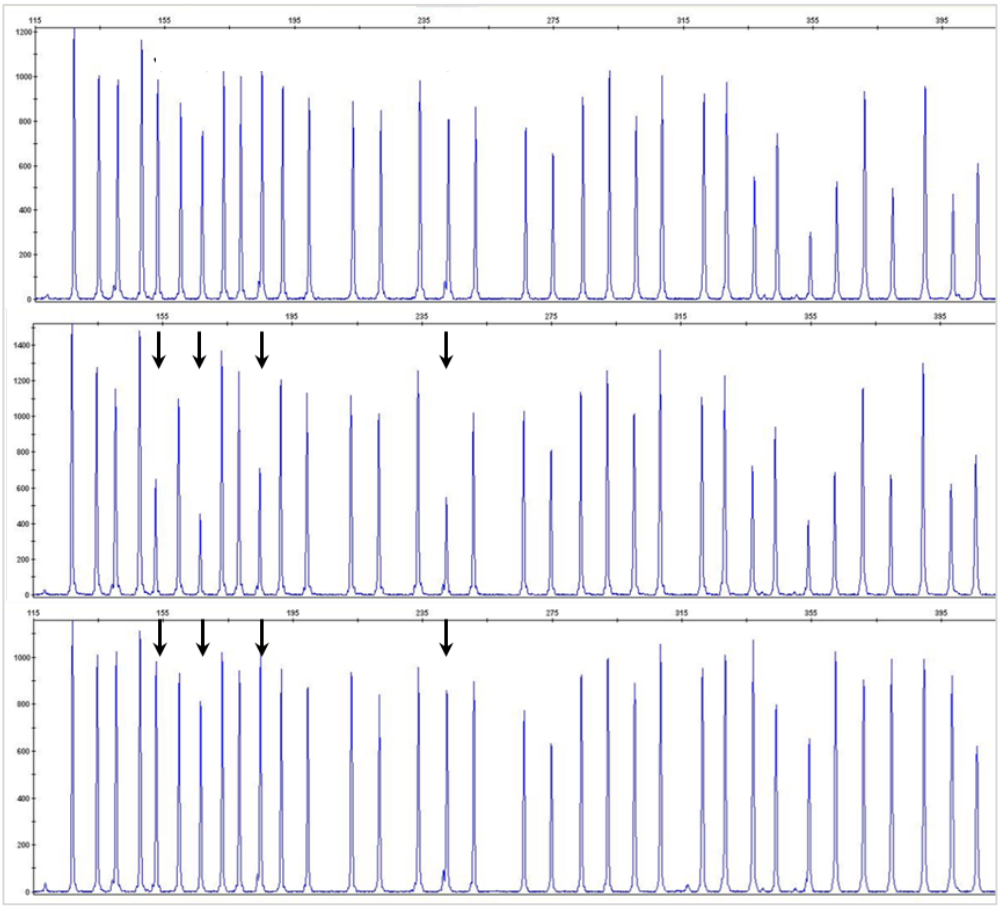

Figure 27. MLPA profile illustrating the presence/absence of LOH in OT0567 family. Upper panel, black arrows indicate the probes for exon 7, 8, 9 and 10 of the WT1 gene in a negative control. Middle panel, MLPA profile of the index case. The peak size of probes for exon 7 to 10 is reduced. Lower panel, normal MLPA profile of the patient's mother.

10.4.1.5. *MAP3K1* variants

We identified a *MAP3K1* gene variant in a 46,XY patient with complete gonadal dysgenesis. Female external genitalia, absence of thelarche and pubarche III were noted. This 16-year-old patient (GN0011) was examined due to primary amenorrhea. No uterus was visualized at ultrasound. She referred that she had grown during the last year when pubarche started. Gonadotropins were elevated (FSH:145U/L; LH:45U/L) in the presence of low testosterone (1.0ng/dL) levels (Supplementary data 3). Bilateral gonadectomy was planned and substitutive oestrogenic treatment would be then prescribed. Patient's father and brother reported delayed pubertal development.

Firstly, *SRY* and *NR5A1* genes were amplified and studied by traditional sequencing but no change was detected. Then, NGS identified a heterozygous *MAP3K1* c.2291T>G change in exon 13, which is predicted to result in a p.Leu764Arg variants. Parents and brother were studied and did not carry the variant, thus is a *de novo* change in the patient. This variant has been previously listed in a publication (82).

However, no functional studies have been done yet probing that the variants causes the clinical features in this type of patients.

10.4.1.6. *GATA4* variants

This male patient (GN0171) was noted at birth to have micropenis and bilateral cryptorchidism. At 11 years of age he still presented micropenis (3.5cm), testes were non-palpable and severe obesity (Body Mass Index (BMI) 36.6, +6.4SD) was observed. His gonadotropins were elevated with low testosterone. Karyotype was 46,XY. Magnetic resonance imaging showed both testes in inguinal canal. At 14 years, replacement testosterone treatment was started with gradual increasing doses. After treatment, LH suppression and normalization of FSH was observed. At 18 years, azoospermia was found. At recent follow-up at 21 years of age, he presented with severe obesity (BMI 44.9, +6.8SD) and short stature (156.1cm). Penis (8cm) was buried in subcutaneous fat and left testis was in scrotum (0.5ml). Biochemically, he presented normal gonadotropins with a testosterone of 11.8ng/ml under treatment. Ultrasound showed an ovoid structure (17mm × 16mm) in right inguinal canal corresponding to atrophic testis. Currently, laparoscopy is planned to investigate the right testis and biopsy both to assess the malignancy risk. Echocardiography revealed no congenital heart disease (CHD).

We found the novel *GATA4* c.677C>T nucleotide change, which is predicted to replace a proline with leucine at position p.226 (p.Pro226Leu). The same heterozygous variant was also found in the healthy mother. The pathogenesis of the *GATA4* gene variant was studied *in vitro* (See 10.5, page 200) and although prediction software and protein alignment (Supplementary data 14) suggested a deleterious effect, transactivation studies showed activation of the CYP17 promoter similar to wild type, thus we classified the p.Pro226Leu protein change as a VUS. The clinical and molecular characterization of the variant found in the index case of the GN0171 family has been described in a publication by our group in 2018 (25).

10.4.1.7. *WWOX* variants

Sequence variants in the *WWOX* gene were found in two patients with 46,XY DSD, presumably with complete and partial gonadal dysgenesis diagnosis.

GN0020 female went to medical consultation for short stature and absence of pubertal development. Physical examination of the patient at the age of 13 years revealed little pubic hair, without any other obvious findings. Analytical results found normal 17-hydroxiprogesterone and DHEA-S (Dehydroepiandrosterone-sulfate) levels after ACTH test, while oestradiol was below age and sex-range (<5pg/mL). Ultrasound

(US) showed no gonads. After being diagnosed, she received replacement treatment with oestrogens and gonadal remaining was removed. At follow up (17 years), she reported normal menses but breast development was unsatisfactory. The second patient (GN0203) harbouring a *WWOX* variant was born in Syria and was referred for gonadal dysgenesis evaluation. When this male child was 5 years old he presented micropenis with fused scrotum and bilateral cryptorchidism. US confirmed he had atrophic testes located in inguinal canal and no female internal genitalia were seen.

Targeted gene panel sequencing found two missense variants in the *WWOX* gene. Subject GN0020 carried the heterozygous c.1096C>G; p.Pro366Ala variants while GN0203 harboured, also in heterozygosis, the c.184G>A; p.Gly62Arg change in exon 3 of the WWOX gene. We were not able to achieve DNA samples of the family.

As none of the variants had been reported, prediction software analysis and comparison between different species was done. Allele variant c.1096C>G had not been listed before in population frequency studies, however comparison of the protein variability across species showed that the proline was slightly conserved (Supplementary data 14). Moreover, prediction software results were not certain. On the other side, c.184A allele was noted with a frequency below 0.01 (MAF <0.01) in 1000 Genomes Phase 3, ESP and ExAC, although the protein seemed to be highly conserved and the *in silico* analyses predicted as potentially pathogenic. Then, we classified c.1096C>G variants as a VUS and c.184G>A as a likely pathogenic change.

10.4.1.8. Variants in Oestrogen receptor 1 and 2

In total, 2 different variants were identified in two independent individuals by NGS in *ESR1* and *ESR2* genes.

10.4.1.8.1. *ESR1* gene

Female GN0198 was referred to clinician due to primary amenorrhea and failure of pubertal development at age 17 years. Her female external genitalia were normal with thelarche stage II and pubarche III at examination. No ovaries were found by magnetic resonance imaging and ultrasound, but a likely uterus remnant was observed. Karyotype was 46,XX. Biochemical and hormonal studies at presentation revealed elevated gonadotropins and a low oestradiol (<5pg/ml) (compared to female age-controls) (Supplementary data 3). Hormone replacement therapy was prescribed. She is healthy otherwise. We found a novel heterozygous variants in *ESR1* (c.1781C>T; p.Thr594Met), which is also carried by her mother. Due to the results given by the prediction software, the protein conservation analysis (Supplementary data 14) and the inheritance pattern, we classified the variants as a VUS.

10.4.1.8.2. *ESR2* gene

Case GN0207 was referred to clinician due to primary amenorrhea and pubertal development delay at age 14 years. At physical examination, this female presented normal external genitalia with Prader scale I and no puberty signs. The karyotype was 46,XX. Ultrasound was done and prepubertal ovaries and uterus were seen. Hormone replacement therapy was planned. Hormonal test revealed a hypergonadotropic hypogonadism (Supplementary data 3). Family history is remarkable for the following: mother was virilised and drugs were taken during pregnancy. We identified a novel variant in exon 5 of *ESR2* gene (c.661A>G; p.Arg221Gly) in heterozygosis with a population frequency <0.01 and as seen in Supplementary data 14, the arginine protein is highly conserved. Additionally, most of the prediction programs catalogue the variants as pathogenic. Samples of the parents were not available.

10.4.1.9. *DMRT2* variants

One single change was found in the *DMRT2* gene. Proband of GN0142 family was examined at birth because the obtained result of the karyotype during the amniocentesis (46,XY) and the female phenotype was discordant. This “female” newborn presented with female external genitalia with a 1-2cm clitoris, urinary meatus on base, non-palpable gonads and a bifid scrotum that seems to be poorly developed. Uterus and gonads were absent at US. Hormonal analysis when the patient was 22 days of life indicated the absence of Sertoli cells due to high levels of FSH (35.4U/L) and undetectable AMH levels.

Interestingly, her cousin presented at 15 years with primary amenorrhea, undeveloped breast and lack of uterus and ovaries. She underwent laparotomy for resection of her male remnant. Vaginoplasty was done and is under hormonal oestrogen treatment. She was diagnosed with 46,XY DSD gonadal dysgenesis.

Index case GN0142 was examined on the target gene panel for DSD and found a missense variant in exon 4 of the *DMRT2* gene (c.1607C>T; p.Ser536Leu). As shown below the amino acid in position 532 is not highly conserved (Supplementary data 14) and only 4 of the *in silico* programs we used categorize the genetic change as deleterious. Family studies were done in parents, maternal uncle and grandmother, and found that only the mother (II.2) is a heterozygous carrier of the c.1607C>T change, as well as the index case (Figure 28). The genetic study of the cousin with complete gonadal dysgenesis would clarify the possible pathogeny of this variant, meanwhile we classified it as a VUS.

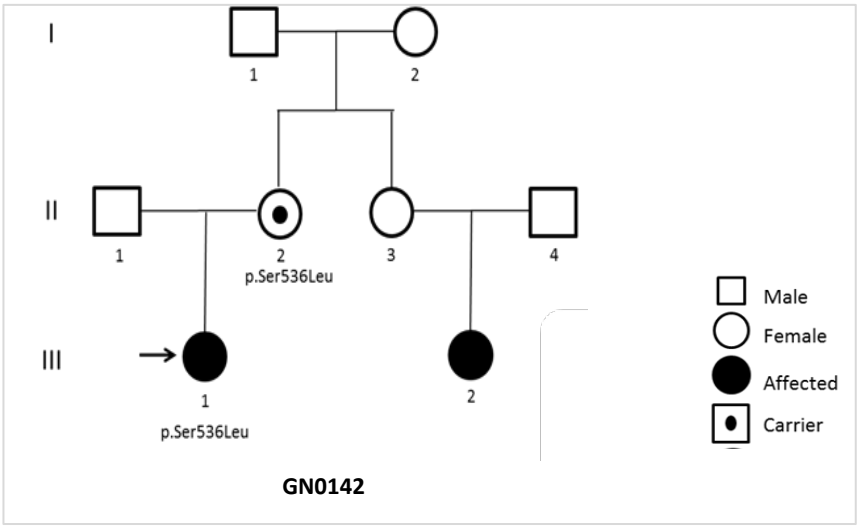

Figure 28. Pedigree and genotype information of the GN0142 family. Index case is indicated with an arrow.

10.4.1.10. *ZFPM2* gene

We found a variant in the *ZFPM2* gene in one patient. The GN0155 individual was born with ambiguous genitalia and was reared as a female. When the patient was 3 years old, surgery was done to adequate genitalia to assigned gender and a disgenetic testicle was removed. At age 7, the patient had a left inguinal hernia. At puberty, the patient was spontaneously masculinized and began expressing male identity. At age 20 years, a karyotype study was done and found 46,XY. Due to low levels of androgens, treatment with testosterone was prescribed (37 years old). At examination, he presented male normal phenotype, obesity and lipomastia. The scan showed remains of uterus and structures looking like copus cavernosum in the perineum. Testosterone levels were very low when removing treatment (1.9ng/dL).

Targeted gene panel sequencing showed the c. 3077C>T; p.Ala1026Val change in heterozygosis in the exon 8 of the *ZFPM2* gene. This variant has not been described and no data is available in population databases. We classified this novel variant as a VUS regarding the incongruent results given by the *in silico* programs and comparison of the protein across species (Supplementary data 14). Moreover, familial studies were not performed.

10.4.1.11. *MAMLD1* gene

We identified a novel *MAMLD1* gene variant in a patient with a gonadal dysgenesis and a 46,XX karyotype. Subject GN0154 was examined because of primary amenorrhea at age 15. She had normal female external genitalia and thelarche at stage II-III. Biochemical studies indicated a hypergonadotropic hypogonadism

(Supplementary data 3). A small uterus was identified by ultrasound and then confirmed by laparoscopy. Fallopian tubes and small ovaries were also observed. Histology revealed many immature primordial follicles and follicular atresia. The patient began with the oestrogenic treatment at that time and continued together with progesterone.

SRY was first studied but no gene amplification was observed. Large rearrangements in the *NR5A1, CXorf21, NR0B1, WNT4, SOX9, SRY* and *ZFY* genes were also discarded by MLPA. Then GN0154 was examined on the target gene panel and found in heterozygosis the missense c. 2009C>T; p.Thr670Ile change in exon 5 in the *MAMLD1* gene. We classified this variant as a VUS. The frequency of the T allele is <0.01 and has not been cited in ClinVar. However, the *in silico* programs disagree and comparison of the protein across species didn't find any similarity between species although no protein homologue is found in many of them (Supplementary data 14). Unfortunately, family studies were not done and therefore, prediction of disease-causing effect is more difficult.

This patient was included in an exome sequencing study to analyse additional genetic variations in cases with *MAMLD1* sequence changes. Flück *et al*, together with our group, identified 14 candidate gene variants that could theoretically contribute to the patient's phenotype (91).

10.4.2. Findings in genes related to genital differentiation

10.4.2.1. *AR* gene

In our DSD cohort a total of 15 different changes in 16 independent patients were identified in the *AR* gene. Twelve individuals were genetically diagnosed after Sanger sequencing, 3 by NGS and one case needed further confirmation due to an exon deletion. Mainly, they had been previously reported, except for three, which were classified either as pathogenic or likely pathogenic (Table 15, page 157).

10.4.2.1.1. Sanger sequencing of the *AR* gene

Among the cases with an AR defect, 12 patients were studied by sequencing the 8 exons that constitute the *AR* gene

Three patients (GN0035, GN0125 and GN0139) were examined by the clinician during puberty due to the presence of primary amenorrhea. They all had bilateral inguinal hernias at birth or during the first months of life. The absence of female

internal genitalia was confirmed by ultrasound in cases GN0125 and GN0139. Karyotype was 46,XY in all the three patients.

The suspicion of androgen insensitivity syndrome was set with the finding of an *AR* variant. Patient GN0035 presented the c.2522G>A; p.Arg841His mutation, described for the first time by McPhaul *et al* in 1992 (254), while GN0125 harboured the nonsense mutation c.865G>T (p.Glu289Ter) (255). Finally, in GN0139 we identified the novel T to G change in position 2642 (c.2642T>G) in exon 8, this p.Leu881Arg variants has not been described before but has been classified as likely pathogenic because of the results given by the prediction programs and the residue seems to be highly conserved across different species (Supplementary data 14). The mother carries the variant in heterozygosis.

The other patients were born with female external genitalia and presented inguinal hernias or masses located in inguinal canal. Foetal karyotyping was done in some individuals for another medical reason, such as phenylketonuria family history in case GN0177. Ultrasound (US) was performed and testicular tissue was defined after gonadectomy or biopsy. Probands of families GN0055 and GN0076 carried in hemizygosis the c.2710G>A; p.Val904Met and c.2566C>T; p.Arg856Cys mutations, respectively (254). The mothers of the patients were carriers of the mutations, but surprisingly mother of proband GN0055 presented the mutation as a mosaicism. More missense mutations were determined in GN0018 (p.Arg775Cys) (256), GN0024 (p.Asp696Asn) (257) and GN0177 (p.Ser434Phe). Unfortunately, family studies were not done in these cases. We also found a novel frameshift variant in proband GN0112, a deletion of a cytosine at position c.827 changed the reading frame (p.Pro276HisfsTer20) and thus we considered it as pathogenic. This change was also found in her healthy mother in heterozygosis.

Furthermore, patient GN0080 harboured another frameshift variant. This female child was examined because of the suspicion of AIS. Karyotyping at amniocentesis was 46,XY. At 5 years of age, testicular tissue was reported after gonadectomy and US imaging showed absence of uterus and Müllerian ducts. We found the novel c.298insC; p.His100ProfsTer3 variants, which was also present in her mother but not in her sister or father. As a disruption of the protein is happening, we considered the variants as deleterious.

Patient GN0146 had ambiguous genitalia and was intervened when she was 10 years to remove male gonads and began treatment. At recent follow-up, she is 19 years old and presents hypertrophic erectile organ and partial labial fusion. No Müllerian ducts were observed after imaging examination by ultrasound and Magnetic Resonance Imaging (MRI). She recently had a genito-plastic surgery. The study of the

AR gene revealed a c.2473C>A; p.Gln825Lys mutation. This variant had been observed before (258, 259). DNA samples of the parents were not available.

Male proband (GN0037) had perineal hypospadias, micropenis and bilateral testes located in scrotal folds at age 4. Testosterone was normal, as well as the rest of hormones. Karyotype was 46,XY. Enlargement of the pseudoclitoris was performed. At last examination (11years), the patient presented the same penis size (3cm). Testosterone and dihydrotestosterone values after stimulation slightly augmented (Supplementary data 3). The suspected diagnosis of PAIS was definite after the finding of the p.Phe726Leu change (260). Mother was a carrier of the same variants.

10.4.2.1.2. Targeted gene panel sequencing

Three more patients were genetically diagnosed with the targeted gene panel. All variants were missense and none was novel.

Patient GN0189 was noted to have palpable gonads in labioscrotal folds at birth. Otherwise, she presented with female external genitalia. Ultrasound ascertained the presence of bilateral testes of approximately 0.3mm. Hormonal values were in the age-and sex-range. Karyotype was 46,XY. Ultrasound was repeated at 3 months and a vaginal pouch was observed. However, no uterus or ovaries were identified. MRI confirmed the absence of uterus and ovaries. Testes (20x9mm and 18x8mm) were located in labia majora. The c.2567G>A; p.Arg856His mutation was found in hemizygosis. The mother of the patient was also a carrier of the *AR* change in one allele. Batch *et al* (1992) found and described this variant in a patient with AIS (261). Case GN0194 presented at 15 years old with primary amenorrhea. At examination, lack of pubarche, axilarche and breast development was observed. She was obese. Karyotype was 46,XY. Remarkably, her sister has a similar phenotype, although sample was not available. NGS revealed the c.2323C>T; p.Arg775Cys change in hemizygosis (256). On the contrary, patient GN0164 was referred to clinician because of hypergonadotropic hypogonadism, bilateral gynecomastia, micropenis and testis of 3mL at 36 years old. Karyotype was 46,XY. In the beginning a Leydig‘s cell hypoplasia was suspected and LH receptor gene was studied. However, nothing was found and the sample was included in the panel. We found the previously reported mutation c.2270A>G; p.Asn757Ser in hemizygosis (262). DNA samples of the parents were not available.

10.4.2.1.3. QMPSF methodology for the analysis of AIS

PCR amplification of *AR* gene exons was done due to the suspicion of androgen insensitivity syndrome in the index case of family GN0041. However, the amplification

of exon 2 was not possible and therefore confirmation was needed. This proband had bilateral gonadectomy performed at age 3 and was referred to clinician due to a suspected androgen insensitivity syndrome when she was 9 years old. She presented a 46,XY karyotype with normal female phenotype and prepuberal external genitalia. Elevated gonadotrophins levels were found, but testosterone and DHEAs were within normal range. Abdominal ultrasound (US) was done when the patient was 12 years old and revealed the absence of female internal genitalia. Treatment with ethinyl-oestradiol was prescribed and completely developed secondary sexual characteristics. Her older sister presented the same phenotype (Figure 29).

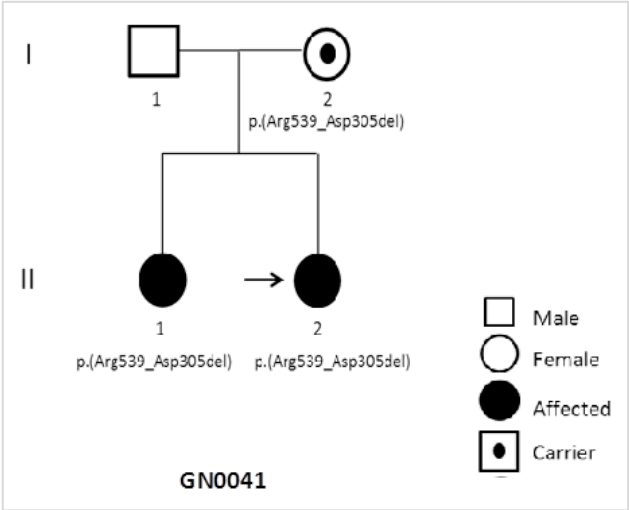

Figure 29. Pedigree and genotype information of the GN0041 family. Index case is indicated with an arrow.

Quantitative Multiplex Polymerase Chain Reaction of Short Fluorescent (QMPSF) was used to confirm the variation in the gene dose. As shown in Figure 29, we identified the deletion of exon 2 in the *AR* gene (c.(1616+1_1617-1)_(1768+1_1767-1)del; p.(Arg539_Asp305del)). Electropherogram of the patient showed the different peaks of the exons 3 to 8 in the *HNF1β*, and the complete absence of exon 2 in the *AR* gene. Co-segregation studies in the family revealed that the older sister (II.1) with a 46,XY karyotype, harboured the same hemizygous deletion while the mother (I.2) was heterozygous for the gene dosage variation and the father (I.1) was normal (Figure 30). Deletions of the exon 2 in *AR* gene, as well as partial gene deletions containing exon 2 or complete deletions of the gene, have been previously described in patients with a defect in androgen action (263).

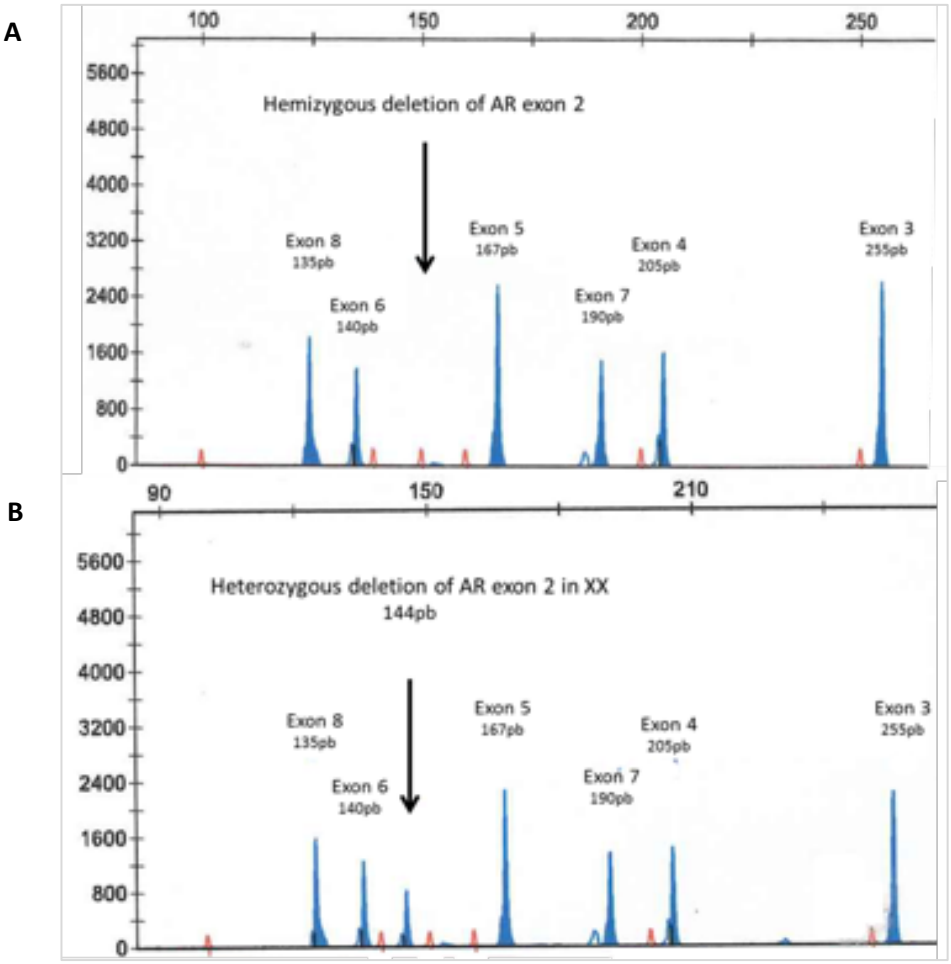

Figure 30. Results of the QMPSF assay in GN0041 family showing the deletion of exon 2 in *AR* gene. A, complete deletion of exon 2 in *AR* gene happening in the index patient and in her sister, both diagnosed with CAIS. B, the peak at 144pb revealls a heterozygous deletion of exon 2 and a 46,XX karyotype in the mother of the index patient.

10.4.2.2. *SRD5A2* gene

We identified recessive mutations in the *SRD5A2* gene in two patients.

Proband individual in family GN0046 was a female in which a bilateral gonadectomy was performed during the first years of life after echography had revealed the presence of inguinal testes. Immature testes and deferent ducts were found. Karyotype was 46,XY. We found a heterozygous A to G transition located at c.377 in exon 2 of the coding *SRD5A2* sequence, replacing a glutamine with arginine at position p.126 (c.377A>G; p.Gln126Arg). The same heterozygous variant was found in the healthy father. This change has been reported before (264). Because of the autosomal recessive inheritance of the gene we performed a QMPSF test to find a gene dosage variants and ensure the phenotype presented by the patient. We designed labelled primers for the 5 exons of the gene (Supplementary data 6) and found a deletion in the exon 1 in one allele (c. (-1+1_1-1)_(281+1_280-1)del;

p.(Met1_Arg94del)). Both parents were studied and did not carry the deletion, thus this variant is a *de novo* change in the patient (Figure 31). This deletion has been previously reported (265). Genetic findings demonstrated that the patient could fit in an androgen biosynthesis defect diagnosis due to a 5α reductase deficiency, instead of an AIS syndrome as firstly supposed.

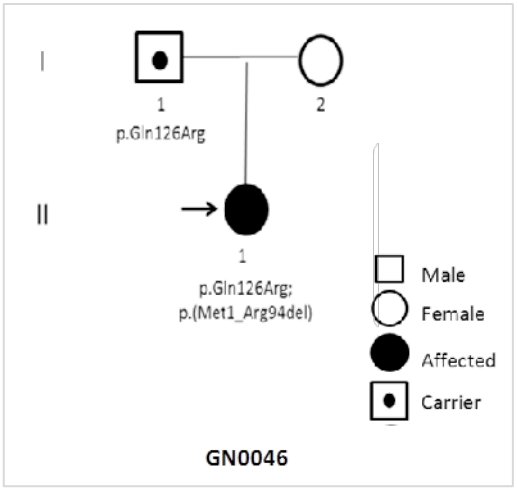

Figure 31. Pedigree and genotype information of the GN0046 family. Index case is indicated with an arrow.

Patient GN0186 was referred to clinician because of high stature at age 14 and elevated testosterone levels (Supplementary data 3). Two years later, this female presented with primary amenorrhea. Karyotype was 46,XY. At medical examination, an underdeveloped breast, pubarche at Tanner stage 4, hypertrophic erectile organ and two palpable masses in inguinal canal were noticed. Pelvic MRI showed two structures suggestive of testes in inguinal canal and a rudimentary vagina of 1cm. Biochemically, she presented normal gonadotropins with a testosterone of 4ng/dL and high AMH (471ng/mL) levels. NGS revealed the c. 271T>G> p.Tyr91Asp change in homozygosis in the exon 1 of the *SRD5A2* gene. Both healthy parents were studied and carried the variant in heterozygosis. This variant has been associated to steroid 5 alpha-reductase type 2 deficiency (266).

10.4.2.3. *LHCGR* gene

We found 6 cases harbouring different variants in *LHCGR* gene, 5 were missense variants and one patient presented a whole gene deletion. Among them, 3 patients were referred to clinician due to a suspicion of precocious puberty and activating mutations were therefore found. In contrast, inactivating variants were observed in patients presenting ambiguous genitalia.

10.4.2.3.1. Gene sequencing in patients with precocious puberty

The first case (GN0068) presented precocious puberty at 3.5 years and family history of the disease. Patient GN0088 was also noticed at age 3 to have pubarche, penis and testes growth and high stature. 6 months later, he was 112cm tall, had 6mL testes and pubertal penis. Magnetic Resonance Imaging (MRI) showed a doubtful image of a pituitary adenoma. Gonadotropins and testosterone were elevated after leuprolide. Few months later, testes were 6 to 8mL and he had grown 3cm in 2 months. He began treatment with triptorelin for 28 days. MRI was repeated and adenoma was discarded. At 4.5 years of age began treatment with ketoconazole. When he was 5 years old, pubarche started to diminish and no axilarche was noticed. Penis was 2.5cm and testes 8mL. Treatment with triptorelin was stopped and the dose of ketoconazole was elevated. Candidate gene sequencing revealed the heterozygous c.1713G>T; p.Met571Ile and c.1193T>C; p.Met398Thr in the *LHCGR* gene in cases GN0068 and GN0088, respectively. The father of patient GN0068 was also a carrier of the mutation, however no clinical information was provided. On the contrary, genetic study of the parents of patient GN0088 showed that the change was *de novo.* These two mutations have been previously described in the same phenotype (267, 268).

Index case GN0157 was born in Equatorial Guinea and was referred to clinician due to the presence of adrenarche when he was 6 years old. Genital development was Tanner 2 and testicular volume 2-3mL. One year later, growth speed was realized as well as acne and body odour. Testes were 1-2mL. At this stage pathology of the adrenal gland was discarded and abdominal scanner was normal. At 9 years of age, he was at Tanner 3 and testes were 3mL. Genetic study showed the heterozygous cytosine to adenine change in position c.568 (c.568C>A; p.Gln190Lys). This variant has not been described before and the protein region is highly conserved (Supplementary data 14). We classified the variants as likely pathogenic. Family studies showed that the father is a carrier of the variant.

10.4.2.3.2. NGS and aCGH to study other phenotypes caused by a *LHCGR* gene variant

We detected variants in *LHCGR* gene in another 3 individuals presenting with different phenotypes. Patient GN0147 was born as a female in Algeria and at 7 years of age had surgery because of a right inguinal hernia. Histology revealed it was a normal teste. She was referred to clinician in her birth country due to primary amenorrhea and provided reports with ovarian and uterine agenesis. Thelarche was spontaneous when she was 13-14 years old. The patient always had male identity. At 24 years of age, the patient presented left inguinal hernia which was a normal teste. When he escaped to Spain clinical examination was done. Karyotype was 46,XY,

hypergonadotropic hypogonadism was observed (T: 0ng/mL after gonadectomies) and US revealed vaginal outline, without uterus and annexes. At 35 years of age, patient presented ambiguous external phenotype, thelarche II-III, female external genitalia, buried erectile organ and an opening for urethra and vagina. Hormonal analysis still indicated a hypergonadotropic hypogonadism (Supplementary data 3). Due to the suspicion of AIS, *AR* gene was studied first but no variants were found. We found by NGS the c.757T>C; p.Ser253Pro variant in L*HCGR* in a homozygous state. Prediction software programs and comparison of the proteins revealed that the region is highly conserved (Supplementary data 14). Therefore, the identified *LHCGR* gene variant was predicted as pathogenic. This prediction was confirmed by testing the transcriptional activity of the variants *in vitro*.

Patient GN0171 has been described before due to the presence of a novel *GATA4* gene VUS variant. We also found in *LHCGR* the reported c.1660C>T;p.Arg554Ter substitution in heterozygosis (269). The mother of the patient did not present this change. The deleterious effect of the alteration was confirmed in the expression studies we performed (See 10.5, page 200).

A last patient was diagnosed with a LH receptor defect by the customized gene panel and validated then by aCGH. Patient GN0034 was noted to have bilateral inguinal hernia at 8 months of age. Her phenotype was female with normal external genitalia. After herniorrhaphy, testicular tissue was found. Absence of upper vagina and uterus was noticed. Her testosterone levels were undetectable with normal levels of gonadotropins (LH: 2mU/mL and FSH: 1mU/mL) (Supplementary data 3). AIS was suspected and karyotype was confirmed to be 46,XY. At 4 years of age, an echocardiography demonstrated the lack of uterus and ovaries. At 9 years old, biochemical and hormonal studies revealed high levels of testosterone with undetectable LH. After puberty gonadotropins, testosterone (60ng/dL) and dihydrotestosterone (0.4ng/mL) increased. After gonadectomy, the patient began treatment with oestradiol, and two years later at medical consultation (15 years old) normal pubarche and thelarche was noticed. *AR* gene was first studied by Sanger sequencing and then, *NR5A1, HSD17B3* and *SRD5A2* were discarded. Finally, it was analysed by the customized gene panel where we found that the *LHCGR* gene had not been covered. PCR of all the exons was done with no amplification of the regions. We made an aCGH and found a 77.45kb deletion in the short arm of chromosome 2 (2p16.3) from base pairs 48,905,663 to 48,983,208 (Figure 32). Genes located in this region included *LHCGR*, among others as *GTF2A1L* (General Transcription Factor IIA Subunit 1 Like) or *STON-GTF1A1L*. The absence of the entire *LHCGR* gene has been previously reported by Richard *et al* in 2011 (270). Unfortunately, DNA samples from the parents were not available.

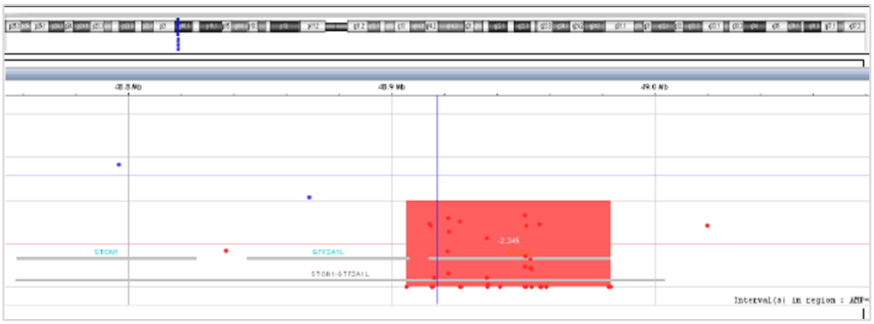

Figure 32. aCGH profile from GN0034 patient. The figure shows a 77.54Kb deletion at 2p16.3 position in chromosome 2.

10.4.2.4. *CYP17A1* gene

We found a unique variant in the *CYP17A1* gene in the RE0045 family. This index female patient presented at puberty with primary amenorrhea and lack of pubertal development. Karyotype was 46,XY. She had presented high blood pressure since childhood. At 17 years old, she was diagnosed with a bilateral pheochromocytoma and a left adrenalectomy was performed later that year. Histology revealed a left adrenal gland of 25g and 5.5x 5cm in size and indicated an adrenal cortical and medullar hyperplasia. After surgery blood pressure was normalized and urinary catecholamines values descended and have been maintained normal. In order to discard a medullary thyroid carcinoma, and then a Multiple Endocrine Neoplasia type II (MEN2), basal calcitonin levels and the response to pentagastrin stimulation studies were performed every year. In 1999, at the age of 19 pentagastrin test is considered pathologic and a total thyroidectomy is done. The consequent pathological anatomy is compatible with the hyperplasia of parafollicular C cells of the thyroid gland. Since then, patient has been receiving thyroid hormone replacement therapy and during the following years has presented normal urine values of noradrenaline and normetanephrine. An abdominal CT scan was performed when the patient was 25 years old, which revealed a normal right adrenal gland and left adrenal nodule of 8mm, suggesting an atrophic kidney. At 32 years old, she was gonadectomised and biopsy showed immature testes with Leydig's cell hyperplasia. No uterus was observed. The genetic study of *RET* (REarranged during Transfection) protooncogene (OMIM 164761) was negative. Alterations in *NR5A1, AR* and *WT1* had been previously ruled out.

Remarkably, the older sister (II.1) of the index case was referred to clinician at age 17 with primary amenorrhea and absence of pubertal development due to a hypergonadotrophic hypogonadism. Karyotype was 46,XX. She had uterine corpus agenesis and cervix hypoplasia and was therefore treated with oestrogens. At 28 years of age, she was noted to have elevated blood pressure. Left and right side adrenalectomy was done when the patient was 33 and 34 years, respectively, due to

the presence of bilateral adrenal hyperplasia. MEN2 was discarded. Genetic analysis of *WT1* and *RET* genes were negative.

DSD targeted gene panel sequencing identified a homozygous *CYP17A1* c.1246C>T (p.Arg416Cys) nucleotide change in the index case (II.2). This variant has been previously listed by Takeda *et al* in 2001 in a patient presenting a defect in 17alpha-hydroxylase (271). The patient's affected sister (11.1) had the same variant in homozygous state. Mother (I.1), younger sister (II.3) and her 15 years old-twin daughters (III.1 and III.2) were carriers of the same heterozygous *CYP17A1* variant (Figure 33). Sample of the father was not available.

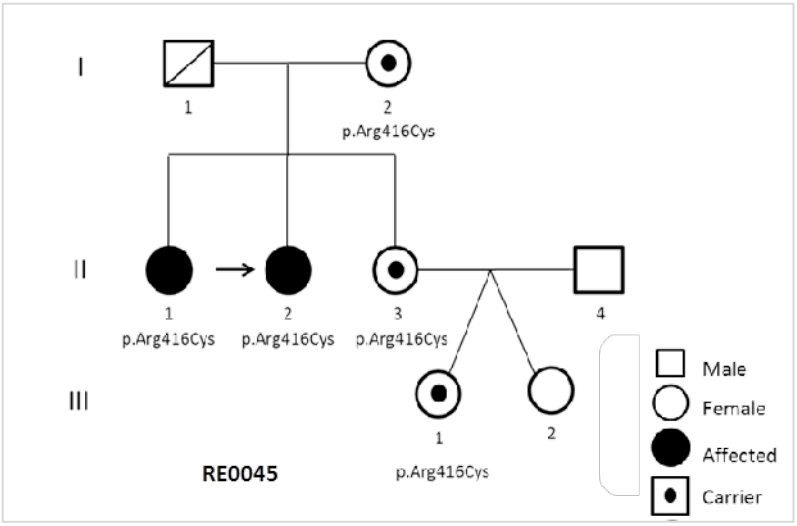

Figure 33. Pedigree and genotype information of the RE0045 family. Index case is indicated with an arrow.

10.4.2.5. *HSD17B3* gene

We only found one mutation in the *HSD17B3* gene, encoding for 17-beta hydroxysteroid dehydrogenase 3. The proband (GN0038) was born with female external genitalia and was assigned female at birth, although foetal karyotyping was performed because of parent's consanguinity and indicated a 46,XY karyotype. When the individual was 10 years old, voice and changes in external genitalia were noticed. At examination, she presented hypertrophic clitoris and palpable gonads at inguinal canal which Magnetic Resonance Imaging (MRI) confirmed as possible testes. Vagina was small (2cm). Laparoscopy showed the absence of female genitalia. Later, clitoroplasty was performed and gonadectomy of both testes. She began treatment.

Sanger sequencing study was done and revealed a C to T change in position c.845 (c.845C>T) with produced a proline to leucine modification at the protein level (p.Pro282Leu). The mutation was found in homozygosis and had been previously

reported by Andersson *et al* in 1996 (130). The mother of the patient was a carrier of the same variants while father could not be analysed.

10.4.2.6. *STAR* gene

Two missense variants were found in the *STAR* gene in patients GN0185 and GN0028.

The GN0185 newborn was referred after birth because of ambiguous genitalia (Supplementary data 2). This "female" presented with a small penis (1.5cm) and a midshaft hypospadias. Labia majora were well feminized but vagina was blind. Right gonad was palpable in inguinal canal. Karyotype was 46,XY. A small vagina and hypoplastic uterus displaced to the left (2x0.8x0.8) were identified by MRI. Laparoscopy showed a normal vagina in size and a small uterus which continuous to the right with a fallopian tube ending in a testis in inguinal canal, and to the left the fallopian tube encloses a small streak gonad. Hormonal studies at presentation revealed more or less normal values (compared to male age-controls), only elevated 17-hydroxiprogesterone and progesterone was remarkable (Supplementary data 3). One year later, both gonads were removed and showed a left streak ovary and a right dysgenetic teste. Clitoroplasty and vaginoplasty was also done. At recent follow up, she presented a slight pubertal development (11 years old) and the induction of puberty is planned. We identified a heterozygous guanine to timine nucleotide change at position 50 (c.50T>G; p.Met17Arg) in exon 1 of the *STAR* gene by NGS. Moreover, the father of the patient presented the same change in heterozygosis.

Regarding the *STAR* gene variant observed in patient GN0028, see 10.4.3.

10.4.3. Identification of additional gene variants

A total of 4 patients had more than one variant in a DSD-related gene (Table 17). Three of these patients had a 46,XY karyotype while the other (GN0198) had 46,XX. Interestingly, two individuals (GN0028, GN0042) had a pathogenic variant in *NR5A1* in combination with changes in genes acting in gonadal differentiation. On the other side, subject GN0171 harboured a pathogenic *LHCGR* gene mutation and a novel missense variant in *GATA4*.

Regarding GN0028 male, targeted gene panel sequencing revealed a second variant in addition to the *NR5A1* gene change. The novel, heterozygous C to T change at position 361 (c.361C>T; p.Arg121Trp) of the *STAR* gene was found in the patient and his father. As no *STAR* gene variants related to an isolated DSD phenotype have been described, we categorize these changes as of unknown significance, even if the

prediction tools and amino acid conservation studies estimated as deleterious (Supplementary data 14). Patient (GN0042) had a known variant in *NR5A1,* but we also found the heterozygous c.428C>T; p.Thr143Ile VUS variant in the testis development gene *AMH*. The mother of the patient did not harbour the sequence changes. Although prediction software and conservation studies lead to a non-pathogenic effect (Supplementary data 14), it is well known that SF1 regulates in Sertoli cells the expression of *AMH* gene and would fit with the phenotype found in this subject.

Table 17. Individuals with additional variants identified by NGS.

Family	Gene	Variant	Classification
GN0028	*NR5A1*	C.88T>A; p.Cys30Ser	Pathogenic
	STAR	c.361C>T; p.Arg121Trp	VUS
GN0042	*NR5A1*	c.614_615insC; p.Gln206ThrfsTer20	Pathogenic
	AMH	c.428C>T; p.Thr143Ile	VUS
GN0171	*LHCGR*	c.1660C>T; p.Arg554Ter	Pathogenic
	GATA4	c.677C>T; p.Pro226Leu	VUS
GN0198	*ESR1*	c.1781C>T; p.Thr594Met	VUS
	HSD17B4	c.524delC; p.Ala175GlufsTer26	VUS

VUS, variant of unknown significance.

Patients GN0171 has been previously described (See 10.4.1.6, page 178).

Among the patients with a 46,XX DSD, only GN0198 was found to have two variants of unknown significance in *ESR1* and *HSD17B4* gene. NGS showed in patient GN0198 the deletion of a cytosine in position c.524 (c.524delC) which leads to the substitution of an alanine and a change in the reading frame of the protein (p.Ala175GlufsTer26) in heterozygosis in *HSD17B4* gene. The mother of the patient is a carrier of this variant, located in exon 8 of the gene. GN0198 case also harboured a novel VUS variant in *ESR1*. We classified the variant as a VUS. Despite the effect that this change could cause in the length of the protein, no mutations have been related to 46,XX gonadal dysgenesis up to date. *HSD17B4* gene variants have been described in D-bifunctional protein deficiency and Perrault syndrome, a condition that alters hearing loss and causes abnormalities in ovaries. Two alterations have been associated to premature ovarian failure too (225).

10.4.4. Family studies and familial cases of DSD

We were able to achieve at least one first degree family member of 37 patients presenting with a rare genetic variants (45.6%).

Nine families among the familial cases described in our cohort presented an alteration in a DSD-related gene, however we could only study 5 families (Table 18). The two siblings in family GN0007 had gonadal dysgenesis and a hemizygous nonsense variant in the *SRY* gene. Moreover, the same pathogenic mutation in the *AR* was identified in the sister of index case GN0041, who presented identical phenotype. Regarding GN0075 family, 3 affected members harboured a reported variant in *NR5A1*. An activating mutation in the *LHCGR* gene was found in the GN0068 proband and his father, with family history of precocious puberty. Finally, the two affected siblings with gonadal dysgenesis in family RE0045 presented a homozygous missense mutation in *CYP17A1* gene. Although other familial cases of DSD were referred to the clinician, we were not able to study the affected family members (GN0078, GN0142, GN0177 and GN0186). A summary of the genetic findings in the affected and non-affected family members is shown in Supplementary data 15.

In other familial cases, no DSD-related variant was found.

Table 18. Analysed familial cases with a genetically positive variant

Case	Gene variant	Affected family members
GN0007	*SRY*, c.289C>T; p.Gln97Ter	Index case and sister
GN0041	*AR*, c.(1616+1_1617-1)_1768+1_1767-1)del; p.(Arg539_Asp305del)	Index case and sister
GN0068	*LHCGR*, c.1713G>T; p.Met571Ile	Index case and father
GN0075	*NR5A1*, c.250C>T; p.Arg84Cys	Index case, mother, aunt and uncle
RE0045	*CYP17A1*, c.1246C>T; p.Arg416Cys	Index case and sister

10.4.5. Yield of identified variants

Sequencing was carried out in the total DSD cohort (125 independent patients). At the beginning genetic analysis was done by a candidate gene approach. After the implementation of NGS in 2015, patients referred with a DSD were analized using a targeted gene panel as well as previous patients with negative results. In total, 89 patients were analysed by targeted panel (Figure 34).

Totally, we made a DSD-associated genetic finding in 53 patients (41.6%). Among these positive cases, 36 (69.2%) were identified by a candidate gene approach by PCR-based sequencing, aCGH, QMPSF or MLPA technique. The DSD panel provided a molecular diagnosis in 17 out of 89 patients (19.1%) tested by NGS with an unknown previous genetic diagnosis after candidate gene approach (Table 15, page 157). From

the total observations, clinically relevant rare variants were reported if the filtering process classified them as pathogenic or likely pathogenic (229), the variant was consistent with the inheritance pattern for that gene and with the patient's phenotype.

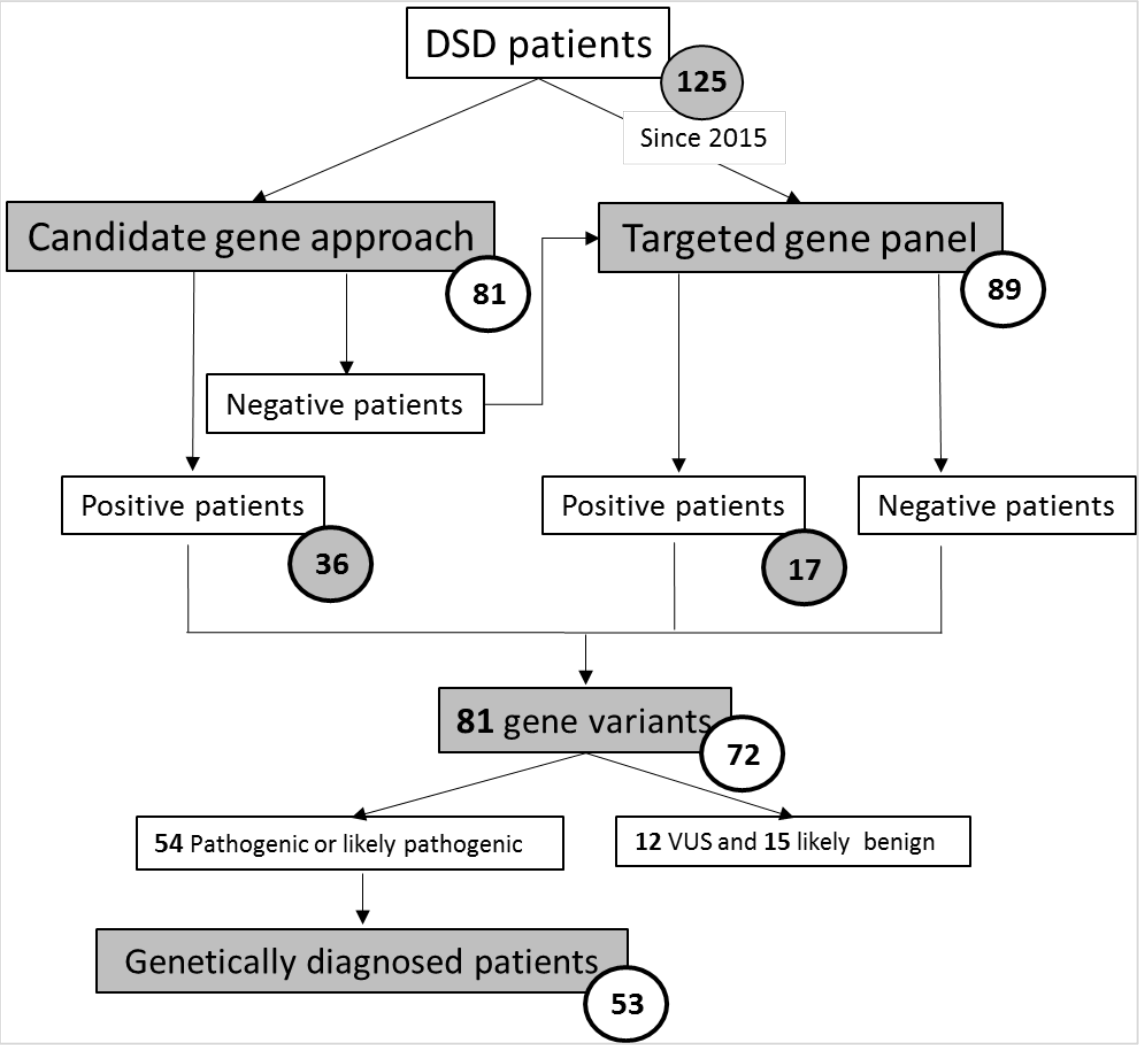

Figure 34. Schematic representation of the molecular study in the DSD cohort. Number of patients is given in cicles.

The diagnostic yield for this panel is therefore 17/89 for all samples. This yield is increased up to 22.0% for 46,XY DSD (15/68) while it only provided a molecular diagnosis in 10.5% of 46,XX DSD individuals (2/19). Variants of unknown significance were found in 11 patients (12.3%) and then, diagnosis could not be confirmed because of the ambiguity around the variant.

Several variants were identified in the remaining 61 patients, but none showed consistency with the phenotype or inheritance mode of the individuals. In these cases, DSD might be caused by other factors that need to be discovered, including those

patients in whom the Gly146Ala protein change in the *NR5A1* gene was observed. Although these cases are not discussed in this work they are being further studied.

10.4.6. Clinical diagnosis after molecular studies

After molecular analysis, initial diagnosis changed in some patients in which a pathogenic or likely pathogenic variant was identified.

Disorders of testicular development were suspected in 47 probands. In 4 out of 17 of those with a complete gonadal dysgenesis this was supported by the identified gene variant. On the other side, 22 patients presented with a suspected partial gonadal dysgenesis and only two cases were found to be genetically consistent. An *AR* gene variant was found in another index case. Two patients referred as gonadal dysgenesis were then reassigned to the androgen biosynthesis defect group due to the detection of *SRD5A2* and *CYP21A2* gene variants (Table 19), while in another case diagnosis was reinforced. The only case with an ovotesticular suspicion presented an *AR* variant and was therefore classified as AIS, however, none of the patients with gonadal regression harboured a gene variant. Twenty-one patients were categorised as AIS (Complete, partial or AIS), however an *AR* defect was only found in 12 of these cases and 4 had another variant. Disorders of androgen synthesis were confirmed in one patient with a *HSD17B3* deficiency suspicion and in three patients with luteinizing hormone defects and activating mutations in *LHCGR* gene. Regarding the group classified as others, adrenal insufficiencies were confirmed with the detection of *NR0B1* gene changes, as well as 3 out of 4 pathologies related to the *WT1* gene (Denys-Drash and Frasier syndrome). Among those individuals with a clinical suspicion of anorchia, cryptorchidism or 46,XY DSD, the genetic aetiology was set in 4 out of seven patients leading diagnosis to AIS in the case of anorchia suspicion, and three cases of 46,XY DSD to gonadal dysgenesis and LH-receptor defect. Overall, a genetic alteration was determined in 21.7% of those with a suspected disorder of testicular development (Table 20, page 198), 70.3% of those with a disorder of androgen synthesis or action and 62.5% of those classified as others.

All the individuals with a suspected 46,XX testicular DSD were genetically confirmed. Furthermore, two patients with a primary cause of 46,XX DSD had a SRY translocation and a likely deleterious change in *ESR2* gene and were reassigned to 46,XX testicular DSD and ovarian dysgenesis, respectively. The remaining cases with a 46,XX karyotype and gonadal dysgenesis, Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome or OT DSD were not confirmed. Overall, 36.8% of patients referred with a related ovarian disorder were molecularly diagnosed.

Table 19. Classification of our DSD cohort at clinical diagnosis and after molecular study in accordance to the Consensus statement on intersex disorder (172). Number of patients with a genetic finding in each group is indicated.

	DSD classification	At initial diagnosis	After molecular study	Patients with variants (%)
46,XY DSD	**Disorders of testicular development**			
	Complete gonadal dysgenesis	17	19	
	Partial gonadal dysgenesis	22	23	
	Gonadal dysgenesis	4	1	
	Gonadal regression	3	3	
	Ovotesticular DSD	1	0	
	Total	47	46	10 (21.7)
	Disorders in androgen synthesis and action			
	Androgen biosynthesis defect	6	9	
	Defects in androgen action	21	21	
	LH receptor defects	5	7	
	Total	32	37	26 (70.3)
	Others			
	Frasier and Denys-Drash syndromes	4	4	
	Severe hypospadias	2	2	
	Primary adrenal insufficiency	7	7	
	Unknown	7	3	
	Total	20	16	10 (62.5)
	Total 46,XY DSD	**99**	**99**	**46 (46.5)**
46,XX DSD	**Disorders of ovarian development**			
	Ovotesticular DSD	3	3	
	Testicular DSD	6	6	
	Gonadal dysgenesis	8	8	
	Total	17	17	6 (35.3)
	Androgen excess			
	Total	0	0	0 (0)
	Others			
	MRKH syndrome	3	3	
	Unknown	4	4	
	Total	7	7	0 (0)
	Total 46,XX DSD	**24**	**24**	**7 (29.2)**
Sex chr DSD	45X/46,XY, mixed gonadal dysgenesis	2	2	0 (0%)
	Total Sex chromosome DSD	**2**	**2**	**0 (0%)**
	Total DSD	**125**	**125**	**53 (42.4%)**

Chr, chromosome; DSD, disorders of sex development.

Table 20. Detail of the patients with a confirmed molecular diagnosis.

Case	Suspected clinical diagnosis	Affected gene	Results reported	Diagnosis after molecular study
GN0004	46,XX T DSD	*SRY*	Consistent with clinical features	46,XX T DSD
GN0007	46,XY DSD CGD	*SRY*	Confirmed diagnosis	46,XY DSD CGD
GN0009	Frasier syndr	*WT1*	Consistent with diagnosis	Frasier syndr
GN0011	46,XY DSD CGD	*MAP3K1*	Confirmed diagnosis	46,XY DSD CGD
GN0018	46,XY DSD PGD	*AR*	Consistent with clinical features	46,XY DSD CAIS
GN0023	46,XX T DSD	*SRY*	Consistent with clinical features	46,XX T DSD
GN0024	46,XY DSD	*AR*	Consistent with clinical features	46,XY DSD PAIS
GN0028	46,XY DSD	*NR5A1*	Consistent with diagnosis	46,XY DSD PGD
GN0034	46,XY DSD AIS	*LHCGR*	Consistent with clinical features	46,XY DSD LH RD
GN0035	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0037	46,XY DSD PAIS	*AR*	Consistent with clinical features	46,XY DSD PAIS
GN0038	46,XY DSD ABD	*HSD17B3*	Confirmed diagnosis	46,XY DSD ABD
GN0041	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0042	46,XY DSD AIS	*NR5A1*	Consistent with clinical features	46,XY DSD CGD
GN0046	46,XY DSD AIS	*SRD5A2* *SRD5A2*	Consistent with clinical features	46,XY DSD ABD
GN0054	46,XX T DSD	*SRY*	Consistent with clinical features	46,XX T DSD
GN0055	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0068	46,XY DSD LH RD	*LHCGR*	Supports diagnosis	46,XY DSD LH RD
GN0075	46,XY DSD	*NR5A1*	Consistent with clinical features	46,XY DSD PGD
GN0076	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0078	PAI	*NR0B1*	Confirmed diagnosis	PAI
GN0080	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0088	46,XY DSD LH RD	*LHCGR*	Supports diagnosis	46,XY DSD LH RD
GN0091	PAI	*NR0B1*	Consistent with diagnosis	PAI
GN0101	PAI	*NR0B1*	Supports diagnosis	PAI
GN0109	46,XY DSD GD	*NR5A1*	Confirmed diagnosis	46,XY DSD CGD
GN0111	46,XY DSD CGD	*NR5A1*	Confirmed diagnosis	46,XY DSD CGD
GN0112	46,XY OT DSD	*AR*	Consistent with clinical features	46,XY DSD CAIS
GN0123	46,XY DSD PGD	*NR5A1*	Consistent with diagnosis	46,XY DSD PGD
GN0125	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0132	Frasier syndr	*WT1*	Consistent with diagnosis	Frasier syndr
GN0133	46,XX T DSD	*SRY*	Consistent with clinical features	46,XX T DSD
GN0139	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0141	46,XY DSD CGD	*SRY*	Confirmed diagnosis	46,XY DSD CGD
GN0146	46,XY DSD PAIS	*AR*	Confirmed diagnosis	46,XY DSD PAIS
GN0147	46,XY DSD AIS	*LHCGR*	Consistent with clinical features	46,XY DSD LH RD
GN0153	PAI	*NR0B1*	Confirmed diagnosis	PAI
GN0157	46,XY DSD LH RD	*LHCGR*	Supports diagnosis	46,XY DSD LH RD
GN0159	46,XX T DSD	*SRY*	Confirmed diagnosis	46,XX T DSD
GN0164	46,XY DSD LH RD	*AR*	Consistent with clinical features	46,XY DSD AIS
GN0171	46,XY DSD	*LHCGR*	Consistent with clinical features	46,XY DSD LH RD

Table 20. Detail of the patients with a confirmed molecular diagnosis (Continuation).

Case	Suspected clinical diagnosis	Affected gene	Results reported	Diagnosis after molecular study
GN0177	46,XY DSD CAIS	*AR*	Consistent with clinical features	46,XY DSD CAIS
GN0186	46,XY DSD GD	*SRD5A2*	Supports diagnosis	46,XY DSD ABD
GN0187	46,XX T DSD	*SRY*	Confirmed diagnosis	46,XX T DSD
GN0189	46,XY DSD CAIS	*AR*	Confirmed diagnosis	46,XY DSD CAIS
GN0194	46,XY DSD AIS	*AR*	Confirmed diagnosis	46,XY DSD AIS
GN0203	46,XY DSD PGD	*WWOX*	Consistent with clinical features	46,XY DSD PGD
GN0207	46,XX DSD	*ESR2*	Consistent with clinical features	46,XX DSD
OT0567	Denys-Drash syndr	*WT1*	Consistent with diagnosis	Denys-Drash syndr
POL0274	PAI	*NR0B1*	Consistent with clinical features	PAI
POL0285	PAI	*NR0B1*	Confirmed diagnosis	PAI
POL0301	PAI	*NR0B1*	Confirmed diagnosis	PAI
RE0045	46,XY DSD GD	*CYP17A1*	Consistent with clinical features	46,XY DSD ABD

ABD, androgen biosynthesis defect; AIS, androgen insensitivity syndrome; CAIS, complete androgen insensitivity syndrome; CGD, complete gonadal dysgenesis; DSD, disorder of sexual development; GD, gonadal dysgenesis; OT, ovotesticular; PAI, Primary adrenal insufficiency; PAIS, partial androgen insensitivity syndrome; PGD, partial gonadal dysgenesis; RD, receptor defect; syndr, syndrome; T, testicular.

Due to the uncertainty around the pathogenicity of VUS variants initial clinical diagnosis could not be confirmed in such patients (Table 21).

Table 21. Detail of the patients with a variant of unknown significance (VUS).

Case	Suspected clinical diagnosis	Affected gene	Results reported
GN0020	46,XY DSD GD	*WWOX*	Consistent with clinical features
GN0028	46,XY DSD	*STAR*	Variant found in addition to NR5A1
GN0042	46,XY DSD AIS	*AMH*	Variant found in addition to NR5A1
GN0142	46,XY DSD GD	*DMTR2*	Consistent with clinical features
GN0150	46,XY DSD AIS	*WT1*	Consistent with clinical features, not with diagnosis
GN0154	46,XX DSD GD	*MAMLD1*	Conflicting evidence
GN0155	46,XY DSD AIS	*ZFPM2*	Consistent with clinical features, not with diagnosis
GN0156	46,XY DSD LH RD	*WT1*	Conflicting evidence
GN0171	46,XY DSD	*GATA4*	Variant found in addition to LHCGR
GN0185	46,XY DSD GD	*STAR*	Consistent with clinical features
GN0198	46,XX DSD	*HSD17B4*	Consistent with clinical features
		ESR1	Consistent with clinical features

AIS, androgen insensitivity syndrome; DSD, disorder of sexual development; GD, gonadal dysgenesis; RD, receptor defect.

10.5. *IN VITRO* FUNCTIONAL STUDIES

Transcriptional activity of different variants located in *GATA4*, *NR5A1* and *LHCGR* genes was tested to elucidate its potential pathogenic role in the development of the disease.

10.5.1. Functional characterization of the GATA4 transcription factor

In the literature, a single *GATA4* variant found in a patient with DSD and CHD (congenital heart defect) has been characterized before (27). The p.Gly221Arg variant, located in the N-terminal zinc finger domain of GATA4, had impaired transactivation activity on the AMH promoter.

Priorization of variants detected by NGS in our cohort revealed a novel *GATA4* gene variant in heterozygosis in patient GN0171. *In silico* programs predicted the gene variant c.677C>T; p.Pro226Leu as likely pathogenic and comparison of the amino acid among different species indicated that the region is highly conserved. This suggested the importance of the residue in the protein. Then, we decided to study the impact of the Pro226Leu variant in the transactivation capacity of GATA4 *in vitro*, and compared to the Gly221Arg mutant.

10.5.1.1. Transactivation capacity of GATA4 in JEG3 cells

Previously, co-transfection of the GATA4 expression construct and the *AMH* and *SRY*-reporter genes promoters had been carried out in different cell models (25) but only JEG3 cells transfected with CYP17 promoter revealed consistent results. Functional activity was tested by studying transactivation of wt and mutant GATA4 on the CYP17A1 promoter in JEG3 cell systems (Figure 35).

Transfection of human placental JEG3 cells with wt and GATA4 mutant together with the CYP17A1 promoter reporter construct indicated that the Pro226Leu variant activated the CYP17 promoter similar to wt, while Gly221Arg mutant lost transcriptional activity.

An original research article was derived from these results (Supplementary data 16).

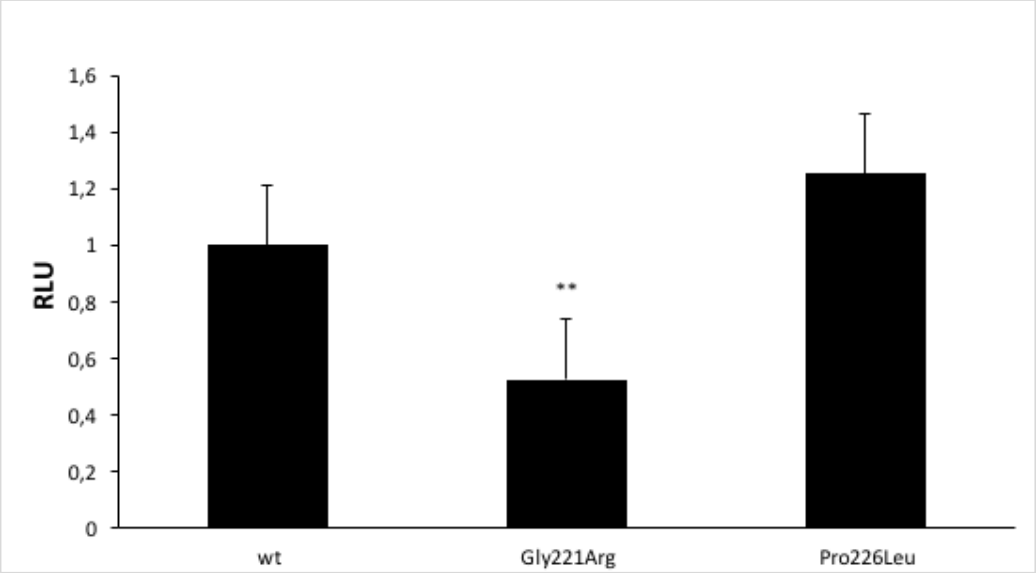

Figure 35. Transcriptional activity of *GATA4* variants on the CYP17A1 promoter. JEG3 cells were transfected with a CYP17 promoter luciferase reporter construct, and the activity of wild type (wt) and mutant GATA4 to transactivate the promoter was tested using the Promega Dual Luciferase readout system. Results are shown as the mean ± SEM of four independent experiments, all performed in duplicate. **p-Value ≤ 0.01. RLU, relative light units; wt, wild type.

10.5.2. Functional characterization of the novel *NR5A1* variants

Many patients with *NR5A1* variants have been studied widely, although no clear genotype-phenotype correlation has been found already. We describe 3 novel *NR5A1* variants (c.71A>T; p.His24Leu, c.88T>A; p.Cys30Ser and c.902G>A; p.Cys301Tyr) found in three 46,XY DSD patients. Comparison of the residues against multiple species indicated its high conservation and *in silico* programs predicted them as likely pathogenic.

10.5.2.1. Promoter transactivation studies

To study the impact of the three changes to the transactivation faculty of SF1 *in vitro*, HEK293 cells were cotransfected with wt or mutant NR5A1 expression construct and a responsive CYP17A1, HSD17B3 and CYP11A1 reporter genes.

As shown in Figure 36, all NR5A1 mutants led to a significant reduced transactivation of the CYP17A1 promoter. These results were confirmed for the His24Leu and Cys30Ser mutants when transactivating the constructs with the CYP11A1 and HSD17B3 promoter stimulation similar to the CYP17A1 promoter. However, for Cys301Tyr mutant no change was observed with HSD17B3 and CYP11A1 reporter genes.

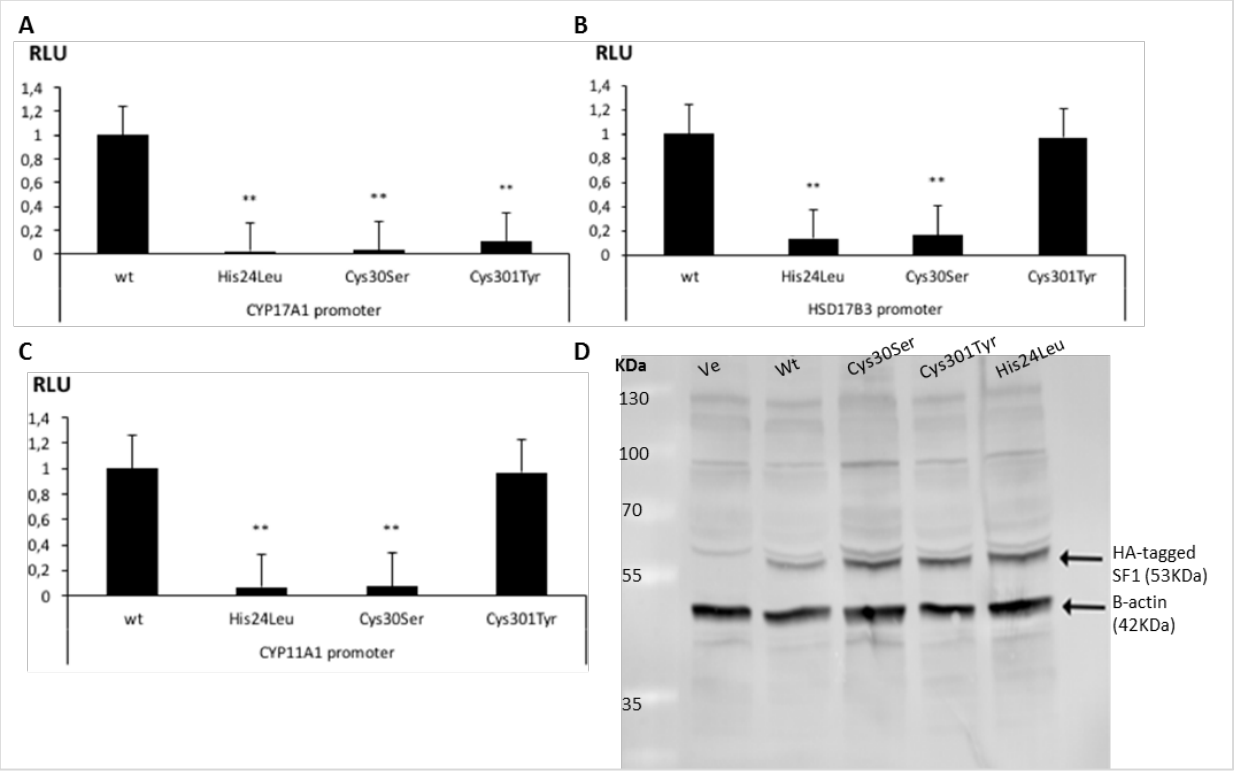

Figure 36. Functional studies of the 3 novel NR5A1 variants. The ability of wt and mutant NR5A1 to activate promoter luciferase reporter constructs was tested in HEK293 cells (A-C). Cells were transiently transfected with NR5A1 expression vectors and promoter reporter contructs CYP17A1 (A), HSD17B3 (B) and CYP11A1 (C). Results are shown as the mean±SEM of five independent experiments, all performed in duplicate. D, Western blot showing expression of wt and mutant SF1 proteins. HA antibody recognized Hemagglutinin-tagged SF1 in the western blot (band at 53 KDa). B-actin was used as a control (band at 42 KDa). **p-Value ≤ 0.01. Ha, hemagglutinin; RLU, relative light units; Ve, empty vector; wt, wild type.

10.5.2.2. Protein expression of *NR5A1* variants

Expression of *NR5A1* variants in HEK293 cells was assessed by a western blot. Cells were cotransfected with the NR5A1 expression vectors containing a HA-tag, then the western blot was performed using an antibody against HA. As shown in Figure 32, we found no difference for SF1 protein expression for the tested variants compared to wt.

10.5.3. Signal transduction of the *LHCGR* gene variants

We analysed *in vitro* two possible *LHCGR* variants found in patients GN0147 and GN0171. The novel missense c.757T>C variant, replacing a serine with proline at position p.253 was found in homozygous state in patient GN0147, while the c.1160C>T;p.Arg554Ter gene variant appeared in heterozygosis in GN0171 case. Although this last variant had been reported previously (269), functional impact was not assessed.

HEK293 cells were transfected with either wt or mutant pSG5 plasmids and a luciferase reporter gene to evaluate the functional impact. As the LHCGR couples to G_s alpha subunit to mediate its functions, we assayed the measure of the second messenger cAMP. We used a reporter gene system of the cAMP response element fused to firefly luciferase (CRE-luc) and cells were then analysed for their responsiveness to different hCG doses (0ng/ml to 100,000ng/ml). The generation of cAMP directly correlates with the activation of the CRE-luc reporter system.

We found that cells transfected with HA-tag and expression vector showed no cAMP generation in response to the increasing tested concentrations of hCG, unlike the cells expressing the wt *LHCGR* receptor (Figure 37).

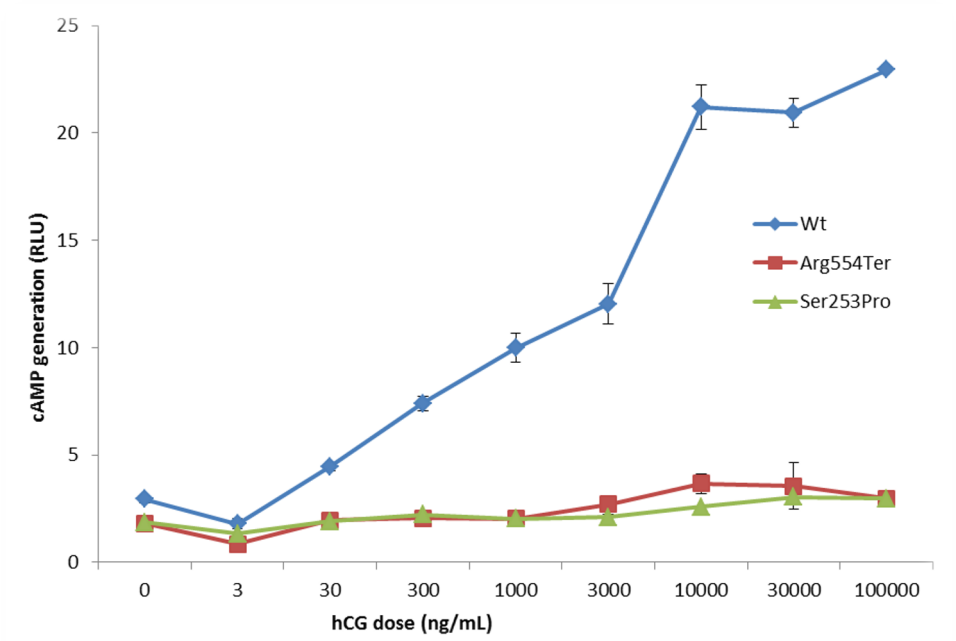

Figure 37. cAMP generation by the wt and the two distinct LHCGR mutants. HEK293 cells were transfected with a pSG5 plasmid carrying either a HA-tagged wt or a HA-tagged mutant (Arg554Ter and Ser253Pro). Cells were stimulated with hCG at different doses and the generation of cAMP was tested using the Promega Dual Luciferase readout system. Each data point represents the mean±SEM of three independent experiments, carried out in duplicates. cAMP, cyclic adenosine monophosphate; hCG, human chorionic gonadotrophin; RLU, relative light units; wt, wild type.

DISCUSSION

In 2006, at the Chicago consensus conference a new term to include the confusing intersex, sex reversal, hermaphroditism and pseudohermafroditism words was created. Such term, DSD, encompasses a range of conditions and avoids mistakes with terms like gender dysphoria or transgender. When genetically caused, reaching a molecular diagnosis supports the management of the patient concerning to gender development, gonadal and adrenal function or other long-term consequences. Since the development of High-throughput screening (HTS) technologies, the understanding of the genetic basis of rare human disorders, including DSD, has increased.

The main objective of this work was the clinical characterization and the molecular diagnosis of an extent cohort of patients with a disorder of sex development either by a candidate gene approach or using next generation sequencing technologies. A targeted gene panel was designed to include genes associated with gonadal development and differentiation, as well as hypogonadism. Hereafter I discuss about the capacity of the customized gene panel as a disgnostic tool. Then, I analyse the phenotype-genotype correlations and the results of the *in silico* and *in vitro* functional studies of the sequence variants found in *NR5A1, LHCGR* and *GATA4.*

DSDs represent a major paediatric concern and a difficult clinical management that might be associated with infertility or cancer. The correct diagnosis of DSD is challenging and may be requested due to a family history of DSD or a phenonotypic presentation suggesting the disorder. Clinical care requires a multidisciplinary team that includes endocrinology, either paediatric or adult since diagnosis ranges from prenatal to adolescent or adults, urology or gynaecology, radiology, pathology in some cases and genetics (174). An important part of the evaluation of the DSD patients is the bioquemical analysis.

Although steroid hormone testing has been used in the past as the first approach to guide the diagnosis, currently is combined with molecular genetic analyses. Hormonal analyses in blood or urine provide differences in the phenotypic expression and offer a fast differential diagnostic orientation. For example, in congenital adrenal hyperplasia it has shown to have a good phenotype-genotype correlation as the quantification of numerous steroid detects rare forms of steroid synthesis disorders. The bioquemical analysis is important for the initial diagnosis of DSD but is also critical to ensure the greatest development at puberty and at hormone replacement treatments (272). It is important to note that in DSDs the homone data interpretation needs to considerate the clinical presentation of the patients and has to be age- and sex- related. In healthy infants, a transient postnatal activation of the hypothalamic-pituitary gonadal axis occurs. This mini-puberty period represents an important diagnostic tool in clinical diagnosis of infants with a suspected disorder of sex development when concentrations of gonadotropins and hormones are measured.

Between the first week and the three months of life, gonadotropins rise and show an obvious sexual dimorphism. In males, predominant levels of LH elevate serum concentrations of testosterone, AMH and inhibin B, while in females FSH gives rise to oestradiol. The following months, concentrations decrease to prepubertal levels. Up to date, little evidence existed on the referenced ranges at this mini-puberty period which difficulted the evaluation of a child with DSD. Now, these ranges have been established (273).

The molecular diagnosis provides guidance for clinical management and further risk, however achieving a genetic diagnosis is difficult due to the wide range of phenotypes that are encompassed in this group of anomalies, the poor genotype-phenotype correlation and the limited knowledge of the pathogenesis that lead to DSD. Currently, 50% of 46,XY DSD patients do not receive a genetic diagnosis (12, 226), although other factors such as environment or epigenetic changes need to be consider also (140). For the elucidation of the genetic causes in DSD, the multiciscilinary team and current technologies should determine firstly the chromosomal sex of the individual. Depending on this, the case will be assigned to one of the three major DSD subclasses established in the Chicago Consensus (174) and together with clinical description, molecular diagnosis will be requested by using the most adequeate molecular testing.

However it is ofter challenging to appropriately classify a DSD patient due to the difficulty to make a correct diagnosis. In this study, this was emphasized by the fact that as a multicentre study, clinical and hormonal data that we obtained from the included patients was highly variable or insufficient. In these cases, the molecular analysis prior to the implementation of NGS led to a poor specific gene study. On the other side, the analysis of these patients, classified as of unknown origin with the targeted gene panel pointed to a more clear DSD classification. In the group of unknown 46,XY DSD, 4 patients were reassigned to disorders of androgen synthesis or action after the identification of an *AR* and *LHCGR* gene variant and to disorders of testicular development due to alterations in *NR5A1*. Nevertheless, a scarce clinical description of the DSD patient drives to a more complicated management of the variants of unknown significance. An individual in the unknown 46,XX DSD cohort presented a sequence variant in *ESR2* but was not further classified because prepubertal ovaries and not dysgenetic had not been reported in the patient. Others have illustrated the requirement of a more detailed biochemical and clinical data for a better understanding of the DSDs (274). To avoid these differences in the clinical descriptions of the patients presenting a DSD, standardized guidelines could be used, such as the External Genitalia Score (275) or recommendations already published as Consensus Statements by COST Action BM1303 working groups. Other papers have

illustrated the requirement of a more detailed biochemical and clinical data for a better understanding of the DSDs (274).

11. EVALUATION OF THE TARGETED GENE SEQUENCING PANEL

We evaluated the performance of the customized panel as a diagnostic gene sequencing tool. We observed that 94.2% of the reads were aligned to the reference genome with a mean depth of 288.3 reads on targeted regions. Although the expected coverage for the targeted regions was 98.55% according to the Ion Ampliseq Designer tool, we observed that the coverage depth was variable among the different genomic regions. For the 48 DSD-related genes we found that 9.3% of the amplicons included in the panel were not covered in most of the samples or were covered by less than 20 reads, an acceptable level for diagnostic use. An example of the latter is the complex sequencing of the *CYP21A2* gene due to the high sequence homology with the *CYP21A1P* pseudogene.

Coverage and targeting efficiency is positively comparable with other studies, using the same and other competing platform (12, 218) Current sequencing platforms, such as Ion Torrent or Illumina use short reads (<400pb) for mapping to the reference genome, which reduces costs and fastens the sequencing but fails to uniquely target the interesting region, Further analysis was done to study those regions without mapped or low amplified regions that were located within exons or exon boundaries.

We also included previously sequenced samples to validate the assay quality. Panel sequencing did not detect a single one-nucleotide deletion in the first exon of the *WT1* gene, located within an amplicon that failed to be read in all the samples. Therefore, sensitivity was set at 81.8%. Although others have included known point variants to ascertain the performance of the panel, the type of variant is not reported and comparison couldn't be made.

CNV were not completely analysed in this study. The customized gene panel failed to detect those rearrangements in patients included as positive controls, therefore few copy number gains or losses were broaden validated. Up to date, the detection of CNV is challenging when using gene panels and although a workflow can be standardized for its use with the Ion Torrent technology, a major expertise is required. For example, it is essential to include control samples without known CNV, as rearrangements will be reported based on their copy number gain or loss relative to the control samples. Four samples from patients with known point mutations in which commercial MLPA had ruled out a CNV were used as controls. However, the presence of other duplications or deletions couldn't be discarded in these samples. A more

exhaustive genetic panel including CNV may improve from 9.1% to 29.8% the disease aetiology, as based on previous reports (28, 222, 276, 277). A specific defect of the Ion Torrent sequencing platform is given by the errors in the number of bases called when a nucleotide or a group of nucleotides are repeated more than once in a sequence (homopolymer errors) (278). Subsequently, alignment is misguided and false small deletions and insertions are called, increasing the false positives and hiding real indels.

Even though the limitations related to the current next generation sequencing technologies or to the Ion Torrent platforms exclusively, we have shown that targeted gene panel sequencing has a great potential as a diagnostic tool to detect point mutations.

12. CLINICAL AND MOLECULAR CHARACTERISTICS OF THE DSD COHORT

For this study we have analysed DNA from 125 patients and 37 family members from DSD patients. Targeted gene panel sequencing together with single gene testing provided a genetic diagnosis in 42.4% of these individuals.

Among the studied patients, 36 were genetically identified by a candidate gene approach and 89 underwent diagnostic testing for DSD using a 48-gene NGS panel. This panel was used for all the samples since 2015, unless a specific gene sequencing or aCGH analysis was requested based on the clinical features of the patient. The diagnostic yield of this targeted gene panel was shown to be 17/89 (19.1%) for all DSDs. Our detection rate is lower to those in previous studies (12, 81, 93, 217, 219, 220, 279) but similar to that obtained by Fan Y *et al* (20.7%) (218). The molecular analysis in our cohort showed a total of 47 pathogenic and 7 likely pathogenic variants in known DSD genes. Among them, 14 were previously unreported, in addition to the 12 VUS variants that were also identified.

As in other publications, a separate analysis of the 46,XY DSD from the whole cohort results in an improved yield to 15/68 (22.05%). It is important to note that prior to the design and implementation of the customized gene panel, a big proportion of our DSD cohort had undergone genetic pre-screening (64.8%) in our laboratory, which definitely affects our overall diagnostic rate. We suppose that if we had applied the gene panel as the first diagnostic approach, we could expect our panel to provide a greater diagnostic percentage. On the other side, reaching a diagnosis essentially depends on the genes included in the panel. Our panel was designed in 2015 and the 48 genes were selected due to published evidence of their role in DSD. Evidence has shown that a higher number of genes do not correlate with a better diagnostic rate. For example, Hughes *et al* demonstrated a molecular diagnosis of 32.1% for their DSD

population by analysing 30 genes (219) and Kim *et al* reached a 29.5% including 67 genes (93). Comparison of the genes in our panel against published genes in other customized panels showed a percentage of common genes from 17.8% (217), 60.9% (12) to 86.6% (219). This may highlight the importance of the selected patients to be analysed through the panel, rather than the genes included in the panel, as the major influence in the diagnostic rate.

Sanger sequencing or testing with a MLPA or aCGH was done depending on the clinician's judgement and following clinical examination. Disease-causative variants were detected in 36 out of the 81 DSD (44.4%) patients studied for a single test. The detection yield of single gene test is dependent on the first selected gene, limited to the phenotype of the patient and therefore to the clinical information given by the clinician. Those patients presenting a DSD with a clear phenotype-genotype correlation were successful after a single gene was sequenced, such as *SRY* in 46,XX testicular DSD or *AR* in AIS suspicion. Patients in which 3 or more genes (or MLPA) were studied had been mainly classified as 46,XY DSD with complete or partial gonadal dysgenesis, which highlights the heterogeneous molecular mechanisms underlying 46,XY DSD.

We conclude that our gene panel is not useful for the molecular diagnosis of 46,XX DSD, at least until the detection of CNVs is upgraded. Among previous studies including 46,XX DSD in their cohorts, massive parallel sequencing led to a diagnostic rate in these patients up to 25% (12, 93, 217-219).

This is one of the largest cohorts of DSD patients analysed for a molecular diagnosis worldwide. To the best of our knowledge this is the first study in which patients with 45X0/46,XY karyotypes and mixed gonadal dysgenesis are included in a targeted gene panel to search for additional variations that explain the phenotype. Altogether this study demonstrated that 42.4% of DSD patients harboured pathogenic or likely pathogenic variants and positively increased the diagnostic yield reached by Fan Y *et al* who analysed 32 patients using a gene sequencing approach and a customized Ion Ampliseq panel (218). Currently, we are working to improve the detection of CNV with the gene panel. An increase in the number of control samples will optimize the workflow and expand the number of patients with a molecular diagnosis. The analysis of both point and big rearrangement in a single run by NGS would improve the use of the technique in a cost-effective mode. Otherwise, a range of methods have been developed to detect these type of genetic variants, such as MLPA or aCGH, and will be performed in those patients with negative genetic results in the near future.

In our series, patients with 46,XY DSD (99, 79.2%) were more common than 46,XX DSD (24, 19.2%), which is consistent with previous studies. The highest

diagnostic rate was observed in patients with 46,XY DSD (46.5%), specifically in disorders of androgen synthesis and action (70.3%), in which described *AR* changes were found in several individuals with identical phenotypes (254-257, 260-262, 280).

As in other populations (140), defects in AR were the most common cause of 46,XY DSD in our cohort. We observed three novel *AR* variants (His100ProfsTer3, Pro276HisfsTer20, Leu881Arg). The two frameshift variants are likely to have no variable expressivity, and then result in a CAIS phenotype. In contrast, Leu881Arg, located in the ligand-binding domain of the receptor is expected to impair androgen binding capacity. In the same codon, another changes support this hypothesis (281, 282) and together with the results shown by *in silico* analyses, it is believed to be the cause of the phenotype in our patient. Moreover, Gln825Lys variant, reported in males with gynecomastia and infertility (259) was identified in a case with different clinical features. However, previously done ligand-binding experiments showed a slight impairment of the mutated AR protein function (258), which could also explain the phenotype in our patient.

Sequence changes in enzymes involved in steroidogenesis were also detected. We identified two novel VUS in *STAR* gene in two cases with discarded adrenal hyperplasia. Up to date, inactivating *STAR* variants led to lipoic congenital adrenal hyperplasia, resulting in adrenal insufficiency and 46,XY DSD, an autosomal recessive disease. Although a non-severe phenotype is expected from missense changes located outside exon 5 to 7 (283), we could hypothesize that heterozygous gene variants, in addition to a second hit, might exclusively affect foetal testicular synthesis without driving to an adrenal steroid deficiency. Indeed, patient GN0028 harboured a heterozygous variant in *NR5A1* together with an unknown significance change in *STAR* gene.

Further reported mutations were detected in *SRD5A2*, *HSD17B3* and *CYP17A1* genes. A number of studies have associated the Gln126Arg change in *SRD5A2* to severe phenotype of the enzyme deficiency in different individuals (264, 284-287), even in compound heterozygosity in four Brazilian patients (288). The deletion of exon 1 in GN0046 was also found in an 46,XY athlete with blind vagina and lack of Müllerian ducts (265). Last variants in *SRD5A2* gene (Tyr91Asp) and in *HSD17B3* (Pro282Leu) were detected in homozygous state and patients were referred with identical phenotypes to those described before (130, 289). Finally, Arg416Cys in *CYP17A1* was identified in a patient with 46,XY karyotype, primary amenorrhea, infantilism and hypertension. Her older sister, with similar clinical features and 46,XX karyotype also presented the mutation. Functional comparison of the wild type and transfected cells of the mutant had shown weak activity of the enzyme (271) in a female with pubertal

delay and hypertension. Additionally, both sisters were diagnosed of pheochromocytoma in their early adulthood.

Nearly 20% of the 46,XY DSD with gonadal dysgenesis are thought to have an inactivating variants in the *SRY* gene (177). However, we only found two patients with hemizygous variants in *SRY,* which represents the 5.7% of the initially suspected cases with 46,XY gonadal dysgenesis. Together with the previously described Gln97Ter variants in the HMG domain of *SRY* in two siblings (238), we also observed the novel and deleterious Pro131Ser variant in a female with complete gonadal dysgenesis. Interestingly, a Proline to Arginine residue change had been described in the same position (239). On the other side, cases of sporadic and familial 46,XY DSD have been related to deleterious variants in *MAP3K1*, a key component of the testis-specific development (81, 82, 290). Eggers *et al* identified 11 patients with 6 different *MAP3K1* variants and extended the phenotype to include hypospadias and undervirilization, to the previous individuals presenting partial and complete gonadal dysgenesis (12). Here, we report one 46,XY gonadal dysgenesis case with a previously reported variant in an individual with similar clinical features, such as female external genitalia, both Müllerian and Wolffian remnants and elevated gonadotropins (82).

WT1 is another early developmental gene identified in 46,XY DSD patients with and without additional syndromes. Historically thought to cause Denys-Drash or Frasier Syndromes, gene alterations in *WT1* encompass a range of clinical phenotypes that may correlate or not with the type of variants and its location (291). Indeed, recent publications have reported *WT1* variants in infants with undervirilized phenotypes, such as hypospadias (38, 291, 292), ambiguous genitalia (219) with and without associated nephropathy, and ovarian failure (19).

In this study we found *WT1* sequence variants in 5 patients with different phenotypes, including two novel missense changes in isolated forms of 46,XY DSDs. The novel Glu75Lys was found in a 3 year old female with immature testes at image and no signs of kidney abnormality. However, both gonadoblastoma and/or kidney disease may be developed later in life (291), therefore patient should be closely followed-up. Although the mother of the patient was a carrier of the same heterozygous variants, incomplete penetrance associated to Denys-Drash syndrome has been previously described (293). The second novel change (Met182Lys) was found in patient GN0156, an infant with precocious puberty. This suggests a greater phenotypic variability than previously thought, however further *in vitro* studies are needed.

Regarding the two syndromes associated with genital malformations and kidney disease, we detected two splice-site variants in intron 9 in two patients with

Frasier syndrome which had been previously described (17, 251). Denys-Drash and Frasier syndromes are characterized by complete or partial gonadal dysgenesis in XY individuals and nephropathy. The difference among these pathologies is the presence of Wilms' tumour in Denys-Drash and gonadoblastoma or streak gonads in Frasier Syndrome. However, since a complete *WT1* gene deletion was found in a newborn with gonadal dysgenesis, gonadoblastoma and bilateral Wilms' tumour (253), the variable phenotype associated to the gene was amplified and enhanced the earlier suggestion that both syndromes are part of the same clinical phenotype (294). In our cohort we found a partial gene deletion, from exon 7 to 10, in a male with bilateral Wilms' tumour and cryptorchidism as the only genital malformation, although hypoplastic testes were visually noted. Nephropathy was not described, but proteinuria was present before surgery. Hence, we support the idea of Frasier and Denys-Drash syndromes as part of a wider phenotypic picture, instead of isolated syndromes, in which milder phenotypes are included, such as less severe kidney anomalies or its absence.

Seven individuals in our cohort with X-linked Adrenal hypoplasia congenita (AHC) were found to have hemizygous missense, nonsense and frameshift changes or complete gene deletions in *NR0B1*. Clinical features of the disease include salt-losing adrenal insufficiency, hypogonadotropic hypogonadism and infertility in both sexes, while boys may also refer early puberty and glucocorticoid or mineralocorticoid deficiency. In our study patients principally manifested typical symptoms of salt-wasting crisis involving dehydration, low sodium levels and hypoglycaemia either with or without skin pigmentation. Abnormal sexual development was only described in POL0274, while two other cases (GN0101 and POL0285) referred hypogonadotropic hypogonadism, when aged 17 and 45 years, respectively. DAX1 and SF1 are the major factors involved in adrenal and reproductive development and function. Most of the different mutations in *NR0B1* are associated with the loss-of-function in the ligand-binding domain or a stop variant that translates to a shorter protein (295).

In the present study we identified two novel variants located in the DNA-binding domain that are susceptible to form shorter protein structures and explain the primary adrenal insufficiency (PAI) observed in these patients. In the newborn males, the condition presents with clinical features that are vaguely different from the salt-wasting form of 21-hydroxylase deficiency. An accurate hormonal and molecular diagnosis has significant consequences for precise prescription of the treatment and genetic screening affects other family members, indeed the risk of symptoms related to adrenal insufficiency in brothers and males in the maternal family need to be consider (295). In our study, patient GN0078 presented a positive family history of unexpected early death in males, and thus had an effect on the early diagnosis of the individual and genetic counselling in an affected cousin. Case POL0285 also reported

family background but was genetically diagnosed in his adulthood when he presented with hypogonadotropic hypogonadism. Similar to GN0101, a complete *NR0B1* gene deletion was found in this patient. Guo *et al* firstly described in 1995 a complete gene deletion in two isolated cases presenting with PAI and mental retardation (249), since then many others have been reported encompassing similar phenotypes to that of our patients (296).

In human, *WWOX* gene germline mutations have been shown to be associated with sex differentiation, but also with spinocerebellar ataxia and WOREE syndrome, characterized by epileptic encephalopathy (WWOX-related epileptic encephalopathy) (297). We identified two missense variants in two patients referred to clinician for 46,XY gonadal dysgenesis suspicion. Gly62 residue, in exon 3, situated in the second WW domain of the protein, is involved in protein interactions and therefore the observed change in patient GN0203 may impair this function. In contrast, Pro366 amino acid is located in the carboxi-terminal of the encoded protein. Although a milder effect would be suspected, the phenotype observed in patient GN0020 is more severe than in subject GN0203. Further molecular analyses are needed to assess the implications of *WWOX* in 46,XY gonadal dysgenesis.

More recently, heterozygous and homozygous missense changes in *ZFPM2* have been noticed in patients with partial and complete gonadal dysgenesis (67). We identified one variant (Ala1026Val) in heterozygosis in GN0155 patient, who had ambiguous genitalia at birth and surgery for a dysgenetic gonad at age 2 years. More analyses are required to test this unreported variants. Other studies have identified 9p deletions, including *DMRT1* and *DMRT2* genes in phenotypic females with Y-chromosome showing gonadal dysgenesis or ambiguous genitalia (75). To the best of our knowledge this is the first *DMRT2* missense variants found in a 46,XY DSD, however its deleterious effect would be probed when the cousin of this patient, also presenting complete gonadal dysgenesis, is analysed.

Regarding the 46,XX DSD cohort, we were able to reach a genetic diagnosis in 7 out of 24 cases (29.2%). The majority of the cases were reported to have a testicular DSD and translocation of the *SRY* gene had been studied prior to gene panel sequencing. We only detected one more patient carrying *SRY* by NGS.

Testicular DSD is caused by the translocation of the *SRY* gene in the 90% of the cases. This type of rearrangements are rare, but about 50 cases have been described and happen throughout the aberrant exchange between homologous sequences (PAR1) on Xp and Yq during meiosis in the spermatocyte. The phenotype of males with Xp:Yq translocations depends on the extend of the Xp deletion. They can present short stature, Leri Weill dyschondrosteosis, chondrodysplasia punctate, mental retardation,

ichthyosis and hypogonadotropic hypogonadism in combination with anosmia (Kallman syndrome) when the deletion is large.

Otherwise, copy number gains and deletions, such as activating variants in *SOX3* driving to testis development or loss-of-function variants in *RSPO1*, an ovary-specific signalling gene (98) (192), contribute to the pathogenesis of the disorders of ovarian development. However, the use of CGHa does not increase the molecular diagnosis of these patients. Kim GJ *et al* identified *SOX9* upstream duplications in three out of nineteen 46,XX DSD individuals (15.8%) (298). Indeed, two ovotesticular patients (GN0090 and GN0158) had undergone for Copy Number Variations (CNV) detection with a commercial MLPA but rearrangements were not found. The analysis of CNV in four patients with SRY-negative 46,XX testicular DSD revealed loss of heterozygosity in 13 regions not known to be associated with testicular development (299). Similar analysis in 46,XX DSD will emphasize the role of CNV in the pathogenesis of the disease. However, disorders other than congenital adrenal hyperplasia (CAH), constitute small groups among 46,XX DSD and obtaining a significant number of patients is challenging.

46,XX DSDs account for approximately the 25% of all DSDs, with CAH due to 21-hydroxylase deficiency being the most usual cause (95%) (300). It is diagnosed during routine newborn screening by detecting baseline and stimulated 17-hydroxy progesterone values at birth. Such screening programs prevent the incorrect gender assignment in females with severe virilization and identify male infants with the salt-wasting type of classic CAH and normal external genitalia to prevent adrenal crisis. However, individuals with CAH were not included in this study due to the difficulty to correctly sequence the gene, as other studies did (12, 218). Patients would have been examined separately for CYP21A1 defects if CAH was suspected. Recent recommendations from the DSDnet state that all neonates should have a complete steroid analysis to prevent emergencies due to severe adrenal insufficiency (174).

Still, relatively few genes have been implicated in the development of ovaries compared to testicular formation, and this unknowledge may be the reason for the 46,XX DSD cases that remain unsolved. Downstream targets of WNT4 and RSPO1, as well as their coactivators seem to be good candidate genes. *LGR4* (Leucine-rich repeat-containing G protein-coupled receptor) for example, codifies for a receptor of R-spondin and mediates WNT4/β-catenin signalling while *Lgr4*-null mice models have shown female-to-male sex reversal (301).

On the other side, traditional and whole exome sequencing have recently identified rare missense heterozygous changes in *ESR2* in two young 46,XX girls with amenorrhea, hypoplastic uterus and non-visible gonads or multicystic ovaries (161).

Although *Esr2* deficient murine models resulted in different phenotypes that suggested its role in fertility in both sexes but not in sexual differentiation (160, 302), loss-of-function mutations have been described in 46,XY DSD (160, 303). Our targeted-gene panel identified a missense change in a highly conserved residue located in the fifth coding exon of *ESR2* in a teenage 46,XX girl. The 14-year old patient presented with puberty delay and primary amenorrhea, infantile uterus and prepubertal ovaries. Estradiol levels were low at <5pg/ml. No other relevant changes were shown in genes linked to ESR signalling or gonadal development. ESR2 is expressed in the foetal ovary and in the Granulosa cells of the small and preovulatory follicles and as reported, negative effect on the receptor led to complete ovarian insufficiency (160). Due to the location of the 221 residue on the DNA-binding domain (DBD) (160), we could hypothesize that those previously reported variants located in the ligand-binding domain (LBD) of the ESR-β cause a more severe phenotype such as absence of ovaries and bilateral cystic lesions. In our case, patient had prepubertal ovaries and although hormone replacement therapy had been planned, proper gonadal development needs to be followed-up. This is the third ESR2 change implicated in the pathogenesis of 46,XX DSD, which further agrees with its involvement on the female gonadal progress.

Regarding *ESR1*, loss-of-function gene variants cause oestrogen resistance, a syndrome characterized by primary amenorrhea with no breast development, polycystic ovaries and infantile uterus in women and delayed bone age, osteoporosis, continuous growth and elevated oestrogens levels in both sexes (161). To date, few patients with *ESR1* mutations have been reported in DSD patients, 1 in males and 4 in females (38, 304). Our 46,XX patient presented with a different phenotype compared to the Mayer-Rokitansky-Küster-Hauser type 1 cases described by Brucker *et al* in 2017 (304). Female KO mice have no pubertal mammary gland development and hypoplastic uterus, then the clinical phenotype of patient GN0198 recapitulates the murine model (162). This young female presented with non-visible ovaries at the referral time but later cystic development is not discarded after treatment has started. On the other side, this patient harboured another variant in *HSD17B4* gene, which function is unknown in isolated DSD, but could modulate the role of ESR1.

Finally, a rare variant in *MAMLD1* was found in a 46,XX DSD. The importance of MAMLD1 in sexual development is unknown. Although it is expressed in testes and ovaries during foetal life and in testicular and adrenal tissue in adults (87, 91), the poor genotype-phenotype correlation led to the search of further variants that explained the clinical broad range. A recent WES analysis in our patient revealed 5 genes related to female gonadal development and 46,XX DSD, such as *MAML3* (91). The identification of multiple variants in a patient suggests the partial contribution of each one and complicates the functional testing of their effect on the pathogenesis. Interactions between genes and proteins involved in sex development were studied

using bioinformatic tools and recent literature in the field, emphasizing the network analysis as a tool to understand complex genetic data. This study underlined that the different clinical features observed in DSD cases might happen due to a group of genetic variants that contribute to sexual development.

In the present study no sequence variants were detected in 72 out of 125 patients. Besides the limitation of targeted gene panel sequencing to detect CNVs, the analysis of deep intronic sequences is also missing with this technology. Genetic studies in DSD have been focused towards the coding regions of the genome, however it is thought that defects in the transcriptional regulation of either well-known genes or yet to be discovered genes resulting from non-coding regions explain a part of the DSD unsolved cases.

In fact, several non-coding defects affecting regulatory elements and the expression of its target genes have been implicated in DSD (194). For example, interesting regulatory mechanisms have been added to the pathogenesis of AIS and have highlighted the significance of non-coding regions in the development of AIS. An identified variant in the 5'UTR of the *AR* gene created an open reading frame (ORF) that modified the transcript and reduced the translation levels of the downstream gene. The expression of the short protein reduced the AR levels, while the mRNA levels remained unchanged (141). Although an intronic variant had been previously identified (140), later in 2016 a deep intronic variant (c.2450-118A>G) creating a *de novo* 5' splice site and two aberrant transcripts was identified by WGS. It was shown that despite total mRNA levels were normal, the stability of the protein was compromised and the AR was undetectable (305). In our cohort, 3 patients with a clinical suspicion of AIS did not reveal a coding *AR* mutation. Our customized panel was designed to span only 25pb of the exon-intron boundaries. Although enough to detect disease-causing intronic variants in patient with a Frasier syndrome suspicion, most diagnostic approaches using gene panels, as well as whole exome sequencing leave a majority of cases unsolved. As we don't use an alternative method to study introns we might miss diagnostic variants that fall within these regions. The analyses of patients with such disorders using genome sequencing will further increase the importance of intronic regions in the pathogenesis of DSD.

As stated before, the number of genes in the panel does not necessarily correlate with a better diagnostic rate. However, the limited number of genes (48 genes) strongly contrast with the 62 and 61 genes known to be involved in 46,XY DSD and 46,XX DSD, respectively (174). Although new primers can be added to the customized panel, the rapid advance in knowledge about the molecular pathogenesis of DSD leads to the discovery of candidate genes from time to time, increasing the cost

of the updating. Sequencing the complete exome appears to be the best cost-eficiency method then.

Investigation by WES or WGS sequencing allows a targeted approach first and the later extension to candidate genes. Indeed, the use of WES has allowed the discovery of genes previously not associated to DSD in humans, such as *ZFPM2* and *SOX8* (53, 67). Nevertheless, the further candidate gene detection needs critical evaluation. The addition of genes without a clear phenotype correlation may only result in an increase of VUS and accidental findings, without a significant improve of the diagnostic rate. Certainly, a challenge associated with the implementation of High-throughput screening (HTS) is the detection of VUS. These may be in genes known to be related to the phenotype but without sufficient evidence to confirm the pathogenicity, affecting sequence changes in candidate genes or in well-known genes not associated to the field of study (306). In our cohort, segregation studies were not performed in some patients with VUS, complicating further categorization of the variant and molecular diagnosis. Although previous analysis of singletons and trios in a large cohort of DSD cases through NGS suggested that the sequencing of the family members is not essential, more single patients had gene variants classified as VUS compared to trios (12).

No sequence variants were found in the two patients with a mixed gonadal dysgenesis and a 45,X0/46,XY karyotype. MGD has been associated with Yp chromosome microdeletions, detected with either sequence tagged site based PCRs or cytogenetic techniques, depending on the size of the missing segment. Besides the inconveniency of detecting microdeletions in the *SRY* gene with the gene panel, one out of 3 individuals with a 45,X0/46,XY karyotype and an apparently normal Y chromosome may have microdeletions (307). Moreover, in 6 out of 15 patients with a 45,X0/46,XY karyotype or its variants and a MGD phenotype microdeletions were found inside the AZF region in Yq11.

The importance of a good clinical description or diagnosis prior to the molecular testing is being questioned. DSD present with variable clinical appearance and hormonal criteria, interfering with the direct causative gene prediction for a concrete phenotype. NGS replaced traditional sequencing allowing the analysis of more than one gene at the time and nowadays the technical cost of targeted panel sequencing is equal to the multiple-exon sequencing of a gene. Traditionally, certain phenotypes were assumed to reliably predict a disruptive change in a DSD-related gene, such as sequence variants in *AR*. However, as mentioned before variants in non-coding or 5'UTR of the *AR* or other genes regulating the AR pathway are now highly suspected to cause AIS (141). In our experience, we have found DSD patients are either misdiagnosed or lack complete clinical information. These findings underline the use of

our customized panel as a fast and less costly method to reach a molecular diagnosis, together with a CNV detection method, when patients are affected with a DSD of unknown basis. Additionally, highlights the importance of the molecular analysis for clinical management. Recently COST Action authors remembered that genetic diagnosis cannot always predict the functional consequences for a particular DSD patient, which are important aspects for individualized management (174). For instance, the range of phenotypes encompassing *LHCGR* or *NR5A1* gene variants in our cohort drives to different managing of the patient.

Finally, not all DSD cases are explained by a genetic variant. The role of environment is evident in the development of genitalia as well as the epigenetic changes that disrupt gene expression in DSDs arising in foetal stage (140). On one hand, oestrogenic and anti-androgenic complexes are responsible of the defects of the urethral closure in humans and the elevated exposure to these compounds explains the increasing incidence of hypospadias in developed countries. Indeed, only 30% of hypospadias cases have a clear genetic cause (183). Exposure to infectious diseases, such as rubella has also sustained the development of ovarian dysgenesis in some individuals (197). Finally, the impact of histone modifications observed in gonad development and the possibility to produce DSD is still challenging. Up to date, the understanding of epigenetic modifications activating or inhibiting gene expression came from *in vitro* analyses (308), although its role in specific processes such as gonadal development keeps being challenging. Recent tools, such as the ChIP-seq (Chromatin Immunoprecipitation Sequencing) allow the investigation of the binding pattern of chromatin-associated proteins in the genome and have led to the hypothesis that a mixture of posttranscriptional modifications, including methylation, acetylation, phosphorylation and ubiquitination are required to obtain an effect on gene expression (216). These advances may provide insight into the development of DSD.

New screening technologies have demonstrated that the phenotype of some individuals with DSD results from the oligogenic inheritance of gene variations. Recent publications suggested that the range of DSD phenotypes associated with *NR5A1* may occur due to the inheritance of heterozygous *NR5A1* gene variants and further hits in other genes involved in the sex development pathway or other cofactors (224). Since Mazen *et al* reported a heterozygous *NR5A1* change in combination with a *MAP3K1* gene variants in a patient with 46,XY gonadal dysgenesis (309), other studies have reinforced the complicated phenotype–genotype correlation observed in this patients (310). Indeed, we found oligogenic inheritance of a heterozygous variant in *NR5A1* in combination with a VUS in testis differentiation genes *AMH* and *STAR* in two of our patients. Both genes are upregulated by the transcription factor in the early human gonad and are involved in the virilization of external genitalia (311).

Oligogenic inheritance has been discovered in other DSDs. Eggers *et al* found segregating variants in a testis development gene and a VUS in a CHH (congenital hypogonadotrophic hypogonadism) gene in 3 patients with severe hypospadias (12). More recently, the contribution of multiple genetic variants to *MAMLD1*-associated 46,XY and 46,XX DSD phenotypes has been published (91). Similarly we found in a 46,XY DSD patient with a mild phenotype, a heterozygote variant in *LHCGR* together with a *GATA4* gene change. Although the *GATA4* variants did not affect the CYP17 promoter activity, it has been suggested that the combination of individually non-deleterious variants contributes to the DSD phenotype (224). Moreover, several GATA4 sites are present in the 5'UTR of *LHCGR,* which indicates that GATA4 could regulate this gene. We propose that *GATA4* may also show oligogenic involvement in DSD. Finally, we found in a 46,XX DSD patient a heterozygous *ESR1* variants in combination with an additional variant in gonadal differentiation gene *HSD17B4*. Although few variants have been identified in *ESR1*, an autosomal recessive inheritance mode is suspected, and then we propose that defects in both genes cause the phenotype of this patient.

13. ANALYSIS OF RARE VARIANTS IN *GATA4, NR5A1* AND *LHCGR* AND FUNCTIONAL STUDIES

The molecular testing of our DSD cohort showed also a VUS in *GATA4*, as well as 6 changes in the *NR5A1* gene and 6 in *LHCGR*. In those novel variants, further studies were performed to test the transcriptional activity of the variants.

13.1. VARIANTS IN *GATA4* AND *IN VITRO* STUDIES

In 2018, we published the characterization of p.Pro226Leu *GATA4* sequence variant together with another two variants (p.Cys238Arg and pTrp228Cys) found by other research groups (25). The three missense changes were found in heterozygosis in three individuals with a 46,XY DSD phenotype with and without congenital heart disease (CHD). *GATA4* mutations associated with cardiac defects are distributed throughout the gene and its contribution to the pathogenesis of the disease was assessed when mutants decrease the ability to transactivate target genes or fail to interact with proteins involved in heart development (312). Moreover, it is known that common variants in a region of GATA4 3'UTR change posttranscriptional gene regulation at miRNA level (313) contributing to the risk of CHD.

In contrast, functional characterization of *GATA4* variants with respect to the 46,XY DSD phenotype has only been performed for the p.Gly221Arg mutation so far (27). *In vitro* studies revealed that p.Gly221Arg lacked DNA binding, had impaired transactivation activity on the AMH promoter, and failed to bind cofactor FOG2. Functional testing of three *GATA4* variants identified in 46,XY DSD individuals of our study showed no effect on transactivation activity on the CYP17 promoter for *GATA4* variants p.Pro226Leu and p.Trp228Cys but similar disruptive effect for the missense mutation p.Cys238Arg.

The few GATA4 missense mutations found in 46,XY DSD individuals with or without CHD are all located in the N-terminal zinc finger domain, responsible for DNA binding and interaction with cofactors (314). Although the three variants are located in the N-terminal zinc finger domain of GATA4, only Cys238 is close to Zn binding sites. The Gly221 also close to such sites, is not directly involved in Zn binding but is situated next to Cys220 which binds the Zn atom, and thus, the mutation Gly221Arg disrupts the Zn binding, leading to a non-functional GATA4. The Cys238 binds Zn and its change to arginine leads to loss of Zn binding.

GATA4 regulates the expression of multiple genes coding for hormones or components of the steroidogenic pathway during testis development and function, such as *NR5A1*. GATA4 interacts with NR5A1 in Sertoli cell cultures to regulate the expression of AMH, and therefore, it has been reported that mutations in *NR5A1* cause 46,XY DSD due to lack of interaction with GATA (23). In families with *GATA4* mutations and isolated CHD no gonadal involvement is mostly detected, possibly because some of the variants retain some DNA-binding activity and exhibit different degrees of transcriptional activation on gonadal promoters and thus, remain able to synergize with NR5A1. The p.Cys238Arg mutation was found in a patient with a complex CHD, genital ambiguity, and persistent Müllerian ducts, which led to female gender assignment. We proposed that p.Cys238Arg mutation in *GATA4* lacks activity to bind DNA reducing the transactivation of AMH critically.

On the contrary, variants p.Pro226Leu and p.Trp228Cys were found in individuals with less severe 46,XY DSD phenotype. They were raised as males, and had no evidence of heart anomalies. Indeed, Harrison *et al* (26) screened patients with 46,XY DSD using an aCGH and found an infant presenting a deletion of 0.22 Mb upstream GATA4. Same deletion was found in his healthy mother and his maternal grandmother, who had CHD. The authors proposed that the identified deletion do not affect the coding sequence of *GATA4*, and may, therefore, not manifest with CHD, but rather disrupt regulatory elements controlling gene expression essential in the developing gonad (28). Thus, the phenotypic variability could be explained by genetic modifiers (315) (67). In addition, GATA4 regulates multiple promoters to variable

degrees. Structural changes caused by both the p.Pro226Leu and pTrp228Cys variations were not predicted to be disruptive and GATA4 structure was not altered. However, the changes were in the DNA interaction sites and it is expected that both p.Pro226Leu and p.Trp228Cys variants could have altered binding and activation of some of GATA4 interaction partners and could also bind to other promoters and potentially change the transcription of several other genes. In fact, NGS revealed segregating genetic variants in two other genes in these two patients with a 46,XY DSD phenotype without CHD. In the GN0171 patient a heterozygote mutation in LHCGR gene was found together with the GATA4 variant. Looking at the 5'UTR of LHCGR, several GATA sites are present suggesting that GATA4 may regulate this gene. Therefore, combined heterozygosity for GATA4 and LHCGR variants in our patient may explain the 46,XY DSD phenotype. The other case harboured an additional heterozygote variant in *LRP4* gene, related to the Cenani–Lenz syndactyly syndrome (OMIM 212780) and disruption of canonical WNT/beta-catenin signalling, important in sexual development (25).

As in other DSD-related genes, phenotypical variability with same heterozygous GATA4 mutation (p.Gly221Arg and p.Trp228Cys) observed within same family manifesting with either 46,XY DSD or CHD only, indicates that there might be incomplete penetrance or variable sensitivity (27) (25). Interestingly, the same observation was made in mice heterozygote for a Gata4 deletion (316). Oligogenic mechanisms may also be considered.

13.2. SEQUENCE VARIANTS IN *NR5A1* GENE AND FURTHER ANALYSES

We found six variants in the *NR5A1* gene, all were detected in heterozygosis and were mainly single-nucleotide variants. A broad range of phenotypes, from hypospadias to severe forms of 46,XY DSD and primary ovarian insufficiency (POI) in 46,XX DSD have been widely described in individuals harbouring heterozygous *NR5A1* variants (14). Clinical picture of these 46,XY patients varied from ambiguous genitalia with micropenis and cryptorchidism to female external genitalia with Müllerian remnants. Hormonal levels were also variable and although adrenal insufficiency has been associated with *NR5A1* mutations, it was not suspected in our patients.

The first three mutations (His24Leu, Cys30Ser and Arg84Cys) are located in the DNA- binding domain (DBD) region of the protein. While His24 and Cys30 amino acids are involved in the binding of the major groove of DNA to promoters, Arg84 is placed in the A box which stabilizes the binding of the protein to DNA (14). The latter, was found in a patient presenting with ambiguous genitalia at birth and has been observed in two 46,XY girls with clitoral hypertrophy so far (240). Functional characterization of

the mutation showed decreased DNA binding and reduced transactivation activity (240). Variants His24Leu and Cys30Ser in *NR5A1* gene are described here for the first time. However, other mutations in the same codons (His24Tyr and Cys30Trp) were reported in Spanish and Turkish girls who presented with bilateral inguinal hernia and female or ambiguous genitalia, respectively (14). We found that these SF1 mutations in the DNA binding domain (DBD) showed impaired transactivation activity when studied in the three different promoters. Gln206ThrfsTer20, located in the hinge region will result in a truncated protein without ligand binding domain (LBD). The same variant in a Spanish girl with similar phenotype had been reported (14). We found two more novel gene variants (Glu304CysfsTer26 and Cys301Tyr) located in the LBD region, which modulates the activity of the protein through the binding of cofactors to the AF2 domain in the C-terminal. Several studies have shown that variants in the DBD may correlate with higher protein impact compared to changes in the LBD region, where effects vary depending on their location and the recognition of the ligand (14, 310). Similar results were observed in our cohort, since the *in vitro* promoter transactivation assay showed a more variable result when analysing Cys301Tyr. However, in this study, we found that patients GN0109 and GN0111 presented with a more severe phenotype compared to those with a sequence variant in the DBD region, which highlights the poor correlation between functional testing and clinical features. In the case of Glu304CysfsTer26 mutation, a premature stop was created and the truncated protein lacked the AF2 region, leading to a loss of protein function.

A clear genotype-phenotype correlation was not seen in patients bearing *NR5A1* mutations and certainly, family members harbouring the same gene variant may present with variable phenotypes (14). In one case (GN0075), family history showed premature menopause or menstrual anomalies and severe hypospadias in some relatives having heterozygous *NR5A1* variants. By contrast, clinically normal carriers of the Cys301Tyr and Glu304CysfsTer26 variants were also observed, as similarly reported before (14). Additional mutations in *NR5A1* or in other genes implied in the sexual development have been proposed as a hypothesis to explain this phenotypic variability. In 2016, Mazen and colleagues found a known Arg313Cys *NR5A1* mutation in compound heterozygous state with a *MAP3K1* variant, suggesting that a digenic inheritance model could explain the phenotypic heterogeneity observed in 46,XY cases with gonadal dysgenesis (309). Since then, other NGS analyses of DSD patients have reported a second rare variant in 46,XY patients (224) (310). Similarly, we analysed the data for additional variants and found the Arg121Trp *STAR* change in patient GN0028 and Thr143Ile variant in *AMH* gene in GN0042. SF1 transcription factor regulates the synthesis of StAR in Leydig cells, causing virilisation of external genitalia as well as testicular descent, while in Sertoli cells it upregulates *AMH*, leading to the regression of Müllerian structures (317). Both patients presented with clinical features explained by these second variants, however the *STAR* variants was also found in the

patient's healthy father. An extensive study of the contribution of these second variants found in this study would provide further insight into the variable genotype-phenotype correlation found in DSD patients harboring a heterozygous *NR5A1* gene variant.

The contribution of the Gly196Ala polymorphism has also been suggested to explain the phenotypic variability, as some of the described mutations in *NR5A1* are in compound heterozygous state with this SNP (318). We did not detect the Gly196Ala change in none of the NR5A1 positive patients, however we found it in four patients presenting another deleterious variants in *AR* and *LHCGR*, which may suggests its benign role in the pathogenesis of DSD. Although functional analyses of the polymorphism demonstrated its null ability to activate transcription (240) we cannot discard an adding effect to the pathogenesis of the disease.

13.3. VARIANTS IN *LHCGR* GENE AND *IN VITRO* STUDIES

Clinical characteristics of patients harbouring sequence variants in *LHCGR* depend on the activating or inactivating effect driven by the change. Based on the inactivation degree of the receptor, consequences of the inactivating variants range from a 46,XY DSD with female phenotype (Leydig Cell Hypoplasia, LCH type 1) to an undervirilized male with hypospadias and micropenis (LCH type 2) (319) or primary amenorrhea in females. Although LCH is thought to be a recessive inheritance disease, we identified the Arg554Ter change in *LHCGR* gene in a patient with incomplete masculinization of external genitalia. This nonsense variant is located in the third intracellular loop of the LH receptor and Latronico *et al* also found it in three affected siblings with female external genitalia (269). Subsequently, we tested the ability of Wt LHCGR and Arg554Ter mutant to mediate cAMP production in respond to increasing concentrations of hCG and found no second messenger generation when cells expressed the mutant LHCGR. The premature interruption of the translation process removes a large part of the receptor and thus, the hormonal signal is not transduced properly. Because heterozygous parents of other described patients were phenotypically normal (269) we could conclude that a non-functioning allele causes no DSD in either sex, but its combined heterozygosity with the GATA4 change might explain the phenotype of this concrete case (75). Moreover, other inactivating missense variants and deletions have been identified throughout the whole gene, which cause the mutant LHCGR unresponsive to hormones and result in female external phenotype, as in two of our cases. Patients GN0034 and GN0147, with typical clinical and hormonal phenotype of LCH, presented a complete gene deletion and a novel missense variant in the extracellular domain of the protein, respectively. This large extracellular N-terminal domain is responsible of the high-affinity binding of the

hormone on the surface formed by leucine-rich repeats and the beginning of signal transduction (320). Functional analyses of the mutations located in this region have resulted in receptor intracellular retention with little surface expression and slight or no hormone stimulation measure (321). Accordingly, transfected cells with the Ser253Pro mutant significantly reduced cAMP accumulation after hCG stimulation and elucidates the severe Leydig cell hypoplasia observed in this patient. Patient GN0034 presented a large deletion in chromosome 2 affecting exons 1 to 10 of the gene and therefore inducing the total loss-of-function of the gene. Indeed, the deletion removed *GTF2A1L* and *STON-GTF2A1L* genes, encoding a germ cell specific subunit of the TFIIA transcription factor involved in the interaction between TATA-binding protein and DNA and essential in testis biology and spermatogenesis (322). The lack of the entire LHCGR gene had been previously reported by Richard *et al* (270) in a newborn with sexual ambiguity described as labial fusion, a genital tubercle with hypospadias and inguinal gonads. We therefore expand the phenotype of patients presenting with a whole gene deletion.

On the other side, male-limited precocious puberty is an autosomal dominant disease characterized by the LH-independent activation that increases testosterone synthesis. Heterozygous activating mutations in the *LHCGR* change the normal configuration of the receptor and activate G protein-coupled receptor (267). Patients commonly present early signs of puberty by the age of 4 years and elevated serum testosterone levels despite low LH (323). In early reported patients, either in familial or sporadic cases, mutations were located in exon 11, encoding for the signal transduction region, but now they are sited in other exons. For instance, we identified the novel and probably deleterious Gln190Lys variants in exon 7 in GN0157 case, affected with pubic hair development, accelerated growth and a slight testicular enlargement without penile growth. Adrenal tumors and aromatase deficiency were excluded. Testosterone level was within normal range at referral and LH was undetectable. We can speculate that the different clinical aspects are related to the location of the c.568C>A (p.Gln190Lys) variant in the *LHC*GR gene.

Previous studies had compared clinical and hormonal data of unrelated patients with gonadotropin-independent precocious puberty harboring the same variant in one and two alleles and no differences were observed (324). However functional studies would clarify if the variant is likely to cause the precocious puberty in this boy. Genetic analysis in the parents confirmed that the father is a carrier, however no clinical data was available to verify the pathogenicity of the alteration. As known, females with activating LHCGR mutations have no manifestations (325). In patients with precocious puberty, sporadic or with family background, genetic analyses as well as clinical and hormonal investigations are valuable for diagnosis. In GN0068 the Met571Ile mutation confirmed the disease in the patient aged 3 years and his

father. In contrast, GN0088 was a sporadic case and the Met398Thr mutation appeared *de novo.* Both mutations had been reported in similar phenotypes by Kremer *et al* in 1993 (267) and Kraaij in 1995 (268) in exon 11 of *LHCGR* gene. This two activating mutations, placed in the third intracellular loop and involved in G-protein interaction, change the sixth and second transmembrane segments of the protein. *In vitro* experiments showed LH- or hCG-independent increased production of cAMP levels in HEK293 cells and elevated production of basal steroids in tumor Leydig cells of a murine model (268).

In conclusion, we have reported the clinical and molecular characteristics of 125 patients with a DSD diagnose. We have extended the number of gene variants identified to date in sex development and contributed to the idea of either incomplete penetrance or oligogenic mechanisms that explain the phenotypical variability observed in these individuals. However, further research is needed and the application of other molecular approaches, such as exome and/or genome sequencing would reveal additional alterations in these DSD patients.

CONCLUSIONS

1. The molecular analysis of DSD-related genes identifies a causative genetic variant in the 40.2% of the patients analysed in this cohort.

2. The molecular analysis using a targeted gene panel is useful for the detection of single-nucleotide variants, mostly in 46,XY DSD patients. Indeed we made a molecular diagnosis in 22% of 46,XY DSD compared to the 10.5% of the 46,XX DSD individuals. However, further improvements are necessary to detect copy number variations with this technology.

3. We observed that disorders of sex development are highly heterogeneous. This, difficults the clinical diagnosis and highlights the need of a genetic study to validate the initial suspicion. Still, little correlation has been found between the phenotype of an individual and the genetic variant that is identified.

4. Clinical and biochemical data guide the genetic analysis when using a single-gene sequencing approach. Next generation sequencing provides a molecular diagnosis when there is incomplete analitical information on the patients.

5. Incomplete penetrance or oligogenic mechanisms might explain the phenotypical variability observed in patients with identical gene variants or within same family.

6. Next generation sequencing techniques have expanded the number of genes and pathways that are involved in sex determination and gonadal development. We propose the use of a targeted gene panel as the first layer of the molecular diagnosis and the sequencing of the whole exome (WES) or whole genome (WGS) for additional analysis. The study of chromosomal rearrangements and intronic regions is also necessary.

7. Functional analysis are necessary to validate the pathogenicity of the novel genetic changes that are identified either through the gene-by- gene approach or next generation sequencing. Accordingly, we found that only variants located in the DNA binding domain of the SF1 protein significantly reduced transactivation activity of the promoters, as well as the two gene alterations in *LHCGR*. However, in patients harbouring heterozygote *NR5A1* and *LHCGR* variants, we cannot confirm its causative effect in the pathogenesis of the disease.

REFERENCES

1. Karl J, Capel B. Sertoli cells of the mouse testis originate from the coelomic epithelium. Dev Biol. 1998;203(2):323-33.
2. Ohnesorg T, Vilain E, Sinclair AH. The genetics of disorders of sex development in humans. Sex Dev. 2014;8(5):262-72.
3. Yang Y, Workman S, Wilson M. The molecular pathways underlying early gonadal development. J Mol Endocrinol. 2018.
4. Munger SC, Natarajan A, Looger LL, Ohler U, Capel B. Fine time course expression analysis identifies cascades of activation and repression and maps a putative regulator of mammalian sex determination. PLoS Genet. 2013;9(7):e1003630.
5. Mamsen LS, Ernst EH, Borup R, Larsen A, Olesen RH, Ernst E, et al. Temporal expression pattern of genes during the period of sex differentiation in human embryonic gonads. Sci Rep. 2017;7(1):15961.
6. Bösze P, Szabó D, László J, Gaál M. Ultrastructure of the fibrous tissue of the streak gonads. Acta Med Acad Sci Hung. 1982;39(3-4):133-5.
7. Ahmed SF, Bashamboo A, Lucas-Herald A, McElreavey K. Understanding the genetic aetiology in patients with XY DSD. Br Med Bull. 2013;106:67-89.
8. Baetens D, Verdin H, De Baere E, Cools M. Update on the genetics of differences of sex development (DSD). Best Pract Res Clin Endocrinol Metab. 2019.
9. Tantawy S, Mazen I, Soliman H, Anwar G, Atef A, El-Gammal M, et al. Analysis of the gene coding for steroidogenic factor 1 (SF1, NR5A1) in a cohort of 50 Egyptian patients with 46,XY disorders of sex development. Eur J Endocrinol. 2014;170(5):759-67.
10. Biason-Lauber A. Control of sex development. Best Pract Res Clin Endocrinol Metab. 2010;24(2):163-86.
11. Achermann JC, Ito M, Hindmarsh PC, Jameson JL. A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. Nat Genet. 1999;22(2):125-6.
12. Eggers S, Sadedin S, van den Bergen JA, Robevska G, Ohnesorg T, Hewitt J, et al. Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. Genome Biol. 2016;17(1):243.
13. Knarston IM, Robevska G, van den Bergen JA, Eggers S, Croft B, Yates J, et al. NR5A1 gene variants repress the ovarian-specific WNT signaling pathway in 46,XX disorders of sex development patients. Hum Mutat. 2019;40(2):207-16.
14. Camats N, Pandey AV, Fernández-Cancio M, Andaluz P, Janner M, Torán N, et al. Ten novel mutations in the NR5A1 gene cause disordered sex development in 46,XY and ovarian insufficiency in 46,XX individuals. J Clin Endocrinol Metab. 2012;97(7):E1294-306.
15. Pérez de Nanclares G, Castaño L, Bilbao JR, Vallo A, Rica I, Vela A, et al. Molecular analysis of Frasier syndrome: mutation in the WT1 gene in a girl with gonadal dysgenesis and nephronophthisis. J Pediatr Endocrinol Metab. 2002;15(7):1047-50.
16. Hatano O, Takakusu A, Nomura M, Morohashi K. Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. Genes Cells. 1996;1(7):663-71.
17. Barbaux S, Niaudet P, Gubler MC, Grünfeld JP, Jaubert F, Kuttenn F, et al. Donor splice-site mutations in WT1 are responsible for Frasier syndrome. Nat Genet. 1997;17(4):467-70.
18. Gomes NL, de Paula LCP, Silva JM, Silva TE, Lerário AM, Nishi MY, et al. A 46,XX testicular disorder of sex development caused by a Wilms' tumour Factor-1 (WT1) pathogenic variant. Clin Genet. 2019;95(1):172-6.
19. Wang H, Li G, Zhang J, Gao F, Li W, Qin Y, et al. Novel WT1 Missense Mutations in Han Chinese Women with Premature Ovarian Failure. Sci Rep. 2015;5:13983.
20. Viger RS, Mertineit C, Trasler JM, Nemer M. Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Müllerian inhibiting substance promoter. Development. 1998;125(14):2665-75.

21. Morrisey EE, Ip HS, Tang Z, Parmacek MS. GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily. J Biol Chem. 1997;272(13):8515-24.
22. Eggers S, Sinclair A. Mammalian sex determination—insights from humans and mice. Chromosome Res. 2012;20(1):215-38.
23. Viger RS, Guittot SM, Anttonen M, Wilson DB, Heikinheimo M. Role of the GATA family of transcription factors in endocrine development, function, and disease. Mol Endocrinol. 2008;22(4):781-98.
24. Flück CE, Miller WL. GATA-4 and GATA-6 modulate tissue-specific transcription of the human gene for P450c17 by direct interaction with Sp1. Mol Endocrinol. 2004;18(5):1144-57.
25. Martinez de LaPiscina I, de Mingo C, Riedl S, Rodriguez A, Pandey AV, Fernández-Cancio M, et al. GATA4 Variants in Individuals With a 46,XY Disorder of Sex Development (DSD) May or May Not Be Associated With Cardiac Defects Depending on Second Hits in Other DSD Genes. Front Endocrinol (Lausanne). 2018;9:142.
26. Harrison SM, Granberg CF, Keays M, Hill M, Grimsby GM, Baker LA. DNA copy number variations in patients with 46,XY disorders of sex development. J Urol. 2014;192(6):1801-6.
27. Lourenço D, Brauner R, Rybczynska M, Nihoul-Fékété C, McElreavey K, Bashamboo A. Loss-of-function mutation in GATA4 causes anomalies of human testicular development. Proc Natl Acad Sci U S A. 2011;108(4):1597-602.
28. White S, Ohnesorg T, Notini A, Roeszler K, Hewitt J, Daggag H, et al. Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. PLoS One. 2011;6(3):e17793.
29. Igarashi M, Mizuno K, Kon M, Narumi S, Kojima Y, Hayashi Y, et al. mutations are uncommon in patients with 46,XY disorders of sex development without heart anomaly. Asian J Androl. 2018;20(6):629-31.
30. Katoh-Fukui Y, Miyabayashi K, Komatsu T, Owaki A, Baba T, Shima Y, et al. Cbx2, a polycomb group gene, is required for Sry gene expression in mice. Endocrinology. 2012;153(2):913-24.
31. Norling A, Hirschberg AL, Iwarsson E, Wedell A, Barbaro M. CBX2 gene analysis in patients with 46,XY and 46,XX gonadal disorders of sex development. Fertil Steril. 2013;99(3):819-26.e3.
32. Biason-Lauber A, Konrad D, Meyer M, DeBeaufort C, Schoenle EJ. Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. Am J Hum Genet. 2009;84(5):658-63.
33. Ma W, Li Y, Wang M, Li H, Su T, Wang S. Associations of Polymorphisms in WNT9B and PBX1 with Mayer-Rokitansky-Küster-Hauser Syndrome in Chinese Han. PLoS One. 2015;10(6):e0130202.
34. Eozenou C, Bashamboo A, Bignon-Topalovic J, Merel T, Zwermann O, Lourenco D, et al. The TALE homeodomain of PBX1 is involved in human primary testis-determination. Hum Mutat. 2019.
35. Piard J, Mignot B, Arbez-Gindre F, Aubert D, Morel Y, Roze V, et al. Severe sex differentiation disorder in a boy with a 3.8 Mb 10q25.3-q26.12 microdeletion encompassing EMX2. Am J Med Genet A. 2014;164A(10):2618-22.
36. Birk OS, Casiano DE, Wassif CA, Cogliati T, Zhao L, Zhao Y, et al. The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature. 2000;403(6772):909-13.
37. Wilhelm D, Englert C. The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. Genes Dev. 2002;16(14):1839-51.
38. Wang H, Zhang L, Wang N, Zhu H, Han B, Sun F, et al. Next-generation sequencing reveals genetic landscape in 46, XY disorders of sexual development patients with variable phenotypes. Hum Genet. 2018;137(3):265-77.
39. Lucas-Herald AK, Bashamboo A. Gonadal development. Endocr Dev. 2014;27:1-16.

40. Croft B, Ayers K, Sinclair A, Ohnesorg T. Review disorders of sex development: The evolving role of genomics in diagnosis and gene discovery. Birth Defects Res C Embryo Today. 2016;108(4):337-50.
41. Biason-Lauber A. WNT4, RSPO1, and FOXL2 in sex development. Semin Reprod Med. 2012;30(5):387-95.
42. Eggers S, Ohnesorg T, Sinclair A. Genetic regulation of mammalian gonad development. Nat Rev Endocrinol. 2014;10(11):673-83.
43. Bashamboo A, McElreavey K. Gene mutations associated with anomalies of human gonad formation. Sex Dev. 2013;7(1-3):126-46.
44. Benko S, Gordon CT, Mallet D, Sreenivasan R, Thauvin-Robinet C, Brendehaug A, et al. Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. J Med Genet. 2011;48(12):825-30.
45. Huang B, Wang S, Ning Y, Lamb AN, Bartley J. Autosomal XX sex reversal caused by duplication of SOX9. Am J Med Genet. 1999;87(4):349-53.
46. Bishop CE, Whitworth DJ, Qin Y, Agoulnik AI, Agoulnik IU, Harrison WR, et al. A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. Nat Genet. 2000;26(4):490-4.
47. Sekido R, Lovell-Badge R. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature. 2008;453(7197):930-4.
48. Kim Y, Kobayashi A, Sekido R, DiNapoli L, Brennan J, Chaboissier MC, et al. Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. PLoS Biol. 2006;4(6):e187.
49. Wilhelm D, Hiramatsu R, Mizusaki H, Widjaja L, Combes AN, Kanai Y, et al. SOX9 regulates prostaglandin D synthase gene transcription in vivo to ensure testis development. J Biol Chem. 2007;282(14):10553-60.
50. Jameson SA, Lin YT, Capel B. Testis development requires the repression of Wnt4 by Fgf signaling. Dev Biol. 2012;370(1):24-32.
51. Barrionuevo F, Bagheri-Fam S, Klattig J, Kist R, Taketo MM, Englert C, et al. Homozygous inactivation of Sox9 causes complete XY sex reversal in mice. Biol Reprod. 2006;74(1):195-201.
52. O'Bryan MK, Takada S, Kennedy CL, Scott G, Harada S, Ray MK, et al. Sox8 is a critical regulator of adult Sertoli cell function and male fertility. Dev Biol. 2008;316(2):359-70.
53. Portnoi MF, Dumargne MC, Rojo S, Witchel SF, Duncan AJ, Eozenou C, et al. Mutations involving the SRY-related gene SOX8 are associated with a spectrum of human reproductive anomalies. Hum Mol Genet. 2018;27(7):1228-40.
54. Weiss J, Meeks JJ, Hurley L, Raverot G, Frassetto A, Jameson JL. Sox3 is required for gonadal function, but not sex determination, in males and females. Mol Cell Biol. 2003;23(22):8084-91.
55. Grinspon RP, Nevado J, Mori Alvarez MeL, Del Rey G, Castera R, Venara M, et al. 46,XX ovotesticular DSD associated with a SOX3 gene duplication in a SRY-negative boy. Clin Endocrinol (Oxf). 2016;85(4):673-5.
56. Schmahl J, Kim Y, Colvin JS, Ornitz DM, Capel B. Fgf9 induces proliferation and nuclear localization of FGFR2 in Sertoli precursors during male sex determination. Development. 2004;131(15):3627-36.
57. Bagheri-Fam S, Ono M, Li L, Zhao L, Ryan J, Lai R, et al. FGFR2 mutation in 46,XY sex reversal with craniosynostosis. Hum Mol Genet. 2015;24(23):6699-710.
58. Chiang HS, Wu YN, Wu CC, Hwang JL. Cytogenic and molecular analyses of 46,XX male syndrome with clinical comparison to other groups with testicular azoospermia of genetic origin. J Formos Med Assoc. 2013;112(2):72-8.
59. García-Acero M, Molina M, Moreno O, Ramirez A, Forero C, Céspedes C, et al. Gene dosage of DAX-1, determining in sexual differentiation: duplication of DAX-1 in two sisters with gonadal dysgenesis. Mol Biol Rep. 2019;46(3):2971-8.

60. Rojek A, Obara-Moszynska M, Malecka E, Slomko-Jozwiak M, Niedziela M. NR0B1 (DAX1) mutations in patients affected by congenital adrenal hypoplasia with growth hormone deficiency as a new finding. J Appl Genet. 2013;54(2):225-30.
61. Seminara SB, Achermann JC, Genel M, Jameson JL, Crowley WF. X-linked adrenal hypoplasia congenita: a mutation in DAX1 expands the phenotypic spectrum in males and females. J Clin Endocrinol Metab. 1999;84(12):4501-9.
62. Merke DP, Tajima T, Baron J, Cutler GB. Hypogonadotropic hypogonadism in a female caused by an X-linked recessive mutation in the DAX1 gene. N Engl J Med. 1999;340(16):1248-52.
63. Meeks JJ, Weiss J, Jameson JL. Dax1 is required for testis determination. Nat Genet. 2003;34(1):32-3.
64. Bouma GJ, Albrecht KH, Washburn LL, Recknagel AK, Churchill GA, Eicher EM. Gonadal sex reversal in mutant Dax1 XY mice: a failure to upregulate Sox9 in pre-Sertoli cells. Development. 2005;132(13):3045-54.
65. Swain A, Narvaez V, Burgoyne P, Camerino G, Lovell-Badge R. Dax1 antagonizes Sry action in mammalian sex determination. Nature. 1998;391(6669):761-7.
66. Mizusaki H, Kawabe K, Mukai T, Ariyoshi E, Kasahara M, Yoshioka H, et al. Dax-1 (dosage-sensitive sex reversal-adrenal hypoplasia congenita critical region on the X chromosome, gene 1) gene transcription is regulated by wnt4 in the female developing gonad. Mol Endocrinol. 2003;17(4):507-19.
67. Bashamboo A, Brauner R, Bignon-Topalovic J, Lortat-Jacob S, Karageorgou V, Lourenco D, et al. Mutations in the FOG2/ZFPM2 gene are associated with anomalies of human testis determination. Hum Mol Genet. 2014;23(14):3657-65.
68. Clark AM, Garland KK, Russell LD. Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules. Biol Reprod. 2000;63(6):1825-38.
69. Ayers K, van den Bergen J, Robevska G, Listyasari N, Raza J, Atta I, et al. Functional analysis of novel desert hedgehog gene variants improves the clinical interpretation of genomic data and provides a more accurate diagnosis for patients with 46,XY differences of sex development. J Med Genet. 2019;56(7):434-43.
70. Lindeman RE, Gearhart MD, Minkina A, Krentz AD, Bardwell VJ, Zarkower D. Sexual cell-fate reprogramming in the ovary by DMRT1. Curr Biol. 2015;25(6):764-71.
71. Ottolenghi C, McElreavey K. Deletions of 9p and the quest for a conserved mechanism of sex determination. Mol Genet Metab. 2000;71(1-2):397-404.
72. Zhao L, Svingen T, Ng ET, Koopman P. Female-to-male sex reversal in mice caused by transgenic overexpression of Dmrt1. Development. 2015;142(6):1083-8.
73. Matson CK, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D. DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature. 2011;476(7358):101-4.
74. Privitera O, Vessecchia G, Bernasconi B, Bettio D, Stioui S, Giordano G. Prenatal diagnosis of del(9)(p24): a sex reverse case. Prenat Diagn. 2005;25(10):945-8.
75. Ounap K, Uibo O, Zordania R, Kiho L, Ilus T, Oiglane-Shlik E, et al. Three patients with 9p deletions including DMRT1 and DMRT2: a girl with XY complement, bilateral ovotestes, and extreme growth retardation, and two XX females with normal pubertal development. Am J Med Genet A. 2004;130A(4):415-23.
76. Takagi M, Yagi H, Fukuzawa R, Narumi S, Hasegawa T. Syndromic disorder of sex development due to a novel hemizygous mutation in the carboxyl-terminal domain of. Hum Genome Var. 2017;4:17012.
77. Tang P, Argentaro A, Pask AJ, O'Donnell L, Marshall-Graves J, Familari M, et al. Localization of the chromatin remodelling protein, ATRX in the adult testis. J Reprod Dev. 2011;57(3):317-21.

78. Cohen-Haguenauer O, Picard JY, Mattéi MG, Serero S, Nguyen VC, de Tand MF, et al. Mapping of the gene for anti-müllerian hormone to the short arm of human chromosome 19. Cytogenet Cell Genet. 1987;44(1):2-6.
79. Jamil Z, Perveen K, Malik R, Avesi L. Serum anti-mullerian hormone: Correlation with the ovarian follicular dynamics in healthy mice. J Pak Med Assoc. 2016;66(9):1084-8.
80. Bogani D, Siggers P, Brixey R, Warr N, Beddow S, Edwards J, et al. Loss of mitogen-activated protein kinase kinase kinase 4 (MAP3K4) reveals a requirement for MAPK signalling in mouse sex determination. PLoS Biol. 2009;7(9):e1000196.
81. Baxter RM, Arboleda VA, Lee H, Barseghyan H, Adam MP, Fechner PY, et al. Exome sequencing for the diagnosis of 46,XY disorders of sex development. J Clin Endocrinol Metab. 2015;100(2):E333-44.
82. Granados A, Alaniz VI, Mohnach L, Barseghyan H, Vilain E, Ostrer H, et al. MAP3K1-related gonadal dysgenesis: Six new cases and review of the literature. Am J Med Genet C Semin Med Genet. 2017;175(2):253-9.
83. Nef S, Verma-Kurvari S, Merenmies J, Vassalli JD, Efstratiadis A, Accili D, et al. Testis determination requires insulin receptor family function in mice. Nature. 2003;426(6964):291-5.
84. Pitetti JL, Calvel P, Romero Y, Conne B, Truong V, Papaioannou MD, et al. Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. PLoS Genet. 2013;9(1):e1003160.
85. Vinci G, Brauner R, Tar A, Rouba H, Sheth J, Sheth F, et al. Mutations in the TSPYL1 gene associated with 46,XY disorder of sex development and male infertility. Fertil Steril. 2009;92(4):1347-50.
86. Puffenberger EG, Hu-Lince D, Parod JM, Craig DW, Dobrin SE, Conway AR, et al. Mapping of sudden infant death with dysgenesis of the testes syndrome (SIDDT) by a SNP genome scan and identification of TSPYL loss of function. Proc Natl Acad Sci U S A. 2004;101(32):11689-94.
87. Camats N, Fernández-Cancio M, Audí L, Mullis PE, Moreno F, González Casado I, et al. Human MAMLD1 Gene Variations Seem Not Sufficient to Explain a 46,XY DSD Phenotype. PLoS One. 2015;10(11):e0142831.
88. Abstracts of the LWPES/ESPE 8th Joint Meeting Global Care in Paediatric Endocrinology, in collaboration with APEG, APPES, JSPE and SLEP. New York City, New York, USA. September 9-12, 2009. Horm Res. 2009;72 Suppl 3:1-547.
89. Kalfa N, Cassorla F, Audran F, Oulad Abdennabi I, Philibert P, Béroud C, et al. Polymorphisms of MAMLD1 gene in hypospadias. J Pediatr Urol. 2011;7(6):585-91.
90. Fukami M, Wada Y, Okada M, Kato F, Katsumata N, Baba T, et al. Mastermind-like domain-containing 1 (MAMLD1 or CXorf6) transactivates the Hes3 promoter, augments testosterone production, and contains the SF1 target sequence. J Biol Chem. 2008;283(9):5525-32.
91. Flück CE, Audí L, Fernández-Cancio M, Sauter K-S, Martinez de LaPiscina I, Castaño L, et al. Broad Phenotypes of Disorders/Differences of Sex Development in MAMLD1 Patients Through Oligogenic Disease. Frontiers in Genetics. 2019;10(746).
92. White S, Hewitt J, Turbitt E, van der Zwan Y, Hersmus R, Drop S, et al. A multi-exon deletion within WWOX is associated with a 46,XY disorder of sex development. Eur J Hum Genet. 2012;20(3):348-51.
93. Kim JH, Kang E, Heo SH, Kim GH, Jang JH, Cho EH, et al. Diagnostic yield of targeted gene panel sequencing to identify the genetic etiology of disorders of sex development. Mol Cell Endocrinol. 2017;444:19-25.
94. Ottolenghi C, Omari S, Garcia-Ortiz JE, Uda M, Crisponi L, Forabosco A, et al. Foxl2 is required for commitment to ovary differentiation. Hum Mol Genet. 2005;14(14):2053-62.
95. Liu CF, Parker K, Yao HH. WNT4/beta-catenin pathway maintains female germ cell survival by inhibiting activin betaB in the mouse fetal ovary. PLoS One. 2010;5(4):e10382.

96. Tomizuka K, Horikoshi K, Kitada R, Sugawara Y, Iba Y, Kojima A, et al. R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. Hum Mol Genet. 2008;17(9):1278-91.
97. Chassot AA, Gillot I, Chaboissier MC. R-spondin1, WNT4, and the CTNNB1 signaling pathway: strict control over ovarian differentiation. Reproduction. 2014;148(6):R97-110.
98. Tallapaka K, Venugopal V, Dalal A, Aggarwal S. Novel RSPO1 mutation causing 46,XX testicular disorder of sex development with palmoplantar keratoderma: A review of literature and expansion of clinical phenotype. Am J Med Genet A. 2018;176(4):1006-10.
99. Mandel H, Shemer R, Borochowitz ZU, Okopnik M, Knopf C, Indelman M, et al. SERKAL syndrome: an autosomal-recessive disorder caused by a loss-of-function mutation in WNT4. Am J Hum Genet. 2008;82(1):39-47.
100. Naillat F, Yan W, Karjalainen R, Liakhovitskaia A, Samoylenko A, Xu Q, et al. Identification of the genes regulated by Wnt-4, a critical signal for commitment of the ovary. Exp Cell Res. 2015;332(2):163-78.
101. Uhlenhaut NH, Treier M. Forkhead transcription factors in ovarian function. Reproduction. 2011;142(4):489-95.
102. Boulanger L, Pannetier M, Gall L, Allais-Bonnet A, Elzaiat M, Le Bourhis D, et al. FOXL2 is a female sex-determining gene in the goat. Curr Biol. 2014;24(4):404-8.
103. Liu L, Rajareddy S, Reddy P, Du C, Jagarlamudi K, Shen Y, et al. Infertility caused by retardation of follicular development in mice with oocyte-specific expression of Foxo3a. Development. 2007;134(1):199-209.
104. Rajpert-De Meyts E, Jørgensen N, Graem N, Müller J, Cate RL, Skakkebaek NE. Expression of anti-Müllerian hormone during normal and pathological gonadal development: association with differentiation of Sertoli and granulosa cells. J Clin Endocrinol Metab. 1999;84(10):3836-44.
105. Barseghyan H, Délot EC, Vilain E. New technologies to uncover the molecular basis of disorders of sex development. Mol Cell Endocrinol. 2018;468:60-9.
106. Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. Endocr Rev. 2011;32(1):81-151.
107. Stocco DM. StAR protein and the regulation of steroid hormone biosynthesis. Annu Rev Physiol. 2001;63:193-213.
108. Miller WL, Huang N, Pandey AV, Flück CE, Agrawal V. P450 oxidoreductase deficiency: a new disorder of steroidogenesis. Ann N Y Acad Sci. 2005;1061:100-8.
109. Häggström M, Richfield D. Diagram of the pathways of human steroidogenesis. Wikiversity Journal of Medicine; 2014.
110. Mindnich R, Möller G, Adamski J. The role of 17 beta-hydroxysteroid dehydrogenases. Mol Cell Endocrinol. 2004;218(1-2):7-20.
111. Carreau S, Lambard S, Delalande C, Denis-Galeraud I, Bilinska B, Bourguiba S. Aromatase expression and role of estrogens in male gonad : a review. Reprod Biol Endocrinol. 2003;1:35.
112. Fitzpatrick SL, Carlone DL, Robker RL, Richards JS. Expression of aromatase in the ovary: down-regulation of mRNA by the ovulatory luteinizing hormone surge. Steroids. 1997;62(1):197-206.
113. Speiser PW, Arlt W, Auchus RJ, Baskin LS, Conway GS, Merke DP, et al. Congenital Adrenal Hyperplasia Due to Steroid 21-Hydroxylase Deficiency: An Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab. 2018;103(11):4043-88.
114. Flück CE, Meyer-Böni M, Pandey AV, Kempná P, Miller WL, Schoenle EJ, et al. Why boys will be boys: two pathways of fetal testicular androgen biosynthesis are needed for male sexual differentiation. Am J Hum Genet. 2011;89(2):201-18.
115. Fukami M, Homma K, Hasegawa T, Ogata T. Backdoor pathway for dihydrotestosterone biosynthesis: implications for normal and abnormal human sex development. Dev Dyn. 2013;242(4):320-9.

116. Miller WL. Androgen biosynthesis from cholesterol to DHEA. Mol Cell Endocrinol. 2002;198(1-2):7-14.
117. Huang N, Miller WL. Cloning of factors related to HIV-inducible LBP proteins that regulate steroidogenic factor-1-independent human placental transcription of the cholesterol side-chain cleavage enzyme, P450scc. J Biol Chem. 2000;275(4):2852-8.
118. Hiort O, Holterhus PM, Werner R, Marschke C, Hoppe U, Partsch CJ, et al. Homozygous disruption of P450 side-chain cleavage (CYP11A1) is associated with prematurity, complete 46,XY sex reversal, and severe adrenal failure. J Clin Endocrinol Metab. 2005;90(1):538-41.
119. Miller WL. Disorders in the initial steps of steroid hormone synthesis. J Steroid Biochem Mol Biol. 2017;165(Pt A):18-37.
120. Flück CE, Pandey AV. Steroidogenesis of the testis -- new genes and pathways. Ann Endocrinol (Paris). 2014;75(2):40-7.
121. Al Alawi AM, Nordenström A, Falhammar H. Clinical perspectives in congenital adrenal hyperplasia due to 3β-hydroxysteroid dehydrogenase type 2 deficiency. Endocrine. 2019;63(3):407-21.
122. Auchus RJ. Steroid 17-hydroxylase and 17,20-lyase deficiencies, genetic and pharmacologic. J Steroid Biochem Mol Biol. 2017;165(Pt A):71-8.
123. Finkielstain GP, Chen W, Mehta SP, Fujimura FK, Hanna RM, Van Ryzin C, et al. Comprehensive genetic analysis of 182 unrelated families with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. J Clin Endocrinol Metab. 2011;96(1):E161-72.
124. Speiser PW, Azziz R, Baskin LS, Ghizzoni L, Hensle TW, Merke DP, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2010;95(9):4133-60.
125. Kamrath C, Hochberg Z, Hartmann MF, Remer T, Wudy SA. Increased activation of the alternative "backdoor" pathway in patients with 21-hydroxylase deficiency: evidence from urinary steroid hormone analysis. J Clin Endocrinol Metab. 2012;97(3):E367-75.
126. Flück CE, Pandey AV, Huang N, Agrawal V, Miller WL. P450 oxidoreductase deficiency - a new form of congenital adrenal hyperplasia. Endocr Dev. 2008;13:67-81.
127. Krone N, Reisch N, Idkowiak J, Dhir V, Ivison HE, Hughes BA, et al. Genotype-phenotype analysis in congenital adrenal hyperplasia due to P450 oxidoreductase deficiency. J Clin Endocrinol Metab. 2012;97(2):E257-67.
128. Tomalik-Scharte D, Maiter D, Kirchheiner J, Ivison HE, Fuhr U, Arlt W. Impaired hepatic drug and steroid metabolism in congenital adrenal hyperplasia due to P450 oxidoreductase deficiency. Eur J Endocrinol. 2010;163(6):919-24.
129. Mushtaq T, Ahmed SF. The impact of corticosteroids on growth and bone health. Arch Dis Child. 2002;87(2):93-6.
130. Andersson S, Geissler WM, Wu L, Davis DL, Grumbach MM, New MI, et al. Molecular genetics and pathophysiology of 17 beta-hydroxysteroid dehydrogenase 3 deficiency. J Clin Endocrinol Metab. 1996;81(1):130-6.
131. Mendonca BB, Gomes NL, Costa EM, Inacio M, Martin RM, Nishi MY, et al. 46,XY disorder of sex development (DSD) due to 17β-hydroxysteroid dehydrogenase type 3 deficiency. J Steroid Biochem Mol Biol. 2017;165(Pt A):79-85.
132. Unal E, Yıldırım R, Taş FF, Demir V, Onay H, Haspolat YK. Aromatase Deficiency due to a Novel Mutation in. J Clin Res Pediatr Endocrinol. 2018;10(4):377-81.
133. Chávez B, Ramos L, Gómez R, Vilchis F. 46,XY disorder of sexual development resulting from a novel monoallelic mutation (p.Ser31Phe) in the steroid 5α-reductase type-2 (SRD5A2) gene. Mol Genet Genomic Med. 2014;2(4):292-6.
134. Adham IM, Steding G, Thamm T, Büllesbach EE, Schwabe C, Paprotta I, et al. The overexpression of the insl3 in female mice causes descent of the ovaries. Mol Endocrinol. 2002;16(2):244-52.
135. Ayers K, Kumar R, Robevska G, Bruell S, Bell K, Malik MA, et al. Familial bilateral cryptorchidism is caused by recessive variants in. J Med Genet. 2019.

136. El Houate B, Rouba H, Sibai H, Barakat A, Chafik A, Chadli eB, et al. Novel mutations involving the INSL3 gene associated with cryptorchidism. J Urol. 2007;177(5):1947-51.
137. Ferlin A, Zuccarello D, Zuccarello B, Chirico MR, Zanon GF, Foresta C. Genetic alterations associated with cryptorchidism. JAMA. 2008;300(19):2271-6.
138. Nef S, Shipman T, Parada LF. A molecular basis for estrogen-induced cryptorchidism. Dev Biol. 2000;224(2):354-61.
139. Jeyasuria P, Ikeda Y, Jamin SP, Zhao L, De Rooij DG, Themmen AP, et al. Cell-specific knockout of steroidogenic factor 1 reveals its essential roles in gonadal function. Mol Endocrinol. 2004;18(7):1610-9.
140. Audi L, Fernández-Cancio M, Carrascosa A, Andaluz P, Torán N, Piró C, et al. Novel (60%) and recurrent (40%) androgen receptor gene mutations in a series of 59 patients with a 46,XY disorder of sex development. J Clin Endocrinol Metab. 2010;95(4):1876-88.
141. Hornig NC, de Beaufort C, Denzer F, Cools M, Wabitsch M, Ukat M, et al. A Recurrent Germline Mutation in the 5'UTR of the Androgen Receptor Causes Complete Androgen Insensitivity by Activating Aberrant uORF Translation. PLoS One. 2016;11(4):e0154158.
142. Panda B, Rao L, Tosh D, Dixit H, Padmalatha V, Kanakavalli M, et al. Germline study of AR gene of Indian women with ovarian failure. Gynecol Endocrinol. 2011;27(8):572-8.
143. Qiao J, Han B. Diseases caused by mutations in luteinizing hormone/chorionic gonadotropin receptor. Prog Mol Biol Transl Sci. 2019;161:69-89.
144. Xu Y, Chen Y, Li N, Hu X, Li G, Ding Y, et al. Novel compound heterozygous variants in the LHCGR gene identified in a subject with Leydig cell hypoplasia type 1. J Pediatr Endocrinol Metab. 2018;31(2):239-45.
145. Martens JW, Verhoef-Post M, Abelin N, Ezabella M, Toledo SP, Brunner HG, et al. A homozygous mutation in the luteinizing hormone receptor causes partial Leydig cell hypoplasia: correlation between receptor activity and phenotype. Mol Endocrinol. 1998;12(6):775-84.
146. Chen C, Xu X, Kong L, Li P, Zhou F, Zhao S, et al. Novel homozygous nonsense mutations in LHCGR lead to empty follicle syndrome and 46, XY disorder of sex development. Hum Reprod. 2018;33(7):1364-9.
147. Boot AM, Lumbroso S, Verhoef-Post M, Richter-Unruh A, Looijenga LH, Funaro A, et al. Mutation analysis of the LH receptor gene in Leydig cell adenoma and hyperplasia and functional and biochemical studies of activating mutations of the LH receptor gene. J Clin Endocrinol Metab. 2011;96(7):E1197-205.
148. Fanelli F. Theoretical study on mutation-induced activation of the luteinizing hormone receptor. J Mol Biol. 2000;296(5):1333-51.
149. McLaren A. Mammalian germ cells: birth, sex, and immortality. Cell Struct Funct. 2001;26(3):119-22.
150. Jiao X, Ke H, Qin Y, Chen ZJ. Molecular Genetics of Premature Ovarian Insufficiency. Trends Endocrinol Metab. 2018;29(11):795-807.
151. Knarston I, Ayers K, Sinclair A. Molecular mechanisms associated with 46,XX disorders of sex development. Clin Sci (Lond). 2016;130(6):421-32.
152. Liu C, Peng J, Matzuk MM, Yao HH. Lineage specification of ovarian theca cells requires multicellular interactions via oocyte and granulosa cells. Nat Commun. 2015;6:6934.
153. Themmen APN, Huhtaniemi IT. Mutations of gonadotropins and gonadotropin receptors: elucidating the physiology and pathophysiology of pituitary-gonadal function. Endocr Rev. 2000;21(5):551-83.
154. Chang HM, Qiao J, Leung PC. Oocyte-somatic cell interactions in the human ovary-novel role of bone morphogenetic proteins and growth differentiation factors. Hum Reprod Update. 2016;23(1):1-18.
155. Sanfins A, Rodrigues P, Albertini DF. GDF-9 and BMP-15 direct the follicle symphony. J Assist Reprod Genet. 2018;35(10):1741-50.

156. Castronovo C, Rossetti R, Rusconi D, Recalcati MP, Cacciatore C, Beccaria E, et al. Gene dosage as a relevant mechanism contributing to the determination of ovarian function in Turner syndrome. Hum Reprod. 2014;29(2):368-79.
157. Norling A, Hirschberg AL, Rodriguez-Wallberg KA, 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 comparative genomic hybridization platform. Hum Reprod. 2014;29(8):1818-27.
158. Chand AL, Harrison CA, Shelling AN. Inhibin and premature ovarian failure. Hum Reprod Update. 2010;16(1):39-50.
159. Corre T, Schuettler J, Bione S, Marozzi A, Persani L, Rossetti R, et al. A large-scale association study to assess the impact of known variants of the human INHA gene on premature ovarian failure. Hum Reprod. 2009;24(8):2023-8.
160. Baetens D, Güran T, Mendonca BB, Gomes NL, De Cauwer L, Peelman F, et al. Biallelic and monoallelic ESR2 variants associated with 46,XY disorders of sex development. Genet Med. 2018;20(7):717-27.
161. Lang-Muritano M, Sproll P, Wyss S, Kolly A, Hürlimann R, Konrad D, et al. Early-Onset Complete Ovarian Failure and Lack of Puberty in a Woman With Mutated Estrogen Receptor β (ESR2). J Clin Endocrinol Metab. 2018;103(10):3748-56.
162. Bernard V, Kherra S, Francou B, Fagart J, Viengchareun S, Guéchot J, et al. Familial Multiplicity of Estrogen Insensitivity Associated With a Loss-of-Function ESR1 Mutation. J Clin Endocrinol Metab. 2017;102(1):93-9.
163. Lerat J, Jonard L, Loundon N, Christin-Maitre S, Lacombe D, Goizet C, et al. An Application of NGS for Molecular Investigations in Perrault Syndrome: Study of 14 Families and Review of the Literature. Hum Mutat. 2016;37(12):1354-62.
164. Jenkinson EM, Clayton-Smith J, Mehta S, Bennett C, Reardon W, Green A, et al. Perrault syndrome: further evidence for genetic heterogeneity. J Neurol. 2012;259(5):974-6.
165. Zangen D, Kaufman Y, Zeligson S, Perlberg S, Fridman H, Kanaan M, et al. XX ovarian dysgenesis is caused by a PSMC3IP/HOP2 mutation that abolishes coactivation of estrogen-driven transcription. Am J Hum Genet. 2011;89(4):572-9.
166. Al-Agha AE, Ahmed IA, Nuebel E, Moriwaki M, Moore B, Peacock KA, et al. Primary Ovarian Insufficiency and Azoospermia in Carriers of a Homozygous PSMC3IP Stop Gain Mutation. J Clin Endocrinol Metab. 2018;103(2):555-63.
167. Fowler PA, Anderson RA, Saunders PT, Kinnell H, Mason JI, Evans DB, et al. Development of steroid signaling pathways during primordial follicle formation in the human fetal ovary. J Clin Endocrinol Metab. 2011;96(6):1754-62.
168. Windley SP, Wilhelm D. Signaling Pathways Involved in Mammalian Sex Determination and Gonad Development. Sex Dev. 2015;9(6):297-315.
169. Eid W, Opitz L, Biason-Lauber A. Genome-wide identification of CBX2 targets: insights in the human sex development network. Mol Endocrinol. 2015;29(2):247-57.
170. Yao HH, Matzuk MM, Jorgez CJ, Menke DB, Page DC, Swain A, et al. Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. Dev Dyn. 2004;230(2):210-5.
171. Britt KL, Kerr J, O'Donnell L, Jones ME, Drummond AE, Davis SR, et al. Estrogen regulates development of the somatic cell phenotype in the eutherian ovary. FASEB J. 2002;16(11):1389-97.
172. Hughes IA, Houk C, Ahmed SF, Lee PA, Group LWPESESfPEC. Consensus statement on management of intersex disorders. J Pediatr Urol. 2006;2(3):148-62.
173. Ahmed SF, Dobbie R, Finlayson AR, Gilbert J, Youngson G, Chalmers J, et al. Prevalence of hypospadias and other genital anomalies among singleton births, 1988-1997, in Scotland. Arch Dis Child Fetal Neonatal Ed. 2004;89(2):F149-51.
174. Audi L, Ahmed SF, Krone N, Cools M, McElreavey K, Holterhus PM, et al. GENETICS IN ENDOCRINOLOGY: Approaches to molecular genetic diagnosis in the management of

differences/disorders of sex development (DSD): position paper of EU COST Action BM 1303 'DSDnet'. Eur J Endocrinol. 2018;179(4):R197-R206.
175. Mendonca BB, Domenice S, Arnhold IJ, Costa EM. 46,XY disorders of sex development (DSD). Clin Endocrinol (Oxf). 2009;70(2):173-87.
176. Arboleda VA, Sandberg DE, Vilain E. DSDs: genetics, underlying pathologies and psychosexual differentiation. Nat Rev Endocrinol. 2014;10(10):603-15.
177. Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, et al. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature. 1990;346(6281):240-4.
178. Heksch RA, Matheson MA, Tishelman AC, Swartz JM, Jayanthi VR, Diamond DA, et al. TESTICULAR REGRESSION SYNDROME: PRACTICE VARIATION IN DIAGNOSIS AND MANAGEMENT. Endocr Pract. 2019.
179. Fischbach BV, Trout KL, Lewis J, Luis CA, Sika M. WAGR syndrome: a clinical review of 54 cases. Pediatrics. 2005;116(4):984-8.
180. Miller-Hodges E. Clinical Aspects of WT1 and the Kidney. Methods Mol Biol. 2016;1467:15-21.
181. Hornig NC, Ukat M, Schweikert HU, Hiort O, Werner R, Drop SL, et al. Identification of an AR Mutation-Negative Class of Androgen Insensitivity by Determining Endogenous AR Activity. J Clin Endocrinol Metab. 2016;101(11):4468-77.
182. Picard JY, Cate RL, Racine C, Josso N. The Persistent Müllerian Duct Syndrome: An Update Based Upon a Personal Experience of 157 Cases. Sex Dev. 2017;11(3):109-25.
183. Bouty A, Ayers KL, Pask A, Heloury Y, Sinclair AH. The Genetic and Environmental Factors Underlying Hypospadias. Sex Dev. 2015;9(5):239-59.
184. Vikraman J, Hutson JM, Li R, Thorup J. The undescended testis: Clinical management and scientific advances. Semin Pediatr Surg. 2016;25(4):241-8.
185. Topaloğlu AK. Update on the Genetics of Idiopathic Hypogonadotropic Hypogonadism. J Clin Res Pediatr Endocrinol. 2017;9(Suppl 2):113-22.
186. Grinspon RP, Freire AV, Rey RA. Hypogonadism in Pediatric Health: Adult Medicine Concepts Fail. Trends Endocrinol Metab. 2019;30(12):879-90.
187. Salonia A, Rastrelli G, Hackett G, Seminara SB, Huhtaniemi IT, Rey RA, et al. Paediatric and adult-onset male hypogonadism. Nat Rev Dis Primers. 2019;5(1):38.
188. Dutta D, Shivaprasad KS, Das RN, Ghosh S, Chatterjee U, Chowdhury S, et al. Ovotesticular disorder of sexual development due to 47,XYY/46,XY/45,X mixed gonadal dysgenesis in a phenotypic male presenting as cyclical haematuria: clinical presentation and assessment of long-term outcomes. Andrologia. 2014;46(2):191-3.
189. van Niekerk WA, Retief AE. The gonads of human true hermaphrodites. Hum Genet. 1981;58(1):117-22.
190. Grinspon RP, Rey RA. Molecular Characterization of XX Maleness. Int J Mol Sci. 2019;20(23).
191. Jordan BK, Mohammed M, Ching ST, Délot E, Chen XN, Dewing P, et al. Up-regulation of WNT-4 signaling and dosage-sensitive sex reversal in humans. Am J Hum Genet. 2001;68(5):1102-9.
192. Parma P, Radi O, Vidal V, Chaboissier MC, Dellambra E, Valentini S, et al. R-spondin1 is essential in sex determination, skin differentiation and malignancy. Nat Genet. 2006;38(11):1304-9.
193. Aleck KA, Argueso L, Stone J, Hackel JG, Erickson RP. True hermaphroditism with partial duplication of chromosome 22 and without SRY. Am J Med Genet. 1999;85(1):2-4.
194. Baetens D, Stoop H, Peelman F, Todeschini AL, Rosseel T, Coppieters F, et al. NR5A1 is a novel disease gene for 46,XX testicular and ovotesticular disorders of sex development. Genet Med. 2017;19(4):367-76.
195. Bashamboo A, Eozenou C, Rojo S, McElreavey K. Anomalies in human sex determination provide unique insights into the complex genetic interactions of early gonad development. Clin Genet. 2017;91(2):143-56.

196. Ledig S, Hiort O, Wünsch L, Wieacker P. Partial deletion of DMRT1 causes 46,XY ovotesticular disorder of sexual development. Eur J Endocrinol. 2012;167(1):119-24.
197. Meyers CM, Boughman JA, Rivas M, Wilroy RS, Simpson JL. Gonadal (ovarian) dysgenesis in 46,XX individuals: frequency of the autosomal recessive form. Am J Med Genet. 1996;63(4):518-24.
198. Ferrari I, Bouilly J, Beau I, Guizzardi F, Ferlin A, Pollazzon M, et al. Impaired protein stability and nuclear localization of NOBOX variants associated with premature ovarian insufficiency. Hum Mol Genet. 2016;25(23):5223-33.
199. Zhao H, Chen ZJ, Qin Y, Shi Y, Wang S, Choi Y, et al. Transcription factor FIGLA is mutated in patients with premature ovarian failure. Am J Hum Genet. 2008;82(6):1342-8.
200. Kaňová N, Bičíková M. Hyperandrogenic states in pregnancy. Physiol Res. 2011;60(2):243-52.
201. Morris LF, Park S, Daskivich T, Churchill BM, Rao CV, Lei Z, et al. Virilization of a female infant by a maternal adrenocortical carcinoma. Endocr Pract. 2011;17(2):e26-31.
202. Narumi S. Rare monogenic causes of primary adrenal insufficiency. Curr Opin Endocrinol Diabetes Obes. 2018;25(3):172-7.
203. Meimaridou E, Kowalczyk J, Guasti L, Hughes CR, Wagner F, Frommolt P, et al. Mutations in NNT encoding nicotinamide nucleotide transhydrogenase cause familial glucocorticoid deficiency. Nat Genet. 2012;44(7):740-2.
204. Ledig S, Schippert C, Strick R, Beckmann MW, Oppelt PG, Wieacker P. Recurrent aberrations identified by array-CGH in patients with Mayer-Rokitansky-Küster-Hauser syndrome. Fertil Steril. 2011;95(5):1589-94.
205. Takahashi K, Hayano T, Sugimoto R, Kashiwagi H, Shinoda M, Nishijima Y, et al. Exome and copy number variation analyses of Mayer-Rokitansky-Küster- Hauser syndrome. Hum Genome Var. 2018;5:27.
206. Gravholt CH, Viuff MH, Brun S, Stochholm K, Andersen NH. Turner syndrome: mechanisms and management. Nat Rev Endocrinol. 2019.
207. Akcan N, Poyrazoğlu Ş, Baş F, Bundak R, Darendeliler F. Klinefelter Syndrome in Childhood: Variability in Clinical and Molecular Findings. J Clin Res Pediatr Endocrinol. 2018;10(2):100-7.
208. Cools M, Pleskacova J, Stoop H, Hoebeke P, Van Laecke E, Drop SL, et al. Gonadal pathology and tumor risk in relation to clinical characteristics in patients with 45,X/46,XY mosaicism. J Clin Endocrinol Metab. 2011;96(7):E1171-80.
209. Ocal G, Berberoğlu M, Sıklar Z, Ruhi HI, Tükün A, Camtosun E, et al. The clinical and genetic heterogeneity of mixed gonadal dysgenesis: does "disorders of sexual development (DSD)" classification based on new Chicago consensus cover all sex chromosome DSD? Eur J Pediatr. 2012;171(10):1497-502.
210. Hornig NC, Demiri J, Rodens P, Murga Penas EM, Caliebe A, Eckstein AK, et al. Reduced androgen receptor expression in genital skin fibroblasts from patients with 45,X/46,XY mosaicism. J Clin Endocrinol Metab. 2019.
211. Patsalis PC, Skordis N, Sismani C, Kousoulidou L, Koumbaris G, Eftychi C, et al. Identification of high frequency of Y chromosome deletions in patients with sex chromosome mosaicism and correlation with the clinical phenotype and Y-chromosome instability. Am J Med Genet A. 2005;135(2):145-9.
212. Silber SJ. The Y chromosome in the era of intracytoplasmic sperm injection: a personal review. Fertil Steril. 2011;95(8):2439-48.e1-5.
213. Hiort O, Birnbaum W, Marshall L, Wünsch L, Werner R, Schröder T, et al. Management of disorders of sex development. Nat Rev Endocrinol. 2014;10(9):520-9.
214. Cools M, Looijenga LH, Wolffenbuttel KP, Drop SL. Disorders of sex development: update on the genetic background, terminology and risk for the development of germ cell tumors. World J Pediatr. 2009;5(2):93-102.

215. Cools M, Looijenga LH, Wolffenbuttel KP, T'Sjoen G. Managing the risk of germ cell tumourigenesis in disorders of sex development patients. Endocr Dev. 2014;27:185-96.
216. Tobias ES, McElreavey K. Next generation sequencing for disorders of sex development. Endocr Dev. 2014;27:53-62.
217. Dong Y, Yi Y, Yao H, Yang Z, Hu H, Liu J, et al. Targeted next-generation sequencing identification of mutations in patients with disorders of sex development. BMC Med Genet. 2016;17:23.
218. Fan Y, Zhang X, Wang L, Wang R, Huang Z, Sun Y, et al. Diagnostic Application of Targeted Next-Generation Sequencing of 80 Genes Associated with Disorders of Sexual Development. Sci Rep. 2017;7:44536.
219. Hughes LA, McKay Bounford K, Webb E, Dasani P, Clokie S, Chandran H, et al. Next generation sequencing (NGS) to improve the diagnosis and management of patients with disorders of sex development (DSD). Endocr Connect. 2019.
220. Özen S, Onay H, Atik T, Solmaz AE, Özkınay F, Gökşen D, et al. Rapid Molecular Genetic Diagnosis with Next-Generation Sequencing in 46,XY Disorders of Sex Development Cases: Efficiency and Cost Assessment. Horm Res Paediatr. 2017;87(2):81-7.
221. Kon M, Fukami M. Submicroscopic copy-number variations associated with 46,XY disorders of sex development. Mol Cell Pediatr. 2015;2(1):7.
222. Ledig S, Hiort O, Scherer G, Hoffmann M, Wolff G, Morlot S, et al. Array-CGH analysis in patients with syndromic and non-syndromic XY gonadal dysgenesis: evaluation of array CGH as diagnostic tool and search for new candidate loci. Hum Reprod. 2010;25(10):2637-46.
223. Croft B, Ohnesorg T, Sinclair AH. The Role of Copy Number Variants in Disorders of Sex Development. Sex Dev. 2018;12(1-3):19-29.
224. Camats N, Fernández-Cancio M, Audí L, Schaller A, Flück CE. Broad phenotypes in heterozygous NR5A1 46,XY patients with a disorder of sex development: an oligogenic origin? Eur J Hum Genet. 2018;26(9):1329-38.
225. Lee Y, Kim C, Park Y, Pyun JA, Kwack K. Next generation sequencing identifies abnormal Y chromosome and candidate causal variants in premature ovarian failure patients. Genomics. 2016;108(5-6):209-15.
226. Cools M, Nordenström A, Robeva R, Hall J, Westerveld P, Flück C, et al. Caring for individuals with a difference of sex development (DSD): a Consensus Statement. Nat Rev Endocrinol. 2018;14(7):415-29.
227. Guerrero-Fernández J, Azcona San Julián C, Barreiro Conde J, Bermúdez de la Vega JA, Carcavilla Urquí A, Castaño González LA, et al. [Management guidelines for disorders / different sex development (DSD)]. An Pediatr (Barc). 2018;89(5):315.e1-.e19.
228. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977;74(12):5463-7.
229. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.
230. Stenson PD, Mort M, Ball EV, Howells K, Phillips AD, Thomas NS, et al. The Human Gene Mutation Database: 2008 update. Genome Med. 2009;1(1):13.
231. McKusick VA. Mendelian Inheritance in Man and its online version, OMIM. Am J Hum Genet. 2007;80(4):588-604.
232. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. VarSome: the human genomic variant search engine. Bioinformatics. 2019;35(11):1978-80.
233. Rivero-Müller A, Potorac I, Pintiaux A, Daly AF, Thiry A, Rydlewski C, et al. A novel inactivating mutation of the LH/chorionic gonadotrophin receptor with impaired membrane trafficking leading to Leydig cell hypoplasia type 1. Eur J Endocrinol. 2015;172(6):K27-36.

234. Kohler PO, Bridson WE. Isolation of hormone-producing clonal lines of human choriocarcinoma. J Clin Endocrinol Metab. 1971;32(5):683-7.
235. Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol. 1977;36(1):59-74.
236. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagn. 2017;19(3):341-65.
237. So J, Suckow V, Kijas Z, Kalscheuer V, Moser B, Winter J, et al. Mild phenotypes in a series of patients with Opitz GBBB syndrome with MID1 mutations. Am J Med Genet A. 2005;132A(1):1-7.
238. Bilbao JR, Loridan L, Castaño L. A novel postzygotic nonsense mutation in SRY in familial XY gonadal dysgenesis. Hum Genet. 1996;97(4):537-9.
239. Y L, M R, J H, A W. Novel Missense mutation (P131R) in the HMG box of S*RY* in XY sex reversal. Hum Mutat; Suppl. 1:S328.1998.
240. Reuter AL, Goji K, Bingham NC, Matsuo M, Parker KL. A novel mutation in the accessory DNA-binding domain of human steroidogenic factor 1 causes XY gonadal dysgenesis without adrenal insufficiency. Eur J Endocrinol. 2007;157(2):233-8.
241. Yu B, Liu Z, Gao Y, Mao J, Wang X, Hao M, et al. Novel NR5A1 mutations found in Chinese patients with 46, XY disorders of sex development. Clin Endocrinol (Oxf). 2018;89(5):613-20.
242. WuQiang F, Yanase T, Wei L, Oba K, Nomura M, Okabe T, et al. Functional characterization of a new human Ad4BP/SF-1 variation, G146A. Biochem Biophys Res Commun. 2003;311(4):987-94.
243. Liu W, Liu M, Fan W, Nawata H, Yanase T. The Gly146Ala variation in human SF-1 gene: its association with insulin resistance and type 2 diabetes in Chinese. Diabetes Res Clin Pract. 2006;73(3):322-8.
244. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, et al. Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. Genet Med. 2017;19(10):1105-17.
245. Rodríguez Estévez A, Pérez-Nanclares G, Fernández-Toral J, Rivas-Crespo F, López-Siguero JP, Díez I, et al. Clinical and molecular characterization of five Spanish kindreds with X-linked adrenal hypoplasia congenita: atypical findings and a novel mutation in NR0B1. J Pediatr Endocrinol Metab. 2015;28(9-10):1129-37.
246. Guoying C, Zhiya D, Wei W, Na L, Xiaoying L, Yuan X, et al. The analysis of clinical manifestations and genetic mutations in Chinese boys with primary adrenal insufficiency. J Pediatr Endocrinol Metab. 2012;25(3-4):295-300.
247. Lehmann SG, Wurtz JM, Renaud JP, Sassone-Corsi P, Lalli E. Structure-function analysis reveals the molecular determinants of the impaired biological function of DAX-1 mutants in AHC patients. Hum Mol Genet. 2003;12(9):1063-72.
248. Yeste D, González-Niño C, Pérez de Nanclares G, Pérez-Nanclares G, Audi L, Castaño L, et al. ACTH-dependent precocious pseudopuberty in an infant with DAX1 gene mutation. Eur J Pediatr. 2009;168(1):65-9.
249. Guo W, Mason JS, Stone CG, Morgan SA, Madu SI, Baldini A, et al. Diagnosis of X-linked adrenal hypoplasia congenita by mutation analysis of the DAX1 gene. JAMA. 1995;274(4):324-30.
250. Muroya K, Kinoshita E, Kamimaki T, Matsuo N, Yorifugi T, Ogata T. Deletion mapping and X inactivation analysis of a non-specific mental retardation gene at Xp21.3-Xp22.11. J Med Genet. 1999;36(3):187-91.
251. Bruening W, Bardeesy N, Silverman BL, Cohn RA, Machin GA, Aronson AJ, et al. Germline intronic and exonic mutations in the Wilms' tumour gene (WT1) affecting urogenital development. Nat Genet. 1992;1(2):144-8.

252. Nakadate H, Yokomori K, Watanabe N, Tsuchiya T, Namiki T, Kobayshi H, et al. Mutations/deletions of the WT1 gene, loss of heterozygosity on chromosome arms 11p and 11q, chromosome ploidy and histology in Wilms' tumors in Japan. Int J Cancer. 2001;94(3):396-400.
253. Finken MJ, Hendriks YM, van der Voorn JP, Veening MA, Lombardi MP, Rotteveel J. WT1 deletion leading to severe 46,XY gonadal dysgenesis, Wilms tumor and gonadoblastoma: case report. Horm Res Paediatr. 2015;83(3):211-6.
254. McPhaul MJ, Marcelli M, Zoppi S, Wilson CM, Griffin JE, Wilson JD. Mutations in the ligand-binding domain of the androgen receptor gene cluster in two regions of the gene. J Clin Invest. 1992;90(5):2097-101.
255. Holterhus PM, Hiort O, Demeter J, Brown PO, Brooks JD. Differential gene-expression patterns in genital fibroblasts of normal males and 46,XY females with androgen insensitivity syndrome: evidence for early programming involving the androgen receptor. Genome Biol. 2003;4(6):R37.
256. Brown TR, Lubahn DB, Wilson EM, French FS, Migeon CJ, Corden JL. Functional characterization of naturally occurring mutant androgen receptors from subjects with complete androgen insensitivity. Mol Endocrinol. 1990;4(12):1759-72.
257. Ris-Stalpers C, Trifiro MA, Kuiper GG, Jenster G, Romalo G, Sai T, et al. Substitution of aspartic acid-686 by histidine or asparagine in the human androgen receptor leads to a functionally inactive protein with altered hormone-binding characteristics. Mol Endocrinol. 1991;5(10):1562-9.
258. Giwercman A, Kledal T, Schwartz M, Giwercman YL, Leffers H, Zazzi H, et al. Preserved male fertility despite decreased androgen sensitivity caused by a mutation in the ligand-binding domain of the androgen receptor gene. J Clin Endocrinol Metab. 2000;85(6):2253-9.
259. Hellmann P, Christiansen P, Johannsen TH, Main KM, Duno M, Juul A. Male patients with partial androgen insensitivity syndrome: a longitudinal follow-up of growth, reproductive hormones and the development of gynaecomastia. Arch Dis Child. 2012;97(5):403-9.
260. Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS. Androgen receptor defects: historical, clinical, and molecular perspectives. Endocr Rev. 1995;16(3):271-321.
261. Batch JA, Williams DM, Davies HR, Brown BD, Evans BA, Hughes IA, et al. Androgen receptor gene mutations identified by SSCP in fourteen subjects with androgen insensitivity syndrome. Hum Mol Genet. 1992;1(7):497-503.
262. Hiort O, Sinnecker GH, Holterhus PM, Nitsche EM, Kruse K. The clinical and molecular spectrum of androgen insensitivity syndromes. Am J Med Genet. 1996;63(1):218-22.
263. Shao J, Hou J, Li B, Li D, Zhang N, Wang X. Different types of androgen receptor mutations in patients with complete androgen insensitivity syndrome. Intractable Rare Dis Res. 2015;4(1):54-9.
264. Wijeratne N, McNeil AR, Doery JCG, McLeod E, Bergman PB, Montalto J. A Teenage Girl with Unexpected Pubertal Changes. Clin Chem. 2018;64(6):892-6.
265. Fénichel P, Paris F, Philibert P, Hiéronimus S, Gaspari L, Kurzenne JY, et al. Molecular diagnosis of 5α-reductase deficiency in 4 elite young female athletes through hormonal screening for hyperandrogenism. J Clin Endocrinol Metab. 2013;98(6):E1055-9.
266. Wilson JD, Griffin JE, Russell DW. Steroid 5 alpha-reductase 2 deficiency. Endocr Rev. 1993;14(5):577-93.
267. Kremer H, Mariman E, Otten BJ, Moll GW, Stoelinga GB, Wit JM, et al. Cosegregation of missense mutations of the luteinizing hormone receptor gene with familial male-limited precocious puberty. Hum Mol Genet. 1993;2(11):1779-83.
268. Kraaij R, Post M, Kremer H, Milgrom E, Epping W, Brunner HG, et al. A missense mutation in the second transmembrane segment of the luteinizing hormone receptor causes familial male-limited precocious puberty. J Clin Endocrinol Metab. 1995;80(11):3168-72.

269. Latronico AC, Anasti J, Arnhold IJ, Rapaport R, Mendonca BB, Bloise W, et al. Brief report: testicular and ovarian resistance to luteinizing hormone caused by inactivating mutations of the luteinizing hormone-receptor gene. N Engl J Med. 1996;334(8):507-12.
270. Richard N, Leprince C, Gruchy N, Pigny P, Andrieux J, Mittre H, et al. Identification by array-Comparative Genomic Hybridization (array-CGH) of a large deletion of luteinizing hormone receptor gene combined with a missense mutation in a patient diagnosed with a 46,XY disorder of sex development and application to prenatal diagnosis. Endocr J. 2011;58(9):769-76.
271. Takeda Y, Yoneda T, Demura M, Furukawa K, Koshida H, Miyamori I, et al. Genetic analysis of the cytochrome P-450c17alpha (CYP17) and aldosterone synthase (CYP11B2) in Japanese patients with 17alpha-hydroxylase deficiency. Clin Endocrinol (Oxf). 2001;54(6):751-8.
272. Kulle A, Krone N, Holterhus PM, Schuler G, Greaves RF, Juul A, et al. Steroid hormone analysis in diagnosis and treatment of DSD: position paper of EU COST Action BM 1303 'DSDnet'. Eur J Endocrinol. 2017;176(5):P1-P9.
273. Johannsen TH, Main KM, Ljubicic ML, Jensen TK, Andersen HR, Andersen MS, et al. Sex Differences in Reproductive Hormones During Mini-Puberty in Infants With Normal and Disordered Sex Development. J Clin Endocrinol Metab. 2018;103(8):3028-37.
274. Ahmed SF, Achermann JC, Arlt W, Balen A, Conway G, Edwards Z, et al. Society for Endocrinology UK guidance on the initial evaluation of an infant or an adolescent with a suspected disorder of sex development (Revised 2015). Clin Endocrinol (Oxf). 2016;84(5):771-88.
275. van der Straaten S, Springer A, Zecic A, Hebenstreit D, Tonnhofer U, Gawlik A, et al. The External Genitalia Score (EGS): A European multicenter validation study. J Clin Endocrinol Metab. 2019.
276. Tannour-Louet M, Han S, Corbett ST, Louet JF, Yatsenko S, Meyers L, et al. Identification of de novo copy number variants associated with human disorders of sexual development. PLoS One. 2010;5(10):e15392.
277. Barbaro M, Cicognani A, Balsamo A, Löfgren A, Baldazzi L, Wedell A, et al. Gene dosage imbalances in patients with 46,XY gonadal DSD detected by an in-house-designed synthetic probe set for multiplex ligation-dependent probe amplification analysis. Clin Genet. 2008;73(5):453-64.
278. Golan D, Medvedev P. Using state machines to model the Ion Torrent sequencing process and to improve read error rates. Bioinformatics. 2013;29(13):i344-51.
279. Arboleda VA, Lee H, Sánchez FJ, Délot EC, Sandberg DE, Grody WW, et al. Targeted massively parallel sequencing provides comprehensive genetic diagnosis for patients with disorders of sex development. Clin Genet. 2013;83(1):35-43.
280. Holterhus PM, Werner R, Struve D, Hauffa BP, Schroeder C, Hiort O. Mutations in the amino-terminal domain of the human androgen receptor may be associated with partial androgen insensitivity and impaired transactivation in vitro. Exp Clin Endocrinol Diabetes. 2005;113(8):457-63.
281. Galani A, Sofocleous C, Karahaliou F, Papathanasiou A, Kitsiou-Tzeli S, Kalpini-Mavrou A. Sex-reversed phenotype in association with two novel mutations c.2494delA and c.T3004C in the ligand-binding domain of the androgen receptor gene. Fertil Steril. 2008;90(5):2008.e1-4.
282. Davies HR, Hughes IA, Savage MO, Quigley CA, Trifiro M, Pinsky L, et al. Androgen insensitivity with mental retardation: a contiguous gene syndrome? J Med Genet. 1997;34(2):158-60.
283. Bose HS, Sugawara T, Strauss JF, Miller WL, Consortium ICLAH. The pathophysiology and genetics of congenital lipoid adrenal hyperplasia. N Engl J Med. 1996;335(25):1870-8.
284. Russell DW, Wilson JD. Steroid 5 alpha-reductase: two genes/two enzymes. Annu Rev Biochem. 1994;63:25-61.

285. Wigley WC, Prihoda JS, Mowszowicz I, Mendonca BB, New MI, Wilson JD, et al. Natural mutagenesis study of the human steroid 5 alpha-reductase 2 isozyme. Biochemistry. 1994;33(5):1265-70.
286. Forti G, Falchetti A, Santoro S, Davis DL, Wilson JD, Russell DW. Steroid 5 alpha-reductase 2 deficiency: virilization in early infancy may be due to partial function of mutant enzyme. Clin Endocrinol (Oxf). 1996;44(4):477-82.
287. Fernández-Cancio M, Audí L, Andaluz P, Torán N, Piró C, Albisu M, et al. SRD5A2 gene mutations and polymorphisms in Spanish 46,XY patients with a disorder of sex differentiation. Int J Androl. 2011;34(6 Pt 2):e526-35.
288. Hackel C, Oliveira LE, Toralles MB, Nunes-Silva D, Tonini MM, Ferraz LF, et al. [5alpha-reductase type 2 deficiency: experiences from Campinas (SP) and Salvador (BA)]. Arq Bras Endocrinol Metabol. 2005;49(1):103-11.
289. Phelan N, Williams EL, Cardamone S, Lee M, Creighton SM, Rumsby G, et al. Screening for mutations in 17β-hydroxysteroid dehydrogenase and androgen receptor in women presenting with partially virilised 46,XY disorders of sex development. Eur J Endocrinol. 2015;172(6):745-51.
290. Pearlman A, Loke J, Le Caignec C, White S, Chin L, Friedman A, et al. Mutations in MAP3K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. Am J Hum Genet. 2010;87(6):898-904.
291. Köhler B, Biebermann H, Friedsam V, Gellermann J, Maier RF, Pohl M, et al. Analysis of the Wilms' tumor suppressor gene (WT1) in patients 46,XY disorders of sex development. J Clin Endocrinol Metab. 2011;96(7):E1131-6.
292. Dabrowski E, Armstrong AE, Leeth E, Johnson E, Cheng E, Gosiengfiao Y, et al. Proximal Hypospadias and a Novel. Pediatrics. 2018;141(Suppl 5):S491-S5.
293. Coopes MJ. The management and biological behavior of Wilms tumor. Int Rev Exp Pathol. 1994;35:149-76.
294. Koziell A, Charmandari E, Hindmarsh PC, Rees L, Scambler P, Brook CG. Frasier syndrome, part of the Denys Drash continuum or simply a WT1 gene associated disorder of intersex and nephropathy? Clin Endocrinol (Oxf). 2000;52(4):519-24.
295. Buonocore F, Achermann JC. Primary adrenal insufficiency: New genetic causes and their long-term consequences. Clin Endocrinol (Oxf). 2019.
296. Kinoshita E, Yoshimoto M, Motomura K, Kawaguchi T, Mori R, Baba T, et al. DAX-1 gene mutations and deletions in Japanese patients with adrenal hypoplasia congenita and hypogonadotropic hypogonadism. Horm Res. 1997;48(1):29-34.
297. Piard J, Hawkes L, Milh M, Villard L, Borgatti R, Romaniello R, et al. The phenotypic spectrum of WWOX-related disorders: 20 additional cases of WOREE syndrome and review of the literature. Genet Med. 2019;21(6):1308-18.
298. Kim GJ, Sock E, Buchberger A, Just W, Denzer F, Hoepffner W, et al. Copy number variation of two separate regulatory regions upstream of SOX9 causes isolated 46,XY or 46,XX disorder of sex development. J Med Genet. 2015;52(4):240-7.
299. Mizuno K, Kojima Y, Kamisawa H, Moritoki Y, Nishio H, Nakane A, et al. Elucidation of distinctive genomic DNA structures in patients with 46,XX testicular disorders of sex development using genome wide analyses. J Urol. 2014;192(2):535-41.
300. Speiser PW, White PC. Congenital adrenal hyperplasia. N Engl J Med. 2003;349(8):776-88.
301. Szenker-Ravi E, Altunoglu U, Leushacke M, Bosso-Lefèvre C, Khatoo M, Thi Tran H, et al. RSPO2 inhibition of RNF43 and ZNRF3 governs limb development independently of LGR4/5/6. Nature. 2018;557(7706):564-9.
302. Krege JH, Hodgin JB, Couse JF, Enmark E, Warner M, Mahler JF, et al. Generation and reproductive phenotypes of mice lacking estrogen receptor beta. Proc Natl Acad Sci U S A. 1998;95(26):15677-82.
303. Beleza-Meireles A, Kockum I, Lundberg F, Söderhäll C, Nordenskjöld A. Risk factors for hypospadias in the estrogen receptor 2 gene. J Clin Endocrinol Metab. 2007;92(9):3712-8.

304. Brucker SY, Frank L, Eisenbeis S, Henes M, Wallwiener D, Riess O, et al. Sequence variants in ESR1 and OXTR are associated with Mayer-Rokitansky-Küster-Hauser syndrome. Acta Obstet Gynecol Scand. 2017;96(11):1338-46.
305. Känsäkoski J, Jääskeläinen J, Jääskeläinen T, Tommiska J, Saarinen L, Lehtonen R, et al. Complete androgen insensitivity syndrome caused by a deep intronic pseudoexon-activating mutation in the androgen receptor gene. Sci Rep. 2016;6:32819.
306. Bertier G, Sénécal K, Borry P, Vears DF. Unsolved challenges in pediatric whole-exome sequencing: A literature analysis. Crit Rev Clin Lab Sci. 2017;54(2):134-42.
307. dos Santos AP, Andrade JG, Piveta CS, de Paulo J, Guerra G, de Mello MP, et al. Screening of Y chromosome microdeletions in 46,XY partial gonadal dysgenesis and in patients with a 45,X/46,XY karyotype or its variants. BMC Med Genet. 2013;14:115.
308. Kuroki S, Matoba S, Akiyoshi M, Matsumura Y, Miyachi H, Mise N, et al. Epigenetic regulation of mouse sex determination by the histone demethylase Jmjd1a. Science. 2013;341(6150):1106-9.
309. Mazen I, Abdel-Hamid M, Mekkawy M, Bignon-Topalovic J, Boudjenah R, El Gammal M, et al. Identification of NR5A1 Mutations and Possible Digenic Inheritance in 46,XY Gonadal Dysgenesis. Sex Dev. 2016;10(3):147-51.
310. Werner R, Mönig I, Lünstedt R, Wünsch L, Thorns C, Reiz B, et al. New NR5A1 mutations and phenotypic variations of gonadal dysgenesis. PLoS One. 2017;12(5):e0176720.
311. Lin L, Achermann JC. Steroidogenic factor-1 (SF-1, Ad4BP, NR5A1) and disorders of testis development. Sex Dev. 2008;2(4-5):200-9.
312. Zhang X, Wang J, Wang B, Chen S, Fu Q, Sun K. A Novel Missense Mutation of GATA4 in a Chinese Family with Congenital Heart Disease. PLoS One. 2016;11(7):e0158904.
313. Pulignani S, Vecoli C, Sabina S, Foffa I, Ait-Ali L, Andreassi MG. 3'UTR SNPs and Haplotypes in the GATA4 Gene Contribute to the Genetic Risk of Congenital Heart Disease. Rev Esp Cardiol (Engl Ed). 2016;69(8):760-5.
314. Molkentin JD. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J Biol Chem. 2000;275(50):38949-52.
315. Bashamboo A, Ledig S, Wieacker P, Achermann JC, Achermann J, McElreavey K. New technologies for the identification of novel genetic markers of disorders of sex development (DSD). Sex Dev. 2010;4(4-5):213-24.
316. Jay PY, Bielinska M, Erlich JM, Mannisto S, Pu WT, Heikinheimo M, et al. Impaired mesenchymal cell function in Gata4 mutant mice leads to diaphragmatic hernias and primary lung defects. Dev Biol. 2007;301(2):602-14.
317. Miyamoto Y, Taniguchi H, Hamel F, Silversides DW, Viger RS. A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. BMC Mol Biol. 2008;9:44.
318. Domenice S, Machado AZ, Ferreira FM, Ferraz-de-Souza B, Lerario AM, Lin L, et al. Wide spectrum of NR5A1-related phenotypes in 46,XY and 46,XX individuals. Birth Defects Res C Embryo Today. 2016;108(4):309-20.
319. Potorac I, Trehan A, Szymanska K, Fudvoye J, Thiry A, Huhtaniemi IT, et al. Compound heterozygous mutations in the luteinizing hormone receptor signal peptide causing 46,XY disorder of sex development. Eur J Endocrinol. 2019.
320. Bhowmick N, Huang J, Puett D, Isaacs NW, Lapthorn AJ. Determination of residues important in hormone binding to the extracellular domain of the luteinizing hormone/chorionic gonadotropin receptor by site-directed mutagenesis and modeling. Mol Endocrinol. 1996;10(9):1147-59.
321. Newton CL, Anderson RC, Katz AA, Millar RP. Loss-of-Function Mutations in the Human Luteinizing Hormone Receptor Predominantly Cause Intracellular Retention. Endocrinology. 2016;157(11):4364-77.

322. Upadhyaya AB, Lee SH, DeJong J. Identification of a general transcription factor TFIIAalpha/beta homolog selectively expressed in testis. J Biol Chem. 1999;274(25):18040-8.
323. Özcabı B, Tahmiscioğlu Bucak F, Ceylaner S, Özcan R, Büyükünal C, Ercan O, et al. Testotoxicosis: Report of Two Cases, One with a Novel Mutation in LHCGR Gene. J Clin Res Pediatr Endocrinol. 2015;7(3):242-8.
324. Latronico AC, Shinozaki H, Guerra G, Pereira MA, Lemos Marini SH, Baptista MT, et al. Gonadotropin-independent precocious puberty due to luteinizing hormone receptor mutations in Brazilian boys: a novel constitutively activating mutation in the first transmembrane helix. J Clin Endocrinol Metab. 2000;85(12):4799-805.
325. Shenker A, Laue L, Kosugi S, Merendino JJ, Minegishi T, Cutler GB. A constitutively activating mutation of the luteinizing hormone receptor in familial male precocious puberty. Nature. 1993;365(6447):652-4.

SUPPLEMENTARY DATA

Supplementary data 1. Expected clinical and biochemical features of the different DSD. Modified from Guerrero- Fernandez *et al* (227).

			Gonads	Internal genitalia		External genitalia	Bioquemical defect
				Wolffian duct derivatives	Müllerian structures		
46,XY DSD	Disorders of testicular development	Gonadal dysgenesis	Dysgenetic	Complete: absent; Partial: hypoplastic	Complete: present; Partial: absent	Complete: female; Partial: variable virilization	Low T[1,2], low AMH
		46,XY OT DSD	Dysgenetic (Same gonad)	Hypoplastic (+/-)	Absent	Partial variable virilization	
		Testicular regression	Anorchia				
	Disorders of androgen synthesis or action	HSD17B3 def	Testes	Present		Female, ambiguous, rarely male. In puberty, virilization signs	Low T, N/high AMH, N ACTH[2]. T/Andros <1
		HSD3B2 def				Male with undervirilization	Low T, N/high AMH, high ACTH[2]
		CYP17A1 def		Present (Normal or hypoplastic)		Female or ambiguous	
		LCAH or CYP11A1 def					
		POR def					
		Cyb5 def					
		5α-reductase def		Present (Hypoplasic vagina)		Variable virilization	N T and AMH. T/DHT>30[2]
		AIS		Complete: absent; Partial: hypoplastic		Complete: female; Partial: variable	N/High T, AMH and LH
		LH receptor defects		Present		Ambiguous genitalia to normal female	High LH, low T[1,2]
		PMDS	Testes		Present	Male	N T, N/low AMH

46,XX DSD	**Disorders of ovarian development**	46,XX OT DSD	Dysgenetic (Same gonad)	Variable, normally absent	Variable. Mostly present	Ambiguous, male or female	HyperH, AMH in male range[4]. Higher T levels[2].
		46,XX Testicular DSD	Atropic testes	Present	Absent	Male, sometimes with ambiguous genitalia	HyperH, low T, high FSH and LH.
		Ovarian dysgenesis	Dysgenetic	Absent	Present	Female	
	Disorders of androgen excess	CYP21A2 def	Ovaries	Absent or scarce		Ambiguous genitalia	High 17OHP. SW in the classic form.
		CYP11B1 def					High 17OHP, DOC, 11-desoxicortisol
		POR def				Ambiguous genitalia to normal female	High 17OHP, T, prog, corticosterone
		HSD3B2 def				Female, sometimes scarce virilization	High 17OHP and 17OH-preg. SW syndrome
		Maternal/Fetoplacental				Variable virilization	High T
Sex chromosome		45,XO/46,XY MGD	Dysgenetic, streak or normal	Hypoplastic (Unilateral)	Hypoplastic (Unilateral)	Partial virilization (Variable)	HyperH, T and AMH[4]
		46,XX/46,XY mosaicism	Testis and ovary (same gonad/two gonads	Variable	Variable		HyperH, T and AMH[4]
		Turner syndr[3]	Dysgenetic	Absent	Present	Female	HyperH, depending on the gonadal dysgenesis
		Klinefelter syndr[3]	Small testes	Present	Absent	Male	HyperH, higher FSH than LH.
		47,XYY karyotype	Testes				Normal testicular function

ACTH, adrenocorticotropic hormone; AIS, Androgen insensitivity síndrome; AMH, anti-Müllerian hormone; Andros, androstenedione; Cyb5, cytochrome b5; CYP11A1, 20-22 desmolasa; CYP11B1, 11-Beta hydroxilase; CYP17A1, 17α-hydroxylase and 17,20-lyase; CYP21A2, 21-hydroxilase; DHT, dihydrotestosterone; Def, deficiency; DOC, 11-desoxicorticoesterona; E, oestradiol; HSD3B2, 3-Beta Hydroxysteroid dehydrogenase; HSD17B3, 17-Beta Hydroxysteroid dehydrogenase; HyperH, hypergonadotropic hypogonadism; FSH, Follicle Stimulating Hormone; LH, Luteinizing hormone; MGD, mixed gonadal dysgenesis; N, normal; OT, ovotesticular; PMDS, Persisten Müllerian Duct defect; POR, Cytochrome P450 Oxidoreductase; Prog, progesterone; Syndr, síndrome; SW, salt wasting; T, testosterone; 17OH-preg, 17-hydroxipregnenolone; 17OHP, 17-hydroxiprogesterone; [1]Basal; [2]After stimulation test; [3]And variants; [4]Depends on the presence of testicular tissue.

Supplementary data 2. Clinical data of the patients included in the genetic analysis.

Clinical data							
Case	**Karyotype /assigned gender**	**Suspected clinical diagnosis**	**First consultation**	**Clinical phenotype of the patients and relatives**			
				Clinical phenotype	**Müllerian ducts**	**Gonadectomy (age)**	**Histology**
GN0001	46,XX/F	46,XX DSD GD	Primary amenorrhea and delayed puberty	17y, primary amenorrhea, pubertal delay, lack of breast development. US: absence of ovaries and uterus. Begins treatment with estrogens. 24y, normal feminization, lack of menarche. US: absence of ovaries and uterus. **Other features:** Sister presenting same phenotype, after treatment lack of breast development	No	No	
GN0003	46,XX/F	MRKH II syndrome	Vaginal agenesis	8y, vaginal and uterine agenesis, confirmed by laparoscopy. 15y, MRI: left renal agenesis. External normal genitalia.	No	No	
GN0004	46,XX/M	46,XX Testicular DSD	Bilateral gynecomastia	13y, bilateral gynecomastia, unilateral cryptorchidism. 26y, azoospermia.	No	No	
GN0007	46,XY/F	46,XY DSD CGD	Suspected gonadal dysgenesis	16y, female external genitalia with normal vagina and small cervix. Non-palpable gonads. 17y, gonadectomy: hypoplastic uterus, streak gonads (4cm). **Other features:** Sister with possible 46,XY DSD GD.	Yes	Yes (17y)	Ovarian stroma with Sertoli and Granulosa cells surrounded by Leydig's. Unilateral gonadoblastoma and paraovarian cyst
GN0009	46,XY/F	Frasier syndrome	Delayed puberty	12y, headaches, vomiting and fatigue, high blood pressure. Fundoscopy: bilateral papilledema. US: glomerulopathy. Biopsy: confirmed the nephronophtisis. Began treatment. 13y, hemodialysis. Renal transplant but kidney rejection. 17y, infantile external genitalia without breast development and pubarche. US: normal uterus, non-visible gonads. GH treatment was prescribed. 24y, gonadectomy.	Yes	Yes (24y)	Streak gonads. Gonadoblastoma
GN0011	46,XY/F	46,XY DSD CGD	Absence of menarche and thelarche	16y, absence of menarche and thelarche. Normal female external genitalia and hypergonadotropic hypogonadism. US: non-visible uterus, normal vagina and bilateral gonads. Laparoscopy: hypoplastic uterus and streak gonads. Gonadectomy. **Other features:** Father and brother with delayed puberty.	Yes	Yes (16y)	Gonadoblastoma (left), Wolff remnants (right)

GN0012	46,XY/M	46,XY DSD Gonadal regression	Anorchia	1y, no response to hCG stimulation test. Laparoscopy: absence of female gonads and testicular tissue.7-8y, testicular prostheses (5cc).11y, begins treatment with testosterone. 15y, mature penis and scrotum, pubarche IV, facial and body hair. Adrenal insufficiency discarded. 17y, changed treatment to testogel. 19y, well virilized. 26y, increased treatment with testosterone.	No	No	Gonadal tissue without differentiation.
GN0013	46,XY/F	46,XY DSD CGD	Primary amenorrhea and delayed puberty	18y, primary amenorrhea and delayed puberty. Female external genitalia. US: vagina (7cm), streak left gonad. Laparoscopy: rudimentary uterus, large fallopian tubes and streak gonads. **Other features:** marfanoid habitus.	Yes	Yes (18y)	Normal fallopian tubes, cortical stroma without oocites and primordial follicules, mesonefric structures.
GN0014	46,XY/F	46,XY DSD PGD	Lack of secondary sex characteristics	13y, delayed development of secondary sexual characteristics. US: vagina and small uterus in the right side (3.8x1.4x1.8cm), small gonad (1.7cm) next to bladder. CT scan: rudimentary uterus after bladder, two rounded images suggesting vas deferens previous to inguinal canals, no image corresponded to inguinal testes. Gonadectomy of streak gonads. **Other features:** sister presenting with primary amenorrhea and female external genitalia (16y). US: small hypoplastic uterus, well defined ovaries. 46,XY. Laparoscopy: right streak gonad and left gonad (3.5cm), hypoplastic uterus and no wolff remnants. Histology: Right streak gonad with ovarian stroma, Sertoli and Leydig cells; left gonad: disgerminoma	Yes	Yes (13y)	Streak gonads. Ovarian stroma and seminiferous tubules with Sertoli cells. Absence of germ cells. Presence of Leydig cells.
GN0017	46,XY/M	46,XY DSD Gonadal regression	Anorchia	At birth micropenis, bilateral cryptorchidism and hypoplastic scrotum. 12mo, micropenis, non-palpable testes, atrophic scrotum and hypergonadotropic hypogonadism. No response of testosterone to hCG stimulation test. Laparoscopy. **Other feature**s: psychomotor retardation, parental consanguinity.	No	No	Atrophic seminiferous tubules
GN0018	46,XY/F	46,XY DSD PGD	Bilateral inguinal hernia	5y, surgery for bilateral inguinal hernia. US: non-visualized uterus and ovaries. 6y, gonadectomy	No	Yes (6y)	Immature testicular tissue
GN0020	46,XY/F	46,XY DSD PGD	Short stature and delayed puberty	13y, pubertal development delay and little pubic hair. US: absence of gonads. 17y, undeveloped breast, normal menses on treatment. **Other features:** short stature.	ND	Yes (ND)	ND
GN0023	46,XX/M	46,XX	Bilateral	2y, bilateral cryptorchidism. Inguinal gonads that descend after hCG	No	No	

		Testicular DSD	cryptorchidism	stimulation.			
GN0024	46,XY/F	46,XY DSD	Anorchia	At birth, female external genitalia and bilateral inguinal hernia. US: inguinal testes, no uterus or ovaries.	No	No	
GN0025	46,XY/F	46,XY DSD CGD	Primary amenorrhea	17y, primary amenorrhea, undeveloped external genitalia and hypergonadotropic hypogonadism. US: hypoplastic uterus (35x7mm) and no gonadal remnants. Gonadectomy: atrophic uterus, streak gonads. 26y, US: uterus 36x22x17mm. **Other features:** Sister with CGD. Family history positive for short stature (mother 149 cm, father 162.6 cm).	Yes	Yes (18y)	Streak gonads, bilateral gonadoblastoma
GN0026	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, US: short vagina, images suggesting inguinal gonads (10x6mm), no uterus. **Other features:** disorder of intermediary metabolism.	No	No	
GN0027	46,XY/F	46,XY DSD CGD	Primary amenorrhea	15y, primary amenorrhea, female normal external genitalia and delayed puberty. US: Small gonads, uterus and cervix. **Other features:** sister presenting with MGD.	Yes	No	
GN0028	46,XY/M	46,XY DSD	Ambiguous genitalia	At birth curved micropenis with scrotal hypospadias and bilateral cryptorchidism. Surgery. 12y, penis (7.6cm and 10cm diameter), testes (5ml), pubarche and axilarche V and acne. US: Right testis of 15x15x28mm, vol 3.2cc, and left testis 14x15x35mm, vol 3.9cc, small cysts in right epididymis 6.4mm. From 12y 10 15y, phenotype remained without significant changes. 15y, US: Right testis of 12x17x28mm, vol 3.2cc, and left testis 9x21x31mm, vol 3.3cc, small cysts in both epididymis 6.4mm and 2.7mm. **Other features:** 13y, Fragile X syndrome study, *FMR1* gene was normal.15y, Asperger syndrome. Mother and brother with unilateral renal agenesis, brother also presenting with vas deferens and epididymis agenesis.	No	No	
GN0029	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, bilateral cryptorchidism, hypospadias and horseshoe kidney. Orchidopexy was done without good results. 15y, atrophic right testis, undescended left testis and curved penis. **Other features:** Horseshoe kidney	No	No	
GN0031	46,XX/F	46,XX DSD GD	Ambiguous genitalia	16d, hyperpigmented genitalia, hypertrophic labia minora and bilateral inguinal hernia with palpable masses. US: bilateral inguinal hernia with neonatal ovaries. Surgery was planned.	Yes	No	

GN0033	46,XY/M	46,XY DSD PGD	Micropenis	3mo, micropenis and cryptorchidism. 12y, small penis (4cm) and testes (2ml).	No	No	
GN0034	46,XY/F	46,XY AIS	Bilateral inguinal hernia	8mo, normal female external genitalia, bilateral inguinal hernia, absence of upper vagina and uterus. Surgery. 4y, US: absence of uterus and gonads. 13y, gonadectomy. Begins treatment with estradiol. 15y, thelarche IV and pubarche V.	No	Yes (13y)	Testicular tissue
GN0035	46,XY/F	46,XY DSD CAIS	Primary amenorrhea	13y, primary amenorrhea, hypertrophic clitoris, bilateral inguinal gonads, breast development begins. 16y, gonadectomy. 22y, vaginoplasty. 42y, hypertrophic clitoris (1.5-2cm). **Other features:** Obesity (BMI 39)	No	Yes (16y)	ND
GN0037	46,XY/M	46,XY DSD PAIS	Ambiguous genitalia	At birth, micropenis with perineal hypospadias, cryptorchidism and palpable gonads at labia majora.	Yes	No	Normal testicular tissue
GN0038	46,XY/F	46,XY DSD Androgen biosynthesis defect	Voice change and variations at external genitalia	At birth, female external genitalia. 10y, change in voice. 11y, hypertrophy of the clitoris, vagina (2cm) palpable gonads in inguinal canal and hirsutism. MRI: absence of female genitalia, possible gonads in inguinal canal. Laparoscopy: absence of female genitalia. Clitoroplasty and gonadectomy.	No	Yes (12y)	Testicular parenchyma with Sertoli and preSertoli cels. Basal spermatogoonias. Leydig mature cells (both).
GN0039	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, scrotal gonads, scrotal hypospadias and bilateral inguinal hernia. US: Gonads with epididymis suggesting testes (right: 1.2x0.6x0.6; left testis: 1.2x0.7x0.8). **Other features:** Heart murmur	No	No	
GN0041	46,XY/F	46,XY DSD CAIS	Suspicion of AIS	9y, female external genitalia, thelarche and pubarche Tanner I. 12y, US: absence of internal genitalia. **Other features:** Older sister presenting the same phenotype	No	Yes (3y)	ND
GN0042	46,XY/F	46,XY DSD AIS	Bilateral inguinal masses, clitoromegaly	11y, bilateral inguinal hernia, palpable gonads (2cc, 3cc), clitoromegaly and pubarche. Growth rate increase. US: vaginal pouch and non-visible uterus. Clitoroplasty. 20y, good breast development after estrogen treatment. 25y, US: prepubertal hypoplastic uterus, no ovaries.	Yes	Yes (10y)	Hypoplastic testes
GN0043	46,XY/M	46,XY DSD PAIS	Ambiguous genitalia	At birth (33w +3), micropenis with hypospadias, bilateral cryptorchidism and non-palpable testes in inguinal canal. Bilateral inguinal hernia. US: testes in inguinal canal (8mm), no uterus.	No	No	

				Other features: Respiratory distress, anemia, hyperlactiacidemia			
GN0046	46,XY/F	46,XY DSD AIS	Vaginal agenesis	At birth, bilateral inguinal hernia and vaginal agenesis. US: images corresponding to testes. 16mo, bilateral gonadectomy. 10y, begins estrogen treatment. 15y, normal female external genitalia, vagina of 3cm. 16y, vaginoplasty. 17y, thelarche V. **Other features:** Obesity.	No	Yes (16m)	Infantile testes, vas deferens, no interstitial cells.
GN0050	46,XY/F	46,XY DSD CGD	Turner phenotype	3y, presenting with some slight Turner features and normal female external genitalia. Laparoscopy: streak right gonad suspecting dysgenetic testes, normal fallopian tube and small uterus. **Other features:** Low hairline, auditory pavilions with low implantation	Yes	Yes (3y)	ND
GN0051	46,XX/M	46,XX DSD	Ambiguous genitalia	6y, micropenis with scrotal hypospadias, bilateral cryptorchidism, palpable right testis in inguinal canal and bifid scrotum. US: abdominal left testis (9.5x3.8mm) and right in inguinal canal, no Müllerian structures were visualized. Orchidopexy and biopsy. Congenital adrenal hyperplasia was discarded. **Other features:** Patient from Mauritania.	Yes	No	Infantile ovary with follicular cysts, fallopian tube, atrophic uterus and mesonefric remnant.
GN0054	46,XX/M	46,XX Testicular DSD	Weakness and lack of body hair	41y, weakness, absence of body hair, hypergonadotropic hypogonadism. Begins treatment with testosterone.	No	No	
GN0055	46,XY/F	46,XY DSD CAIS	Right inguinal hernia	At birth, right inguinal hernia. Gonadectomy. 13y, US: non-visible uterus, vaginal pouch (2cm), 2 possible testes (17.8 and 13.8 mm).14y, gonadectomy.	No	Yes (1m; 14y)	Testicular dysgenetic tissue, seminiferous tubules with plenty Sertoli cels. Leydig cell hyperplasia.
GN0056	46,XY/M	Denys-Drash syndrome	Ambiguous genitalia, Wilms' tumour	1mo, micropenis with perineal hypospadias, cryptorchidism and Prader III. Laparoscopy: intra-abdominal small gonads, vas deferens and prostatic utricle are visualized. Wilms' tumor.	No	No	
GN0059	46,XY/F	46,XY DSD CGD	Ambiguous genitalia	At birth, micropenis, non-palpable testes and atrophic scrotum. US: normal vagina, non-visible gonads and uterus. 7d, US: non-visible gonads and uterus. 1mo, hCG test with no increase of testosterone. 6mo, laparoscopy. 18mo, feminizing genitoplasty.	Yes	Yes (6m)	Epididimus and fallopian tube is identified, germ cells.
GN0064	47,XY/F	46,XY DSD	Discordant	3mo, US: normal vagina, uterus and gonads.	Yes	No	

		CGD	karyotype	**Other features:** Down syndrome			
GN0066	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, unilateral cryptorchidism, bifid scrotum and scrotal hypospadias. US: inguinal left testis (9.2x5.5x5.1mm). 7y, after surgery and treatment, presents a 3cm penis. Biochemical analysis is normal. 8y, precocious adrenarche. **Other features:** 8y, hearing loss	No	No	
GN0068	46,XY/M	46,XY DSD LH RD	Precocious puberty	3y, precocious puberty and family background.	ND	No	
GN0070	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	9y, penis (2 cm) with scrotal hypospadias and bilateral inguinal testes. US: Images corresponding to testes in inguinal canal (right: 15.6x8.7x6.7mm, 0.5cc; left 17x8.5x8, 0.6cc). **Other features:** Patient from Madagascar.	No	No	
GN0075	46,XY/M	46,XY DSD	Ambiguous genitalia	At birth, micropenis with scrotal hypospadias, undescended testes and bifid scrotum. 5y, normal penis (2.5cm) after testosterone treatment, scrotal right testis (2ml), and hydrocele in left testis. 6y, smaller penis and both testes in scrotum. 7y, small penis (2.5cm). **Other features:** Mother and aunt with precocious menopause and uncle with scrotal hypospadias.	No	No	
GN0076	46,XY/F	46,XY DSD CAIS	Unilateral mass in labia majora	8d, palpable gonad in labia majora. US: 2 visible gonads similar to testes, no uterus.	No	No	
GN0078	46,XY/M	Adrenal Insufficiency	Adrenal Insufficiency	Familial background. 45d, scan: Hypoplastic adrenal glands. **Other features:** Glycemia NA; Na 107mEq/L; K 6mEq/L; Aldosterone 14ng/dL; PRA 66ng/mL/H. Cousin presenting with adrenal insufficicency.	-	-	-
GN0080	46,XY/F	46,XY DSD CAIS	Suspected AIS	5y, gonadectomy. 6y, US: Absence of uterus and fallopian tubes	No	Yes (5y)	Male gonads
GN0084	46,XX/F	MRKH II syndrome	Primary amenorrhea	16y, primary amenorrhea, MRKH type II phenotype, uterine, vaginal and kidney agenesis, ectopic ovaries, hyperandrogenism. 18y, US: left normal gonad, absence of right gonad, uterus and fallopian tubes. MRI: unique pelvic kidney (11x74mm), visible ovaries in pelvic region (60x25mm and 43x22mm). No hearing loss. **Other features:** kidney agenesis.	No	No	
GN0088	46,XY/M	46,XY DSD	Precocious	3y, pubertal penis and testes growth, pubarche. 3y, testes 6mL. 4y, testes 8cc	No	No	

		LH RD	puberty	and begins treatment. 5y, pubarche is diminishing, testes 8mL and penis 25x60mm.			
GN0090	46,XX/M	46,XX OT DSD	Ambiguous genitalia	6y, curved penis (2cm) with scrotal hypospadias, bifid scrotum, right inguinal hernia, right gonad (<1ml) in inguinal canal. US: absence of female internal genitalia, right gonad (10.7mm) in inguinal canal and intra-abdominal left gonad (9.7mm). Laparoscopy: bilateral ovotestes located in inguinal canal and in iliac fosse, attached by tubes and atrophic uterus. Absence of lower portion of uterus or vagina. **Other features:** Patient from Mauritania.	No	No	Bilateral ovotestes
GN0091	46,XY/M	Adrenal Insufficiency	Respiratory distress	At birth, respiratory distress, hyperpigmentation, hypoglycemia. Testes 1ml and penis 3.6cm. MRI: non-visualized adrenals. **Other features:** Glycemia 39mg/dL; Na 126mEq/L; K 6.2mEq/L; Aldosterone 49ng/dL; PRA >21.3ng/mL/H.	-	-	-
GN0096	46,XX/M	46,XX OT DSD	Ambiguous genitalia	3y, rudimentary penis (<0.5cm), palpable right gonad (0.5-1ml) in scrotum, non-palpable left gonad and pubarche I. US: inguinal bilateral gonads (1ml), absence of female internal genitalia. 4y, bilateral orchidopexy and laparoscopy. **Other features:** Patient from Mauritania	No	No	Ovarian tissue (left), testicular and ovarian tissue (right)
GN0100	46,XY/F	46,XY DSD CGD	Primary amenorrhea	16y, primary amenorrhea and lack of secondary sex characteristics. US: structure at the vaginal location, non-visible uterus and ovaries. MRI: visible right gonad (1cm), no uterus. 17y, normal external genitalia. Laparoscopy: streak gonads and absence of uterus. Begins treatment with estradiol.	Yes	Yes (17y)	Fallopian tubes and ovarian tissue. No testicular tissue.
GN0101	46,XY/M	Adrenal Insufficiency	Suspected primary renal insuficiency	3y, primary renal insuficiency. Hypoglycemia, vomiting, asthenia. 17y, testes 4mL and penis 8cm. **Other features:** Glycemia 28mg/dL; Na 119mEq/L; K 6.8mEq/L; Aldosterone 11ng/dL; PRA NA.	-	-	-
GN0103	46,XY/M	46,XY DSD	Ambiguous genitalia	At birth, hypoplastic penis and bilateral cryptorchidism. US: inguinal testes. 6mo, after testosterone treatment penis increased to 5cm and 1.2cm in diameter. US: inguinal both testes. Presents pericentric inversion of chr. 9 (inv[9][p11q13]).	No	No	
GN0104	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, micropenis buried in fat with scrotal hypospadias, bilateral cryptorchidism and well developed scrotal sacks. Palpable testes in inguinal canal, confirmed by US. 22mo, surgery for hypospadias.	No	No	

GN0108	46,XY/F	46,XY DSD CAIS	Discordant karyotype	At birth, female external genitalia. 3y, US: absence of uterus, no gonads in inguinal canal.12y, US: absence of uterus. 13y, MRI: blind short vagina (14.2mm), absence of uterus or annexes, near to left iliac vasses presence of soft structure that could correspond to undescended teste, no structures in inguinal region.	No	No	
GN0109	46,XY/F	46,XY DSD GD	Virilization at puberty	14y, primary amenorrhea, previously virilized (8-9y) and increased body and face hair. Erectile hypertrophic organ with perineal urethra and scrotal rests. US: rudimentary uterus and two images in left inguinal canal suggesting atrophic testes. MRI: endocervical cavity, next to another cystic cavity corresponding to cervix without ovaries. 15y, orchidopexy. Begins treatment.	Yes	Yes (15y)	Testes with Sertoli cells and epididymis.
GN0111	46,XY/F	46,XY DSD CGD	Primary amenorrhea, obesity	14y, primary amenorrhea, stenotic and enlarged vagina. US: non-visible uterus and gonads. MRI: rudimentary uterus and vagina, ovaries not identified. 15y, Laparoscopy: rudimentary uterus, streak gonads. **Other features:** Obesity.	Yes	Yes (15y)	Testicular parenchyma, normal fallopian tubes
GN0112	46,XY/F	46,XY OT DSD	Discordant karyotype	Male karyotype in amniocentesis. At birth, female external genitalia and bilateral inguinal gonads. 1d, US: hypoplastic uterus (2cm), normal vagina and inguinal gonads (right: 1x0.8x0.4c, left: 1.1x0.5x0.6cm).	Yes	No	
GN0114	46,XY/F	46,XY DSD CGD	Suspected GD	Absence of uterus and streak gonads. **Other features:** Niece presenting with same phenotype	No	Yes (17y)	Absence of germ cells
GN0118	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	2y, curved penis with scrotal hypospadias, normal testes, rectal malformation. MRI: absence of uterus and ovaries. **Other features:** Anal agenesis, iron deficiency. Patient from Mauritania	No	No	
GN0119	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	6y, curved penis (4.4-4.5cm) with scrotal hypospadias, cryptorchidism and bifid scrotum. US: Right testis (15x9mm) in scrotum, left testis (13x6mm) in inguinal canal. **Other features:** Patient from Equatorial Guinea.	No	No	
GN0122	46,XY/M	46,XY DSD Androgen biosynthesis defect	Scrotal hypospadias	2y, curved penis (3.1cm) with scrotal hypospadias and testes (0,5-1ml). Peno-scrotal transposition. 5y, penis 5cm, and testes volume 0.5ml.	No	No	
GN0123	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, micropenis with perineal hypospadias, palpable gonads and atrophic bifid scrotum. 2y, laparoscopy: absence of Müllerian ducts. 10y, small penis (2.6cm), right testis in scrotum and left in inguinal canal. 14y, penis (5.5cm)	No	No	2y, normal testicular tissue. 10y, absence of germ cells and

				with both testes in scrotum (2-3cc). 17y, biopsy. 38y, penis 3-4cm, testes 2cc. **Other features:** Obesity.			Leydig cell hypoplasia.
GN0124	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, curved penis (2.8cm) with scrotal hypospadias, palpable inguinal gonads and bifid scrotum. US: Left inguinal hernia, inguinal testes (left: 1.6x1cm; right: 1.5x0.7). Absence of Müllerian ducts. **Other features:** Horseshoe Kidney.	No	No	
GN0125	46,XY/F	46,XY DSD CAIS	Primary amenorrhea	17y, primary amenorrhea, vaginal pouch, absence of uterus and annexes, confirmed by US. Patient refers that she presented bilateral inguinal hernia at birth, corrected by surgery (5y). 20y, laparoscopy: testicular gonads, no uterus and fallopian tubes	No	Yes (20y)	Testicular tissue
GN0132	46,XY/F	Frasier syndrome	Suspected GD, gonadoblastoma and renal insuficiency	18y, gonadal dysgenesis and renal insufficiency. 19y, laparoscopy: Müllerian ducts. **Other features:** Renal insufficiency, obesity	Yes	Yes (19y)	Gonadoblastoma
GN0133	46,XX/M	46,XX T DSD	Incongruent karyotype	39y, male external phenotype with gynecomastia, unilateral cryptorchidism, azoospermia and hypergonadotropic hypogonadism. US: Right testis (30mm) with cyst (5.5mm), inguinal left testis.	No	No	
GN0138	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	At birth, micropenis (30mm and 40mm of diameter) with scrotal hypospadias and undescended testes. US: visible uterus (3.5cm), non-visible gonads.	Yes	No	
GN0139	46,XY/F	46,XY DSD CAIS	Primary amenorrhea	16y, primary amenorrhea, female external genitalia and vaginal pouch (2cm). Bilateral inguinal hernia at 2 months of age. US: non-visible right gonad, left gonad (1cm).	No	Yes (15y)	Testicular tissue
GN0141	46,XY/F	46,XY DSD CGD	Suspected gonadal dysgenesis	15y, gonadal exeresis: streak gonad with tumoral aspect and gonadoblastoma. Müllerian remnants.	Yes	Yes (15y)	Gonadoblastoma
GN0142	46,XY/F	46,XY DSD CGD	Incongruent karyotype	At birth, female external genitalia, clitoris 1-2cm with urinary meatus on base, non-palpable gonads in inguinal canal and undeveloped bifid scrotum. US: Absence of uterus and gonads. Gonadectomy is planned. **Other features:** Cousin diagnosed with 46,XY DSD gonadal dysgenesis (15y)	No	No	
GN0144	46,XY/M	46,XY DSD PGD	Hypospadias	At birth, male external genitalia, hypospadias and scrotal testes. 2y, surgery for hypospadias. 5y, small testes and unilateral cryptorchidism. US: left testis in scrotum and right testis in inguinal canal. **Other features:** Hypothyroidism.	No	No	

GN0145	46,XY/M	46,XY DSD PAIS	Sexual reassignment	20y, small penis (3.5cm) and testes (20cc). Request for sexual reassignment. **Other features:** 20y, identified as female since childhood.	No	No	
GN0146	46,XY/F	46,XY DSD PAIS	Suspected AIS	10y, gonadectomy. 14y, US: no uterus and gonads. 19y, hypertrophic erectile organ, partial labial fusión and non-palpable masses.	No	Yes (10y)	Testicular tissue
GN0147	46,XY/F	46,XY DSD AIS	Primary amenorrhea	7y, right inguinal hernia. At puberty, primary amenorrhea, clinical history from Algeria reporting ovarian and uterine agenesis, genitalia not described. 13-14y, thelarche. 24y, left inguinal hernia. 30y, in Melilla, hypergonadotropic hypogonadism with testosterone 0ng/ml after gonadectomies, never had treatment. US: vaginal outline, non-visible uterus and annexes. 35y, female external genitalia, no facial hair, thelarche II-III, erectile organ is buried, one opening for urethra and vagina. **Other features:** Patient from Algeria.	No	Yes (7 and 24y)	7y, normal testes. 24y, normal testes
GN0148	46,XY/F	46,XY DSD CGD	Primary amenorrhea	3y and 5y, surgery for bilateral inguinal hernia. 14y, primary amenorrhea. 17y, pubarche. 20y, female external genitalia, rudimentary cervix. US: hypoplastic uterus (2.5cm), non-visible gonads, streak annexes. Confirmed by laparoscopy. 25y, refers menses under treatment. **Other features:** Father with inguinal hernia (not intervened), sister with intervened bilateral inguinal hernias	Yes	No	Normal uterine tubes
GN0150	46,XY/F	46,XY DSD AIS	Bilateral inguinal hernia	3y, female external genitalia with bilateral inguinal hernia. Gonadectomy. US: non-visible gonads and uterus. 9y, US: small uterus (7x10x24mm) and normal kidneys.	Yes	Yes (3y)	Immature testes presenting with Sertoli cells and spermatogonias
GN0151	46,XX/F	46,XX DSD GD	Suspected GD	16y, hypergonadotropic hypogonadism. US: hypoplastic uterus and annexes (17 and 18mm). 21y, US: uterus (57x32x22) and non-visible annexes. Bilateral mammoplastia.	Yes	No	
GN0152	46,XY/F	46,XY DSD PGD	Ambiguous genitalia	5mo, scrotal hypospadias, unilateral cryptorchidism and a vaginal pouch. Laparoscopy: Right testis in inguinal canal, no Müllerian ducts.19mo, micropenis with scrotal hypospadias, normal labia majora without minor labia and bilateral inguinal gonads. US: right gonad (0.4cc), left gonad (0.3cc). **Other features:** Ventricular hypertrophy, at 16m normal. Growth delay, hypothyroidism, delayed psychomotor development with macrocephaly. 2y, Fulminant sepsis of digestive origin. Born in Rumania.	No	No	Prepubertal gonad with seminiferous tubes, Sertoli cells and few spermatogonias.
GN0153	46,XY/M	Adrenal Insufficiency	Hypoaldosteronism	8mo, hypoaldosteronism, hyponatremic dehydration and hyperkalemia. Begins treatment with hydrocortisone until analytical results remain normal.	-	-	-

				Actually under fludrocortisone treatment.			
GN0154	46,XX/F	46,XX DSD GD	Primary amenorrhea	15y, primary amenorrhea, female external genitalia, thelarche II-III and hypergonadotropic hypogonadism. US: small uterus, non-visible gonads. Laparoscopy: small uterus, fallopian tubes and 2 small ovaries. **Other features:** Childish habitus, short neck, lots of nevus	Yes	No	Ovarian structure with immature primordial follicles and follicular atresia
GN0155	46,XY/F	46,XY DSD AIS	Ambiguous genitalia	At birth, clitoral hypertrophy. 2y, clitoroplasty and gonadectomy. 7y, left inguinal hernia. At puberty, virilization and male role. 20y, gender reassignment. 37y, male phenotype, low androgens levels and begins treatment. Scan: uterus remnants, other structure like corpus cavernosum in the perineum. **Other features:** 37y, obesity.	No	Yes (2y)	Testicular disgenetic gonad
GN0156	46,XY/M	46,XY DSD LH RD	Progressive increase in genital development	7y, development of genitalia. Penis 8cm length and adult diameter, testes 10mL, pubarche II. **Other features:** Pituitary adenoma, hearing loss. Indian origin.	No	No	
GN0157	46,XY/M	46,XY DSD LH RD	Adrenarche	6y, precocious puberty, genital development P2, testicular volume 2-3ml. 7y, absence of other pubertal signs, pubarche with normal testes, changes in body odor, acne. US: normal. 9y, genital development T3, testicular volume 3ml. US: normal testes located in scrotum. **Other features:** Patient from Equatorial Guinea.	No	No	
GN0158	46,XX/M	46,XX OT DSD	Ambiguous genitalia	3y, curved penis with scrotal hypospadias, bilateral cryptorchidism and atrophic scrotum. Non-palpable testes. US: prepubertal uterus, images suggesting ovaries. Biopsy. **Other features:** Patient from Sierra Leone.	No	No	Ovarian tissue (left), testicular and ovarian tissue (right)
GN0159	46,XX/M	46,XX Testicular DSD	Ambiguous genitalia	At birth, curved penis with scrotal hypospadias and unilateral cryptorchidism. FISH: Confirmed *SRY* translocation. 3mo, normal penis and testes (2ml). 9y, augmented penis (6.3cm circumference x3.8cm length), normal right testis and left in inguinal canal. 11y, surgery for hypospadias. 12y, US: right testis (22.7x7.3mm), left in inguinal canal (24x7.5mm). 13y, left orchidopexy. Right testis 3ml and left 5ml in size, lack of pubarche and axilarche. **Other features:** 3mo, Smith Lemli Opitz syndrome suspected. Psychomotor delay. 11y, cafe au lait spots, microcephaly, corpus callosum agenesia.	No	No	
GN0160	46,XX/M	46,XX DSD	Discordant	1y, male external genitalia, testes in scrotum.	No	No	

			karyotype	**Other features:** 5mo, occipital plagiocephaly, delayed motor development, bilateral congenital cataracts.			
GN0162	46,XY/M	Scrotal hypospadias	Scrotal hypospadias	16mo, curved penis (3.8cm) with scrotal hypospadias and normal testes. **Other features:** Renal litiasis.	No	No	
GN0163	46,XY/M	46,XY DSD Androgen biosynthesis defect	Ambiguous genitalia	11mo, distal hypospadias, surgery. 10y, micropenis (3.6cm) buried in fat, testes (3mL) in scrotum. After testosterone (for 19 days), right testis (2-3ml), left testis (4ml), penis (5.2cm length and 6cm diameter), no pubarche. After treatment (for 1 month), testes 5ml, penis 5cm and 6.4 cm diameter, pubarche III. 12y, penis (5.2cm and normal diameter) buried in fat, testes 10ml and 12ml, pubarche III. 13y, penis (9cm and 8.9cm diameter), testes>25ml, pubarcheIII-IV. 14y, US: intraescrotal testes with normal size and morphology. Surgery for hypospadias 5 times.	No	No	
GN0164	46,XY/M	46,XY DSD LH RD	Bilateral gynecomastia	36y, bilateral gynecomastia, micropenis, small testes (3mm) and hypergonadotropic hypogonadism.	No	No	
GN0167	46,XY/M	46,XY DSD Androgen biosynthesis defect	Hypospadias	At birth, normal penis with distal hypospadias and testes in scrotum (1-2mL).	No	No	
GN0169	46,XY/M	46,XY DSD PGD	Hypospadias	At birth, curved penis with scrotal hypospadias, undescended testes, bilateral inguinal hernia and bifid scrotum. Orchidopexy. 8y, US: testes (0.7ml and 0.3ml) in scrotum.	No	No	
GN0171	46,XY/M	46,XY DSD	Ambiguous genitalia	At birth, micropenis. 11y, micropenis (3.5 cm), non-palpable testes. MRI: testes in inguinal canal. 14y, begins increasing doses of testosterone therapy. 18y, azoospermia. 21y, penis (8cm) buried in subcutaneous fat, left testis in scrotum (0.5ml). US: ovoid structure (17x16mm) in right inguinal canal, corresponding to atrophic testes. Laparoscopy and biopsy are planned. **Other features:** 11y, Obesity (BMI 36.6, +6.4SD) and short stature. 21y, obesity (BMI 44.9,+6.8SD), short stature (156.1 cm),	No	No	
GN0173	46,XY/M	46,XY DSD gonadal regression	Ambiguous genitalia	At birth, micropenis, hypoplastic scrotum with non-palpable gonads. 3mo, micropenis (1cm), developed scrotum. US: non-visible gonads. Begins treatment with testosterone. 5mo, penis (2.3cm). Continues treatment with testosterone. 9mo, penis (3cm). 1y, MRI: absence of gonads. 16mo, penis (4cm) and developed scrotum.	No	No	

GN0174	46,XY/M	46,XY DSD Androgen biosynthesis defect	Scrotal hypospadias	2mo, curved micropenis (1,8-2cm) with scrotal hypospadias and palpable small testes (0,5ml) in scrotum. **Other features:** Hypothyroidism.	No	No	
GN0175	46,XX/F	46,XX DSD	Secondary amenorrhea	16y, complete pubertal development with secondary amenorrhea. Patient refers menarche at age 13. Normal internal genitalia. US: normal uterus (5,3cm), gonads (9-10cm). **Other features:** Ventricular septal defect.	Yes	No	
GN0176	46,XY/M	46,XY DSD Androgen biosynthesis defect	Ambiguous genitalia	At birth, hypospadias, palpable scrotal testes and bifid scrotum. US: non-visible Müllerian ducts.	No	No	
GN0177	46,XY/F	46,XY DSD CAIS	Discordant karyotype	Male karyotype in amniocentesis. At birth, female external genitalia. 6d, US: bilateral inguinal gonads corresponding to testes. 13d, biopsy. **Other features:** Brother affected with phenylketonuria, she is a carrier of the mutation. Maternal aunt with 46, XY karyotype. Patient from Morocco.	Yes	No	Germinal and immature Sertoli cells, no Leydig cells (both)
GN0178	46,XY/M	46,XY DSD PGD	Micropenis	3mo, micropenis, good response to treatment. 6y, penis 2,5cm and normal diameter. 9y, normal penis, testes (2ml) and hypoplastic scrotum.10y, smaller penis (2cm), testes (2ml) and hypoplastic scrotum. **Other features:** 8y, Autoimmune thyroiditis.	No	No	
GN0179	46,XX/F	46,XX DSD GD	Suspected GD	18y, hypergonadotropic hypogonadism, atrophic uterus and anexial atresia.	Yes	No	
GN0182	46,XY/M	Unilateral cryptorchidism	Ambiguous genitalia	At birth, left scrotal hypoplasia, bifid scrotum and left renal agenesis. 4y, orchidopexy. 11y, male phenotype, penis 3.5-4cm, testes (right 4ml, left 2ml), bifid scrotum, bilateral gynecomastia and pubarche I. **Other features:** Lipomeningocele. Chinese origin.	No	No	
GN0183	46,XY/M	Cryptorchidism	Absence of developed genitalia	Undescended right testis. 12y, surgery. 14y, hormonal treatment to induce puberty. 20y, testes of 10 and 12ml in scrotum.21y, testes of 12ml in size, normal penis. **Other features:** Delay of psychomotor development, mental retardation, lack of GH. Father with schizophrenia.	No	No	
GN0185	46,XY/F	46,XY DSD GD	Ambiguous genitalia	At birth, penis (1.5cm) with hypospadias, two individualized labia majora with blind vagina and palpable right gonad in inguinal canal. US: Presence of	Yes	Yes (1y)	Streak left gonad (left), dysgenetic

				uterus and vagina, non-visible ovaries. MRI: hypertrophic clitoris, normal vagina, hypoplastic uterus (2x0.8x0.8), non-visible gonads. 1y, laparoscopy: hypoplastic uterus which continues to fallopian tubes, right one reaches testis in inguinal canal and left tube encloses streak abdominal gonad. Vaginoplasty and clitoroplasty. 11y, thelarche MI, pubarche P1.			testis (right).
GN0186	46,XY/F	46,XY DSD GD	Primary amenorrhea, high stature	14y, high stature. 16y, primary amenorrhea, pubarche P4, undeveloped mammals, hypertrophic erectile organ, palpable gonads (2cm) in inguinal canal. MRI: inguinal bilateral testes. **Other features:** High stature. Brother with gynecomastia.	No	No	
GN0187	46,XX/M	46,XX Testicular DSD	Azoospermia	34y, male phenotype, azoospermia. US: small testes (right 17x8mm, left 21x10mm).	No	No	
GN0189	46,XY/F	46,XY DSD CAIS	Presence of masses in labia majora	At birth, female external genitalia, normal clitoris and palpable male gonads in labioscrotal folds. 3mo, US: vaginal pouch, bilateral testes (0.3cc), without uterus and ovaries. MRI, testes in labioscrotal folds (20x9 and 18x8mm).	No	No	
GN0190	46,XX/F	46,XX DSD GD	Primary amenorrhea	14y, primary amenorrhea, infantile female phenotype, thelarche M2, pubarche P1 and hypergonadotropic hypogonadism. Normal vagina. 15y, US: infantile uterus and left streak gonad. MRI: small uterus, normal vagina, no ovaries. **Other features:** Small stature. Aunt of the mother presented precocious menopause (15y).	Yes	No	
GN0191	45,X/46,XY /F	46,XY/X0 MGD	Ambiguous genitalia	At birth, inguinal hernia, labioscrotal fusion, labioescrotal and inguinal gonads. Prader IV. MRI: visualized hypoplastic uterus and upper vagina, absence of ovaries. 1y, gonadectomy, vaginoplasty and clitoroplasty. 11y, pubertal signs.	Yes	Yes (1y)	ND
GN0192	45,X/46,XY /F	46,XY/X0 MGD	Primary amenorrhea, small stature	19y, amenorrhea. MRI: hypoplastic uterus, absence of gonads and lower vagina. **Other features:** 19y, short stature, multiple nevus, Turner phenotype.	Yes	No	
GN0194	46,XY/F	46,XY DSD AIS	Primary amenorrhea	15y, primary amenorrhea. 44y, absence of developed breast, pubic and axillary hair. Patient refers that at the age of 30y, doctor said she had infantile uterus and ovaries. **Other features:** Abdominal obesity. Sister (30y) with primary amenorrhea, absence of development breast and absence of pubic and axillary hair.	Yes	No	

GN0195	46,XY/M	Hypospadias	Hyspospadias	2y, Hypospadias	No	No	
GN0196	46,XY/F	46,XY DSD CGD	Ambiguous genitalia	At birth, micropenis (2x1.5) with vaginal opening and bifid scrotum looking like labia majora without hyperpigmentation. Non-palpable gonads. 12d, US: normal uterus (3.5cm), in intra-abdominal location an ovoid structure (11.8mm) suggesting right testis.	Yes	No	
GN0198	46,XX/F	46,XX DSD GD	Primary amenorrhea	17, primary amenorrhea, delayed puberty and hypergonadotropic hypogonadism. US and MRI: non-visible ovaries and uterus	No	No	
GN0199	46,XY/M	46,XY DSD CGD	Micropenis	7d, micropenis with fused labia minora, non-palpable gonads, tight vagina. US: normal uterus, absence of gonads.	Yes	No	
GN0200	46,XY/M	46,XY DSD PGD	Obesity, micropenis	At birth, hypospadias. 8mo, micropenis. 2y, surgery for hypospadias. 10y, normal male phenotype, penis buried in fat, testes in scrotrum (2-3ml). **Other features:** Obesity.	No	No	
GN0201	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	15d, micropenis (2.5cm), non-palpable testes and hypoplastic scrotum. 2mo, micropenis and bilateral crytorchidism. US: absence of gonads. 8mo, good response to the treatment with testosterone. 10mo, right orchidopexy. 3y, left orchidopexy. 6y, surgery for hypospadias. 7y, normal penis and testes (1-1.5ml)	No	No	
GN0202	46,XY/F	46,XY DSD CGD	Primary amenorrhea	15y, primary amenorrhea, female external genitalia, pubertal delay and hypergonadotropic hypogonadism. 17y, MRI: normal vagina, rudimentary uterus, left ovary (<1cm) and absence of right gonad. **Other features:** 15y, Obesity (IMC 37,52). 18y gender dysphoria	Yes	No	
GN0203	46,XY/M	46,XY DSD PGD	Ambiguous genitalia	5y, micropenis with fused scrotum and bilateral cryptorchidism. US: atrophic testes in inguinal canal. **Other features:** Born in Syria	No	No	
GN0204	46,XX/F	46,XX DSD GD	Primary amenorrhea	17y, primary amenorrhea, normal female external genitalia and hypergonadotropic hypogonadism, US: small uterus, inconsistent findings for gonads.	Yes	No	
GN0205	46,XX/F	MRKH II syndrome	Vaginal agenesis	16y, primary amenorrhea, agenesia of vagina, uterus and left kidney. MRI: absence of cervix, uterus and upper vagina.	No	No	
GN0207	46,XX/F	46,XX DSD	Primary amenorrhea, puberty delay	14y, primary amenorrhea and delayed puberty. US: prepubertal ovaries and uterus. **Other features:** Mother was virilized during pregnancy	Yes	No	
OT0567	46,XY/M	Denys-Drash	Bilateral	At birth, bilateral cryptorchidism. 2y, bilateral Wilms' tumour. Received	No	No	

		syndrome	cryptorchidism, Wilms' tumour	chemotherapy. Proteinuria. At surgery, found hypoplastic testes. **Other features:** Patient from Tunisia.			
POL0274	46,XY/M	Adrenal Insufficiency	Macrogenitosomia, pubic hair development	1mo, adrenal insufficiency. 9mo, hyperpigmented genitals, hypertrophic penis (5cm) and testes 2mL, pubic hair. CT: lack of left gland. 15mo, penis 6cm. **Other features:** Patient from Mexiko	-	-	-
POL0285	46,XY/M	Adrenal Insufficiency	Salt wasting	At birth, salt wasting, hyperpigmentation. 45y, pubertal failure, hypogonadotropic hypogonadism. Family background..	-	-	-
POL0301	46,XY/M	Adrenal Insufficiency	Acute hyponatremia	2y, acute hyponatremia. Normal penis, testes (0.5mL) and no hyperpigmentation. **Other features:** Glycemia 70mg/dL; Na 119mEq/L; K 4.6mEq/L; Aldosterone 0.15ng/dL; PRA NA.	-	-	-
RE0045	46,XY/F	46,XY DSD GD	Primary amenorrhea	14y, primary amenorrhea, delayed puberty and hypergonadotrophic hypogonadism. 32y, gonadal biopsy revealed immature testes. **Other features:** 17y, diagnosed with bilateral pheochromocytoma. Atrophic kidney. Sister presenting with 46,XX DSD gonadal dysgenesis.	Yes	Yes (32y)	Immature testes with Leydig cells hyperplasia

AIS, androgen insensitivity syndrome; CAIS, complete androgen insensitivity syndrome; CGD, complete gonadal dysgenesis; CT, computed tomography; D, days; DSD, disorder of sex development; F, female; GD, gonadal dysgenesis; M, male; MGD, mixed gonadal dysgenesis; Mo, months; MRI, magnetic resonance imaging; MRKH, Mayer-Rokitansky-Küster-Hauser; N, normal; ND, not determined; OT, ovotesticular; PAIS, partial androgen insensitivity syndrome; PGD, partial gonadal dysgenesis; PRA, plasma renin activity; RD, receptor defect; US, ultrasound; Y, years.

Supplementary data 3. Biochemical data of the patients included in the genetic analysis. Out of range values are given in bold.

Biochemical data																
			Adrenal function								Gonadal function					
Case	Karyotype	Age at evaluation	ACTH (pg/mL)	Cortisol (µg/dL)	11-Deoxycortisol (ng/mL)	P4 (ng/ml)	17OHP4 (ng/mL)	DHEA-S (ng/mL)	Δ4-A (ng/mL)	PRL (ng/mL)	Testosterone (ng/dL)	FSH (U/L)	LH (U/L)	E2 (pg/mL)	DHT (ng/ml)	AMH (ng/mL)
GN0001	46,XX	17y										**65.9/**108*	**37/**150*			
		24y										**92**	**41**	71		
GN0003	46,XX	ND														
GN0004	46,XX	13y					1.8	867			**320**	**23.9**	**21.5**	21		
GN0007	46,XY	16y					1.5				29	**41.4**	**10.7**	26.5		
GN0009	46,XY	17y										**10.4**	**144**	**<10**		
GN0011	46,XY	16y				0.5	**2.5**	**183**	2.6	5.1	1	**145**	**45**	12		
GN0012	46,XY	At birth	<25								**34.7**					
		8y						**1295**			**<10**	**24.8**	**<0.5**			
		11y									12	**86**	**20**			
		13y									**55**	**85**	**42.3**			
		15y								7	**87.8/**281*	**97.9**	**14.8**			
		17y									856.2	**22.3**	**12.2**			
		19y									381	**64.6**	**34.6**			
		21y									532	**69.3**	**33.5**			
		24y									783	0.3	<0.1			
		26y								**16.6**	348	**27.5**	**10.2**			
		27y									**899**	0.3	0.2			
GN0013	46,XY	18y									**2540**	**109**	**50.3**	12.2		
GN0014	46,XY	13y						2600			16	**50**	**35**	<10		
GN0017	46,XY	12m	29	10.7							**40/**30*	**41.7/**4*	**19.1/**8.7*			
GN0018	46,XY	ND														
GN0020	46,XY	13y					N*	N*						<5		
GN0023	46,XX	ND														
GN0024	46,XY	1m							**1.4**		**73**				**0.5**	**11**
		2m							**1.5**		**54**					**11.8**
GN0025	46,XY	17y								8.7		**127**	**54**	<10		
GN0026	46,XY	16d					**14.7**									

GN0027	46,XY	15y					**665**				**<20**	**58.8**	**16.3**	>20		
GN0028	46,XY	10y							1.7/1.9*		86.3/187.4*	11.9	0.9		0.2/0.1*	
		11y					1.3		2.5		290.5	30.4	7.1		0.5	
		14y						1900			**412.5**	**35.9**	**11.4**			**<0.1**
		15y						2010	4.1		518	**38.9**	**14.6**		0.5	**0.4**
GN0029	46,XY	ND														
GN0031	46,XX	16d				**3.49**					**19**	0.8	**<0.1**	<10		
GN0033	46,XY	12y		11.2							16.1	0.4	**<0.1**			10.2
GN0034	46,XY	8m									<<	1	2			
		9y									**17**	2.8	<<			
		11y									**60**	15	22		**0.4***	
		13y						**2166**			**23.4**	125	30.6	<<		
		15y										110	31.4			
GN0035	46,XY	42y		9.2		**<0.2**	0.3	1866	0.7	1.8	**16.8**	**69.1**	19.3	11		
GN0037	46,XY	4y	N/N*	N			N				0.1/3.6					
		11y									5.7/6.8*				**0.4/**0.5*	
GN0038	46,XY	11y						1400	3.9		200	18.1	11.8	53		
GN0039	46,XY	ND														
GN0041	46,XY	10y						368			<10	**31**	**2.5**	<10		
		12y									<10	**76**	**9**	<10		
		13y									<10	**112**	**108**	<10		
		17y						627			15	**23**	**10**	<10		
GN0042	46,XY	10y				**4400**	1.9		**1.0**		**250**	**24**	**95**	16		
		27y					0.6		2.5			**21.4**	8.44	26.4		
GN0043	46,XY	5d		16	9		**2.4**		**2.1**		**160**				2.1	
GN0046	46,XY	ND														
GN0050	46,XY	3y									**29/**22	**111**	**6**			
GN0051	46,XX	6y					**<0.1**	<180			<10/44.1*				**0.3/**0.1*	
GN0054	46,XX	41y		23.8						7.5	**63.5**	**31.7**	5.3			
		52y (After treatment)					0.8				616.3					
GN0055	46,XY	ND														
GN0056	46,XY	14d										**1.8**	**1.8**			
		24d						0.2/0.2*	0.5/0.5*		**110/**325*				0.5/0.4*	
GN0059	46,XY	At birth			4.7		**1.55**		**1.7**		**218**	0.6	<0.5	**199**	**1.8**	

		7d							**0.7**		**56**	**9.1**	**3.9**	15	**0.5**	
		13d							**0.8**		**40**	**9.8**	1	20		
		1m							0.5		**14**	**18.5**	**2.4**	15	0.3	
		2y	37.4	18.5/30.7 *	2.1/4.2*		**1.1/4.1***		0.5/0.6*		**28**	**217.0**	**58.2**			
GN0064	47,XY	At birth									<<	41	<<			
GN0066	46,XY	4d						689			**65.9**	<1.5	<1.5			
		7d									**21.5**	<1.5	**3.7**			
		8y					0.1	**1690**			**16.7**					
GN0068	46,XY	ND														
GN0070	46,XY	9y									80*				N*	
GN0075	46,XY	2d		6.9				202			**2360**	<0.5	<0.5		1.5	
		1m		10.4				400			**6750**	**3.7**	**2.5**			
GN0076	46,XY	8d									**105**	0.4	0.2	17		
		15d									**121/**900*					
GN0078	46,XY	ND		17.8							**170**					
GN0080	46,XY	ND														
GN0084	46,XX	16y	N/N*				1.5									
		21y				**1.3**	1	3480			25	7	7.2	78		
		22y				**0.8**	**1.5**	3490			25.3	6.5	5	37		
GN0088	46,XY	3y									**92.2**	**3.9***	**2.7***			
		4y									**170.1**	0.3	0.1/0.7*			
GN0090	46,XX	6y					0.2	<180	<0.3		<10/125*	<1.5	<1.5		<0.1/0.1 *	
GN0091	46,XY	8m	**1550**	**1.63**			**1.8**	**8.32**			**<0.2**	**4.5**	**4.2**			
GN0096	46,XX	3y						**24.9**			<0.12/143.7*	**2,3**	<0.1		**0.1**	
		4y							**0.5/<0.3 ***		**14/59**	1,3	<0.1		**0.1/0.1***	8.8
GN0100	46,XY	16y	19.8	8.5			0.90	3032		6.7	**34**	**145**	**41**			
GN0101	46,XY	17y	**1747**	**4**			0.2				**170**	**4.6**	**2.5**			
GN0103	46,XY	At birth						**<17**	0.1		<8	**99.5**	**26.9**			
		6m		1.32							<9/9*	**98.43**	**14.04**			**0.2**
		6m (During/afte r treatment)										0.2/13.2	<0.1/0.7			

GN0104	46,XY	16m					0.8		0.1	**30.6**	**20**	0.7	0.2		<0.1	
		21m					0.1		**<0.1**							
GN0108	46,XY	12y									**95**	**11.4**	6.3			
		14y						2180	**7.8**		**833**	**16.2**	**22**	**42**	0.7	**16.7**
GN0109	46,XY	14y	**52**	**16.3**			**0.9**	1690	1.7	**12.8**	**193**	**112.5**	**37.8**	16.7	0.5	**0.4**
GN0111	46,XY	14y		18/27*				1960	2.1		**36.6/**46*	**44.3**	12.2	31.0		<0.1
GN0112	46,XY	4d		**2.3**			**5**	492			122	**0.2**	**<0.1**	22.2		
		17d								**94.4**	43	**0.2**	**<0.1**	<11.8	0.2	**105.2**
GN0114	46,XY	ND										>>	>>			
GN0118	46,XY	2y		17.7		**0.2**		<150	<0.3/<0.3*		<10/110.4*				**<0.1/<0.1***	
GN0119	46,XY	6y		11			0.4	270	<0.3/<0.3*		<10/263.4	0.8	<0.1		0.1/0.3*	22.8
GN0122	46,XY	2y		4.9		0.1	0.1	<150	<0.3/1.2*		<10/534.2*	0.9	<0.1	15	<0.1/0.4*	
GN0123	46,XY	10y									2/13					
		14y	N	N			N				18/35	**55**	**15**			
		15y									**140**/500*	**39/**72*	**11/**137*			
		18y								9	**200**	**33**	**21**			
		38y	36	14.3			0.4	1293	1.5	7	**500**	**25**	**15.6**	20		<0.1
GN0124	46,XY	4m						374	0.5		**205**	0.6	4.8	<12	0.2	99.4
		6m							0.3/0.8*		**77.3/**680*				0.1/0.3*	112.9
GN0125	46,XY	17y	**69**	12.3								**10**	**70**			
GN0132	46,XY	ND														
GN0133	46,XX	42y								7.8	335		2.09	**41.7**		
		44y								12.6	**238**		3.96	30.8		
GN0138	46,XY	At birth					**5.3**				**365**	**2.2**	**3.1**			**19.2**
		4d									**97**				0.6*	
		8d									**69**	**6.2**	**14.5**			
GN0139	46,XY	16y				0.8	**1.1**	3060	2.1	17.4	463.3	**14**	**52.5**	**22**		
		17y										**89.3**	**53.7**	**19**		
GN0141	46,XY	25y										**71**	**33**	**49**		
GN0142	46,XY	4d		N					**1.3**		39	**3.5**	0.3		0.3	<<
		12d							**1.8**		33				0.2	<<
		22d							0.5		1	**35.4**	**3.2**		0.1	<<

GN0144	46,XY	5y						150			5	0.6	<0.1			
		6y							0.4		<5/650*	0.5	<0.1		1.5*	
GN0145	46,XY	20y									607	N	N			
GN0146	46,XY	14y	**79.3**	11.6			0.3	575	0.5	3.8	6	1.87	**0.3**	**30.6**		
		19y (During treatment)									8	2.2	0.6	30.5		
		19y	**80.8**	10.3			<0.01	**806**	0.5	4.4		3.5	1.3	5		
GN0147	46,XY	35y	27.9	21.3			0.2	1390	1.3	7.6	**19**	**67.0**	**48.5**	5		
GN0148	46,XY	20y		N				N	N	N	31	51	28	15.8		
		25y (During treatment)	113	17.8	3.4	0.9		344		23.9	20	24.8	16.7	26	0.3	
GN0150	46,XY	9y		6.5			0.4	683	<0.3		18	33.1	1.6	<12	0.1	<0.1
GN0151	46,XX	ND														
GN0152	46,XY	5m		7.2			0.9	<150			9	**3**	1	<12	0.2	
		19m	22.7	22.8			0.3/0.3*	**1260**	<0.3/<0.3*		3/17*	**3.5**	1.1	5	0.2/0.2*	
GN0153	46,XY	8m	**2297**	**2.9**			0.3									
GN0154	46,XX	15y		N					N			**70**	**18**	22		
GN0155	46,XY	37y									**19**					
GN0156	46,XY	6y					1.4	0.6		14.1	**434.7**	4.6	4.6	22		
		7y	N	N			0.9	N	N		**406**	3.5/7.1*	2.5/**25.0***	N		
GN0157	46,XY	6y				0.3	0.1	1100	<0.3		<20					
		7y			1.4		0.3	1100			<20	0.5	<10			
GN0158	46,XX	3y		13.9			0.1	<150	<0.3/<0.3*		**<10/**72.8*	**3**	<0.1		**<0.1**	7
GN0159	46,XX	9y		13.1		0.1		480	<0.3	17	**<10**	1	0.1		<12/0.1*	
		11y									**<10**	1	0.3			
		12y									**<10**	0.7	1.2			
		13y									35	**0.8**	2	<12		
GN0160	46,XX	ND														
GN0162	46,XY	16m				**0.2**	0.4	168	<0.3		<10	0.7	<0.1		<0.1	
GN0163	46,XY	10y							<0.3		<10	1.3	<0,1		0,1	193
		14y				0.4	0.9	1370	1.3		510	1.7	2.3	22	0.3	9
GN0164	46,XY	ND														

GN0167	46,XY	5d					**1.7**	**740**	**1.6**		**72**	0.5	**1.6**	<5	**0.3**	66.3
		2m									**108**	1	**5**	<5	0.1	**115.1**
GN0169	46,XY	At birth							1.3		178					
		8y							0.3		20	1.5	1.5		0.1	52.3
GN0171	46,XY	11 y									**130**	**63.3**	**14**			
		14 y									**130**	**36.7**	**18.5**			
		15 y									**590**	1.5	<0.5			
		21 y									**1180**	**9.5**	4			
GN0173	46,XY	At birth									211					**0.5**
		3m	34.9*	15.2*	2.1*	0.6*	37*	202.7*			**3***	**21.2***	**0.8***		<0.1*	0.5*
		6m	13*	6.3*	1.6*	**0.1***	8*	77.6*			3*	**40.4***	6.0*			<0.1*
GN0174	46,XY	5m				**0.8**	**1.7**	197	<0.3	**43**	**134**	0.8	**1.8**	<12.0	0.1	**240**
GN0175	46,XX	14y										**9/**22*	**9/**99*	34		
GN0176	46,XY	3m					N		0.4/0.6*		**146.8/**624.3*				0.1/0.2*	N
GN0177	46,XY	10d	30.5	12.8		4.4	1.2	13.1	6.4		122	4.6	2.3	8.7	0.5	75.6
		16d				2.5*	1.4*	2852*	4.5*		**141/65***	20.6*	5*		**0.4/0.2***	
GN0178	46,XY	3m									387*	<</<<*	<</<<*		0.1*	
		10y		17.8						29						
GN0179	46,XX	ND														
GN0182	46,XY	ND														
GN0183	46,XY	20y								14.1	**72.9**	**8.2**	6.5	14		
		21y					1.2	2270	1.2	8.8	**71.3**	**9.9**	6.3	14.3	0.5	
GN0185	46,XY	5d		14	4.3		**1.9**	92	0.9		0.1	11	1.5	<5	**0.4**	
GN0186	46,XY	14y									**450**					
		15y		10.4							**5**	3.1	6.7	**51**		**471**
GN0187	46,XX	34y									**416**	**46.4**	**20.8**	36		
GN0189	46,XY	4d														>21
		20d					**3.0**	**710**			1.6	0.5	**3.4**	<10	<0.1	
GN0190	46,XX	14y		**35**	4.5	0.4						**169.4**	**61.3**	5		
GN0191	46,XY/X0	19y		11.6							**46**	**46**	**15**	5		
GN0192	46,XY/X0	ND														
GN0194	46,XY	ND														
GN0195	46,XY	ND														
GN0196	46,XY	4d			9.2		**3.1**				**45**	**4.6**	1.2			
		12d														**6.5**

GN0198	46, XX	17y										**103**	**33.4**	**<5**		
GN0199	46,XY	15d		21.4		**0.6**	**8.2**	247	**3.2**		1.5	**17.3**	**17.1**	16.2		
GN0200	46,XY	10y					0.6/1.6*		**<</0.3***		**2/40***	**0.3**	0.29			
GN0201	46,XY	2mo		12			12	863	1.1		9/50*	**4.3**	**7.8**			942
		5y		6.5			0.4	**168**	<0.3		3	**0.7**	**<0.1**	<10	<0.1	>23
GN0202	46,XY	15y								10	**38**	**77.3**	**35.3**	<12		
		17y		9.9			0.6	1580	1.5	8	**0.3**	**63.7**	**30.5**	25		**<0.1**
GN0203	46,XY	ND														
GN0204	46,XX	17d				<0.1					<10	**103**	**31.5**	<10		<0.2
GN0205	46,XX	16y				0.4				16.4	27	6.0	6.8	35.1		
GN0207	46, XX	14y		12.2			0.4	240	<0.3		0.1	**50.6**	**22.1**	**<5.4**		<0.2
OT0567	46,XY	ND														
POL0274	46,XY	9mo	**<10**	20.2			334	50	**0.3**		**235.3**	1.1/2.2*	1.1/5.6*			
POL0285	46,XY	ND														
POL0301	46,XY	2y	**777**	**2.3**												
RE0045	46,XY	ND														

ACTH, adrenocorticotropic hormone; AMH, anti-Müllerian hormone; D, days; DHEA-S, dehydroepiandrosterone sulfate; DHT, dihydrotestosterone; E2, estradiol; FSH, follicle-stimulating hormone; hCG, human chorionic gonadotropin; LH, luteinizing hormone; Mo, month; N, normal; ND, not determined; PRL, prolactin; P4, progesterone; Y, years; Δ4-A, delta 4-androstenedione. (*) Values after stimulation with hCG or ACTH

Supplementary data 4. Hospitals in which the patients involved in the study were clinically diagnosed.

- Hospital Universitario Cruces (Barakaldo, Spain)
- Hospital Universitario Basurto (Bilbao, Spain)
- Hospital Universitario Araba (Vitoria-Gasteiz, Spain)
- Hospital Universitario Donostia (San Sebastián, Spain)
- Hospital Bidasoa (Hondarribia, Spain)
- Complejo Hospitalario de Navarra (Pamplona, Spain)
- Complejo Hospitalario La Mancha Centro (Ciudad Real, Spain)
- Hospital Universitario de Getafe (Getafe, Spain)
- Hospital Manises (Manises, Spain)
- Hospital Clínico Universitario Valladolid (Valladolid, Spain)
- Hospital Comarcal de la Axarquia (Vélez-Málaga), Spain)
- Hospital de Mérida (Mérida, Spain)
- Hospital General Universitario de Alicante (Alicante, Spain)
- Hospital General Universitario Gregorio Marañón (Madrid, Spain)
- Hospital Público General de Tomelloso (Tomelloso, Spain)
- Hospital Regional Universitario de Málaga (Málaga, Spain)
- Hospital Sierrallana (Torrelavega, Spain)
- Hospital Universitario Central de Asturias (Oviedo, Spain)
- Hospital Universitario de La Princesa (Madrid, Spain)
- Hospital Universitario Germans Trias i Pujol (Badalona, Spain)
- Hospital Universitario Infanta Cristina (Parla, Spain)
- Hospital Universitario Marqués de Valdecilla (Santander, Spain)
- Hospital Universitario Nuestra Señora de la Candelaria (Santa Cruz de Tenerife, Spain)
- Hospital Universitario Príncipe de Asturias (Alcalá de Henares, Spain)
- Hospital Universitario Puerta del Mar (Cádiz, Spain)
- Hospital Universitario Ramón y Cajal (Madrid, Spain)
- Hospital Universitario Vall d'Hebron (Barcelona, Spain)
- Hospital Universitario Virgen de las nieves (Granada, Spain)
- Hospital Universitario Virgen del Rocío (Sevilla, Spain)
- Hospital Virgen del Puerto (Plasencia, Spain)
- Inselspital, Universitätsspital Bern (Bern, Switzerland)

Supplementary data 5. Clinical data sheet for the inclusion of a patient in the study.

CLINICAL QUESTIONAIRE FOR THE GENETIC ANALYSIS OF DISORDERS OF SEX DEVELOPMENT

FAMILY CODE**:** Referring doctor:

DATA AT BIRTH

Date of birth:

Gestational age (Weeks):

Weight (gr): Length (cm):

Twins/Multiple birth: Yes: No:

ADDITIONAL DATA

Karyotype: SRY (FISH):

Age at diagnosis (years/months/days):

Assigned sex: Male Female

Gender reassignment: Yes No Unknown

If Yes, age at gender reassignment: (Years/months)

REASONS FOR THE FIRST REFERRING

Ambiguous genitalia: Yes No Unknown

Inguinal hernia: Yes No Unknown

Gynecomastia: Yes No Unknown

Prenatal diagnosis: Yes No Unknown

Adrenal insufficiency/Salt loss: Yes No Unknown

Insufficient virilization: Yes No Unknown

Primary amenorrhea: Yes No Unknown

Others: Yes No Unknown

COMMENTS:

FAMILY BACKGROUND

PARENTS: Consanguinity: Yes No Unknown

Adopted: Yes No Unknown

MOTHER: Normal: Yes No Unknown

Virilization during pregnancy: Yes No Unknown

Drugs during pregnancy: Yes No Unknown

If yes, which drug:

FERTILISATION:

-Coital: -IVF (*in vitro fertilization*): -

ICSI (*intracytoplasmatic sperm injection*): -Others:

BROTHERS/SISTERS:

If affected, gender: M F Unknown

Brief description:

OTHER AFFECTED FAMILY MEMBERS:

Sex: M F Unknown

Brief description:

FAMILY BACKGROUND OF INFERTILITY: Yes No Unknown

If yes, which family member?

M (male); F (female)

EXAMINATION

EXTERNAL VIRILIZATION:

Penis/Clitoris	**Hypospadias**	**Labioscrotal fusion**	***Gonads***	***Right/Left***
Normal for M	Normal for M	Yes	Labioscrotal	/
Small for M	Distal Hyposp	No	Inguinal	
Big for M	Midpenile Hyposp	Unknown	Abdominal	
Normal for F	Proximal Hyposp			
Big for F	N for F			

VIRILIZATION IN 46, XX:

Prader scale: I II III IV V

FEMALE INTERNAL GENITALIA

Are present: Yes No Unknown

Vagina: Yes No Unknown Type:

Uterus: Yes No Unknown Type:

Fallopian tubes: Yes No Unknown

Ultrasound: Age: Report:

Laparoscopy: Age: Report:

PUBERTAL SIGNS Spontaneous Induced Age (Years/months):

FEMALE

Thelarche (Tanner): M1 M2 M3 M4 M5 Unknown

Pubarche (Tanner): P1 P2 P3 P4 P5 Unknown

Menarche: Yes No Unknown

MALE

Volume right testis (ml): Volume left testis (ml):

Penis length (cm): Gynecomastia (diameter cm):

Pubarche (Tanner): P1 P2 P3 P4 P5 Unknown

BIOCHEMICAL ANALYSIS

Date:

Karyotype:

-Blood: -Gonad:

AMH (ng/ml):

INHIBINA B (pg/ml):

Beta-HCG (mUI/ml):

Alpha-Fetoprotein (ng/ml):

Basal

Age:

LH (UI/l): FSH (UI/l):

Testosterone (ng/dl): Androstendione (ng/ml): Dihidrotestosterone (ng/ml):

Estradiol (pg/ml): Estrone (pg/ml):

17-OH Progesterone (ng/ml): Progesterone (ng/ml):

17-OH Pregnenolone (ng/ml): 11-Desoxicortisol (ng/ml):

DHEA-S (ng/ml): DHEA (mcg/dl): Cortisol (mcg/dl):

Extraction date:

After-stimulating

Age

-**GnRH/LHRH** test:

FSH peak (min) (UI/l): **LH** peak (min) (UI/l)

Basal testosterone (ng/dl): Basal estradiol (pg/ml):

-**HCG** test**:** dose (UI) y number of doses:

Testosterone (ng/dl): Androstendione (ng/ml):

Dihidrotestosterone (ng/ml):

-**ACTH test**: peak at 60 mins

17-OH Progesterone (ng/ml): Cortisol (mcg/ml):

17-OH Pregnenolone (ng/ml) : 11-Desoxicortisol (ng/ml):

IMAGING

Ultrasound:

MRI Scan

LAPAROSCOPY: Yes No

REPORT:

GONADAL BIOPSY: Yes No

REPORT:

MOLECULAR STUDY

Previously studied genes:

Affected gene and variant:

PROGRESS

-Treatment for micropenis: Yes/No

Testosterone for topical use: Dosage: Number of days:

IM testosterone: Type: Dosage (mg): Number of doses/frecuency:

-Treatment for hypospadias: Yes/No

Report:

-Vaginoplasty: Yes/No

Report:

-Clitoroplasty: Yes/No

Report:

-Gonadectomy: Yes/No

Report:

-Other surgeries: Yes/No

Report:

-Hormone replacement therapy during puberty: Yes/No

Testosterone IM/topical use: Type: Dosage (mg): Number of doses/frecuency:

GNRH: Dosage

FSH: Drug name: Dosage: Frecuency:

LH: Drugname: Dosage: Frecuency:

Estrogens: Drugname: Dosage: Frecuency:

Gestagens: Drugname: Dosage: Frecuency:

DIAGNOSIS:

RELATED PATHOLOGIES:

Supplementary data 6. Primers, probes and conditions used for the amplification by PCR, QMPSF, genetic markers and site-directed mutagenesis.

Primers and conditions used for the PCR-based amplification.

Gene	Amplicon name	Primers Name	Primers sequence (5' to 3')	Conditions	Amplicon size (pb)	Others (DMSO…)
SRY	Amplicon 1	F1-SRY R1-Int-SRY	gcttgagaatgaatacattgt ggtatttctctctgtgcatg	94ºC3'35X(94ºC1'15''58,7ºC1'15''72ºC2')72ºC7'4º	574	
	Amplicon 2	F1-Int-SRY R1-SRY	gcgaaactcagagatcagca gcatgttaatcgtgttgacag	94ºC3'35X(94ºC1'15''58,7ºC1'15''72ºC2')72ºC7'4º	559	
HSD17B3	Exon 1	M13F1-HSD17B3 M13R1-HSD17B3	ggaacacgtccagtgact gccattgcactccagcct	95ºC5'35X(95ºC30''64ºC30''72ºC30'')72ºC7'4º	382	
	Exon 2	F2-HSD17B3 R2-HSD17B3	tgaattctgtcttttaaaagca aatacaagggaggagaaagtcccca	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	127	
	Exon 3	F3-HSD17B3 R3-HSD17B3	gctcatcatccttgtctcttggttt gagggctccacacacatctccctta	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	160	
	Exon 4	F4-HSD17B3 R4-HSD17B3	tggatccctgttcattaaaaaaact gatgtatgacaacaagctttgcatc	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	189	
	Exon 5	F5-HSD17B3 R5-HSD17B3	ctgatcttctgacacatttttgttt ccaggttctgggtcccctggct	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	168	
	Exon 6	F6-HSD17B3 R6-HSD17B3	gagaatttctctaatcatccggctg acatgttaatgcatttcgcaca	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	126	
	Exon 7	F7-HSD17B3 R7-HSD17B3	agttccttgtcgggcttacctttgg agggcagggaggccatgttgctcca	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	121	
	Exon 8	F8-HSD17B3 R8-HSD17B3	caacaaagccatgggaac aaggaagagacttggaagtcatgac	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	154	
	Exon 9	F9-HSD17B3 R9-HSD17B3	agctcactctggggcctcaggtgtc gatgacaaggactccacagctg	95ºC5'35X(95ºC30''65ºC30''72ºC30'')72ºC7'4º	140	
	Exon 10	F10-HSD17B3 R10-HSD17B3	gattgcttctgtgccatggtctttg ttcaagaaaaggagaagtt	95ºC5'35X(95ºC30''51,1ºC30''72ºC30'')72ºC7'4º	216	
	Exon 11	F11-HSD17B3 R11-HSD17B3	tgaactgaggtacttgttattcc gaggaaaaggttgtgctggactcct	95ºC5'35X(95ºC30''55,7ºC30''72ºC30'')72ºC7'4º	206	
AR	Exon 1, Amplicon 1A	F1A-AR	gcctgttgaactcttctgagc	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	535	DMSO (10%)

		R1A-AR	ggaggtgctggcagctgct			
	Exon 1, Amplicon 1B	F1B-AR R1B-AR	cacaggctacctggtcctgga ctgccttacacaactccttggc	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	416	DMSO (10%)
	Exon 1, Amplicon 1C	F1C-AR R1C-AR	gctcccacttcctccaaggac cgggttctccagcttgatgcg	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	528	DMSO (10%)
	Exon 1, Amplicon 1D	F1D-AR R1D-AR	ccagagtcgcgactactacaactt ggactgggatagggcactctgct	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	474	DMSO (10%)
	Exon 1, Amplicon 1E	F1E-AR R1E-AR	gacttcaccgcacctgatgtg ccagaacacagagtgactctgcc	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º	193	DMSO (10%)
	Exon 2	F2-AR R2-AR	gcctatttctgccattca cctgggccctgaaaggt	94ºC5'35X(94ºC30''52,9ºC30''72ºC30'')72ºC7'4º	288	
	Exon 3	F3-AR M13R3-AR	ttatcaggtctatcaactcttgt ctgatggccacgttgcctatgaa	94ºC5'35X(94ºC30''50ºC30''72ºC30'')72ºC7'4º	313	
	Exon 4	F4-AR R4-AR	gataaattcaagtctctcttcct gatccccctatctcatgctccc	94ºC5'35X(94ºC30''54ºC30''72ºC30'')72ºC7'4º	360	DMSO (10%)
	Exon 5	M13F5-AR M13R5-AR	caacccgtcagtacccaga agcttcactgtcaccccatcacca	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	290	DMSO (10%)
	Exon 6	M13F6-AR R6-AR	ctctgggcttattgtaaacttcc gtccaggagctggcttttcccta	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	296	DMSO (10%)
	Exon 7	M13F7-AR M13R7-AR	ctttcagatcggatccagctat ctctatcaggctgttctccctgat	94ºC5'40X(94ºC45''58ºC45''72ºC45'')72ºC10'4º	416	
	Exon 8	M13F8-AR M13R8-AR	gaggccacctccttgtcaaccctg ggaacatgttcatgacagactgtacatca	94ºC5'35X(94ºC30''52,9ºC30''72ºC30'')72ºC7'4º	347	DMSO (10%)
WT1	Exon 1	F1-WT1 R1-WT1	gggcgtccgggtctgagcc gcggagagtccctggcgc	94ºC5'35X(94ºC60''62,3ºC60''72ºC60'')72ºC7'4º	526	DMSO (10%)
	Exon 2	F2-WT1 R2-WT1	cgagagcaccgctgacact gagaaggactccacttggttccg	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	199	
	Exon 3	F3-WT1 R3-WT1	ccaggctcaggatctcgtgt caaggacccagacgcagag	94ºC5'35X(94ºC60''59,5ºC60''72ºC60'')72ºC7'4º	238	
	Exon 4	F4-WT1 R4-WT1	tgcttttgaagaaacagttgtg ggaaaggcaatggaatagaga	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	178	
	Exon 5	F5-WT1 R5-WT1	gggcttttcactggattctg ccatttgctttgccatctcc	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	174	

	Exon 6	F6-WT1 R6-WT1	gtgagccacactgagccttt ggccggtaagtaggaagagg	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	200	
	Exon 7	F7-WT1 R7-WT1	gacctacgtgaatgttcacatg cttagcagtgtgagagcctg	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	286	
	Exon 8	F8-WT1 R8-WT1	gagatccccttttccagtat caacaacaaagagaatca	94ºC5'35X(94ºC60''55,7ºC60''72ºC60'')72ºC7'4º	220	
	Exon 9	F9-WT1 R9-WT1	aagtcagccttgtgggcctc tttccaatccctctcatcac	94ºC5'35X(94ºC60''55,7ºC60''72ºC60'')72ºC7'4º	246	
	Exon 10	F10-WT1 R10-WT1	tgtgcctgtctctttgttgc gttcacacactgtgctgcct	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	224	
NR5A1	Exon 1	M13F-UTR-NR5A1 M13R-UTR-NR5A1	gcgagaggcctgcagagt gcactgcagaaggaggct	95ºC5'35X(95ºC45''64ºC45''72ºC45'')72ºC7'4º	304	
	Exon 2+3	M13F1-NR5A1 M13R2-NR5A1	ggaccccaggctgccggtct gcgaaggccaatggtact	95ºC5'35X(95ºC45''64ºC45''72ºC45'')72ºC7'4º	488	DMSO (10%)
	Exon 4	M13F3-NR5A1 M13R3-NR5A1	ggtcagtgggagccatga ggacagtcgggctaaggct	94ºC5'35X(94ºC30''58,7ºC30''72ºC30'')72ºC7'4º	832	DMSO (10%)
	Exon 5	M13F4-NR5A1 M13R4-NR5A1	ggtgagaggaaggtccctgga cctgaatcctggaagtgca	95ºC5'35X(95ºC45''64ºC45''72ºC45'')72ºC7'4º	250	DMSO (10%)
	Exon 6	M13F5-NR5A1 M13R5-NR5A1	cctccaatccatgccctca cctctggctgtctccacct	95ºC5'35X(95ºC45''64ºC45''72ºC45'')72ºC7'4º	291	DMSO (10%)
	Exon 7	M13F6-NR5A1 M13R6-NR5A1	ggaaggtggtattggtgatgct cctcggtgggcatcaga	95ºC5'35X(95ºC45''64ºC45''72ºC45'')72ºC7'4º	454	DMSO (10%)
FSHR	Exon 1	M13F1-FSHR M13R1-FSHR	ggtcacatgaccctacca cgcaatgcacaaatgcca	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	332	
	Exon 2	M13F2-FSHR M13R2-FSHR	ggtaactgtttgctga ggtctgaggttgctccct	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	223	
	Exon 3	M13F3-FSHR M13R3-FSHR	ccagacacaaggctct ccaggaatgtagaaga	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	234	
	Exon 4	M13F4-FSHR M13R4-FSHR	gcattccttaccatcaaga ggcttgcactatttaatca	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	174	
	Exon 5	M13F5-FSHR M13R5-FSHR	gccagtttcattttcactca gcaagacagatactga	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	166	
	Exon 6	M13F6-FSHR	ggaaggttacatgagtaga	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	215	

	Exon 7+ 8	M13R6-FSHR M13F7-FSHR M13R8-FSHR	ggagcatccaattatgaga cccgtgtattgtttgca gctgcagagagttgacttct	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	381	
	Exon 9	M13F9-FSHR M13R9-FSHR	gcctgctaaccaagagcaga ggacaaagttctacattgggga	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	311	
	Exon 10	M13F10-FSHR M13R10-FSHR	cctgcacaaagacagtga ccttcaaaggcaagactga	94ºC5'35X(94ºC60''51,1ºC60''72ºC60'')72ºC7'4º	1367	
WNT4	Exon 1	M13F1-WNT4 M13R1-WNT4	cagcagcgggcaggctgccggcagg agcgagcgagcctccggtcc	94ºC5'35X(94ºC45''65ºC45''72ºC45'')72ºC7'4º	365	DMSO (10%)
	Exon 2	M13F2-WNT4 M13R2-WNT4	tggatccagagagaagtcgg ttgctcacgagcgtctcatt	94ºC5'35X(94ºC45''65ºC45''72ºC45'')72ºC7'4º	460	
	Exon 3+4	M13F3-WNT4 M13R4-WNT4	ccagctctgccctccctctgc tctgagtggccgtgtgggt	94ºC5'35X(94ºC45''65ºC45''72ºC45'')72ºC7'4º	482	
	Exon 5	M13F5-WNT4 M13R5-WNT4	ggcacaacggcaaatctgactg ttatcggccttccctg	94ºC5'35X(94ºC45''65ºC45''72ºC45'')72ºC7'4º	652	
SRD5A2	Exon 1	M13F1-SRD5A2 M13R1-SRD5A2	gcagcggccaccggcgagg agcagggcagtgcgctgcact	94ºC5'35X(94ºC45''68ºC45''72ºC45'')72ºC7'4º	358	
	Exon 2	M13F2-SRD5A2 M13R2-SRD5A2	tgaatcctaacctttcctccc agctgggaagtaggtgagaa	94ºC5'35X(94ºC45''58ºC45''72ºC45'')72ºC7'4º	235	
	Exon 3	M13F3-SRD5A2 M13R3-SRD5A2	tgtgaaaaaagcaccacaatct cagggaagagtgagagtctgg	94ºC5'35X(94ºC45''58ºC45''72ºC45'')72ºC7'4º	208	
	Exon 4	M13F4-SRD5A2 M13R4-SRD5A2	tgattgaccttccgattctt tggagaagaagaaagctacgt	94ºC5'35X(94ºC45''58ºC45''72ºC45'')72ºC7'4º	232	
	Exon 5	M13F5-SRD5A2 M13R5-SRD5A2	tcagccactgctccattatat cagttttcatcagcattgtgg	94ºC5'35X(94ºC45''58ºC45''72ºC45'')72ºC7'4º	166	
LHCGR	Exon 1	F1-LHCGR R1-LHCGR	ccgcactcagaggccgtccaaga ggcatagagcgatggagggtcctgca	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	279	DMSO (10%)
	Exon 2	F2-LHCGR R2-LHCGR	cctcagcctgaatccagttct ccccttttcaaatgtgttttctct	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	290	
	Exon 3	F3-LHCGR R3-LHCGR	ccagttgttgggtcacacacatagct ccaagtgggctccagccagtga	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	236	
	Exon 4	F4-LHCGR R4-LHCGR	gccagcaacttctggtgacca cagaaggctgaagaggaaca	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	301	

	Exon 5+ 6	F5-LHCGR R6-LHCGR	cctgataacaccaaactca cctagtagtgagactagcagga	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	449	
	Exon 7	F7-LHCGR R7-LHCGR	ggagaactagatattatgaatgcct cctgagttagttgctgaaga	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	286	
	Exon 8	F8-LHCGR R8-LHCGR	ggagctacgaatgcttacacatgaggt ggacaccctaagcagtcctgt	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	203	
	Exon 9	F9-LHCGR R9-LHCGR	ggcgacggagcaagactccgt ggagtgaaggggagtggagct	94ºC5'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	366	
	Exon 10	F10-LHCGR 110-LHCGR	cgcacagtcaggtttagcctgaagt ggtgcacacagaacaagatacga	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	200	
	Exon 11	F11-LHCGR R11-LHCGR	gctattatggctttgtttcct ggtctcttgcctaatgtacct	94ºC5'35X(94ºC30''55,5ºC30''72ºC30'')72ºC7'4º	1196	
CYP17A1	Exon 8	M13F8-CYP17A1 M13R8-CYP17A1	ggtgctattttcataggt ggtgttgaaagaatgagt	94ºC10'35X(94ºC30''56,4ºC30''72ºC30'')72ºC7'4º	758	
ESR1	Exon 5	M13F5-ESR1 M13R5-ESR1	ggagagccacttgttgaacact gacatcacaacaagttc	94ºC5'35X(94ºC30''55ºC30''72ºC30'')72ºC7'4º	548	
	Exon 9	M13F9-ESR1 M13R9-ESR1	gaaagccctcagctttccca gctataataaacccttga	94ºC10'35X(94ºC30''58,9ºC30''72ºC30'')72ºC7'4º	886	
HSD17B4	Exon 8	M13F8-HSD17B4 M13R8-HSD17B4	gcttttgataggtgcagt ccacaatgaacacaatttaca	94ºC10'35X(94ºC30''58,9ºC30''72ºC30'')72ºC7'4º	432	
WWOX	Exon 3	M13F3-WWOX M13R3-WWOX	gttgatgtgacaactgct ccatgcctggcctccctac	94ºC5'35X(94ºC30''61ºC30''72ºC30'')72ºC7'4º	425	
	Exon 9	M13F9-WWOX M13R9-WWOX	ggcctgctaatgcccaggca cctcaggctattctataacag	94ºC10'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	618	
STAR	Exon 1	M13F1-STAR M13R1-STAR	gcagaacaccaggtccaggct ccagtaagaggcacaact	94ºC5'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	316	
	Exon 4	M13F4-STAR M13R4-STAR	ggcgtgaaccaccatg gcattcacactgggctct	94ºC5'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	285	
AMH	Exon 2+3+4	M13F2-AMH M13R4-AMH	ggtgtggccttcaatggct ggaggagttgagagcggcca	94ºC10'35X(94ºC30''64,7ºC30''72ºC30'')72ºC7'4º	794	DMSO (10%)
AMHR2	Exon 3+4	M13F3-AMHR2 M13R4-AMHR2	ggaagggacgcctctga ggaatcaggctatagagat	94ºC5'35X(94ºC30''63,8ºC30''72ºC30'')72ºC7'4º	757	
MAP3K1	Exon 1	M13F1	ggcgcaggaggccttacgct	94ºC10'35X(94ºC30''58ºC30''72ºC30'')72ºC7'4º	696	DMSO (10%)

	Exon 8+9	M13R1 M13F8 M13R9	cgtagcctcgctgcgctga ggttccttctatataaattct ggaggcaagtaaaagtgt	94ºC10'35X(94ºC30''58ºC30''72ºC30'')72ºC7'4º	577	DMSO (10%)
	Exon 13	M13F13-MAP3K1 M13R13-MAP3K1	ggccttacaccatttaatca cctattcatgctctgagt	94ºC5'35X(94ºC30''60ºC30''72ºC30'')72ºC7'4º	380	
DMRT2	Exon 4, Amplicon A	M13F4A-DMRT2 M13R4A-DMRT2	ggatatctttcaccctccca cctgatgaccctgaaggtca	94ºC10'35X(94ºC30''56,4ºC30''72ºC30'')72ºC7'4º	857	
	Exon 4, Amplicon B	M13F4B-DMRT2 M13R4B-DMRT2	cctctaccagcaatgcct tcatcatctttacattaga	94ºC10'35X(94ºC30''56,4ºC30''72ºC30'')72ºC7'4º	742	
RXFP2	Exon 16	M13F16-RXFP2 M13R16-RXFP2	ggaaactgatgacataca ggacacagtcttgactatgt	94ºC5'35X(94ºC30''60ºC30''72ºC30'')72ºC7'4º	567	
GATA4	Exon 3	M13F3-GATA4 M13R3-GATA4	gctcaggggaactctcagt cctggatcattctggtggct	94ºC10'35X(94ºC30''58,7ºC30''72ºC30'')72ºC7'4º	551	
NR0B1	Exon 1, Amplicon 1A	M13F1A-NR0B1 R1A-NR0B1	ggtataaataggtcccagga agcatgctgtagaggatgc	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º	375	DMSO (10%)
	Exon 1, Amplicon 1B	F1B-NR0B1 R1B-NR0B1	gcatcctctacagcatgct cctgcgcttgatttgt	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º	441	DMSO (10%)
	Exon 1, Amplicon 1C	F1C-NR0B1 R1C-NR0B1	ctacttcgcgcagaggcc ggtggtgaggatcttct	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º	445	DMSO (10%)
	Exon 1, Amplicon 1D	F1D-NR0B1 M13R1D-NR0B1	gcagctggtgctggtgcgca cccgatgcttttgtgagct	94ºC5'35X(94ºC60''56ºC60''72ºC60'')72ºC7'4º	376	DMSO (10%)
MAMLD1	Exon 5, Amplicon 5A	M13F1A-MAMLD1 M13R1A-MAMLD1	ccaagggcgggaggtgaga ggctgcccagcaggctga	94ºC5'35X(94ºC30''61ºC30''72ºC30'')72ºC7'4º	734	
	Exon 5, Amplicon 5B	M13F1B-MAMLD1 M13R1B-MAMLD1	gcgttcagcagcactca ccactcaaattgtaacggga	94ºC5'35X(94ºC30''61ºC30''72ºC30'')72ºC7'4º	693	
ATRX	Exon 9, Amplicon 9A	M13F9A-ATRX M13R9A-ATRX	gagttttgttgggaata ctttatgctctttggt	94ºC10'35X(94ºC30''56,9ºC30''72ºC30'')72ºC7'4º	735	
	Exon 9, Amplicon 9B	M13F9B-ATRX M13R9B-ATRX	caaattgtaatggaga caggagtgagtttaaca	94ºC10'35X(94ºC30''56,9ºC30''72ºC30'')72ºC7'4º	813	
	Exon 9, Amplicon 9C	M13F9C-ATRX M13R9C-ATRX	gaagatttagacatgga gaagactcagactgggt	94ºC10'35X(94ºC30''56,9ºC30''72ºC30'')72ºC7'4º	914	
	Exon 9, Amplicon 9D	M13F9D-ATRX M13R9D-ATRX	tcaaaattcagattctga cagttgttccattctta	94ºC10'35X(94ºC30''56,9ºC30''72ºC30'')72ºC7'4º	960	

	Exon 9, Amplicon 9E	M13F9E-ATRX M13R9E-ATRX	ccttcagactttaaga gtagtaactcaagag	94ºC10'35X(94ºC30''56,9ºC30''72ºC30'')72ºC7'4º	854	
	Exon 27	M13F27 M13R27	cctctgaggtggcattct ccatgttgtaaacctc	94ºC10'35X(94ºC30''55ºC30''72ºC30'')72ºC7'4º	489	
	Exon 35	M13F35 M13R35	gatcttactaactggt gttctgttaagtcattga	94ºC10'35X(94ºC30''55ºC30''72ºC30'')72ºC7'4º	543	
CBX2	Exon 1+2	M13F1 M13R2	cgcacgggatttgggca ccgctctgcccgcgggc	94ºC10'35X(94ºC30''61,1ºC30''72ºC30'')72ºC7'4º	493	
DHH	Exon 3	M13F3 M13R3	ggaaaattgtgttgtccatga cctaagccaggcatagccccat	94ºC10'35X(94ºC30''51ºC30''72ºC30'')72ºC7'4º	873	DMSO (10%)
ZFPM2	Exon 8, Amplicon 8A	M13F8A M13R8A	ctctaggagtgaaaatgga ccttcccatttggccca	94ºC5'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	963	
	Exon 8, Amplicon 8B	M13F8B M13R8B	gctgctcattctgctgatcct ccttgcagggttgctaggt	94ºC5'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	995	
	Exon 8, Amplicon 8C	M13F8C M13R8C	ggcccacaagcagaatttct gataccaaaggtgactga	94ºC5'35X(94ºC30''65ºC30''72ºC30'')72ºC7'4º	877	
KISS1	Exon 3	M13F3 M13R3	ggatgggatgacaggaggt gcagaccacacgtcagtga	95ºC5'35X(95ºC30''58,9ºC30''72ºC30'')72ºC7'4º	549	
KISS1R	Exon 2	M13F2 M13R2	cactcggaccaaggtg ccagcttctgagtgcatct	94ºC5'35X(94ºC30''55,9ºC30''72ºC30'')72ºC7'4º	443	
CYP11B1	Exon 3+4	M13F3 M13R4	ccagggcccccagtcagc cctccattccccactggg	94ºC10'35X(94ºC30''61,1ºC30''72ºC30'')72ºC7'4º	750	
ESR2	Exon 2	M13F2 M13R2	gcaatgaagagatgaatga ctatgtaattaatacaat	94ºC10'35X(94ºC30''58ºC30''72ºC30'')72ºC7'4º	770	DMSO (10%)
	Exon 5	M13F5-ESR2 M13R5-ESR2	gctagtacggtcacgacca ggtatgccaagacagaagga	94ºC5'35X(94ºC30''63.8ºC30''72ºC30'')72ºC7'4º	604	
SOX3	Exon 1, Amplicon 1A	M13F1A-SOX3 M13R1A-SOX3	cctccgggttgcgaggggc ggcccagccgttcacgt	94ºC10'35X(94ºC30''53,2ºC30''72ºC30'')72ºC7'4º	895	
	Exon 1, Amplicon 1B	M13F1B-SOX3 M13R1B-SOX3	cgccgactggaaactgct ggcaagcaaagctaaaca	94ºC10'35X(94ºC30''53,2ºC30''72ºC30'')72ºC7'4º	916	
SOX9	Exon 3	M13F3-SOX9 M13R3-SOX9	gggagggtccccggagggt ggtgaaggtggagtaga	94ºC5'35X(94ºC30''60ºC30''72ºC30'')72ºC7'4º	802	DMSO (10%)
RET	Exon 10	F10-RET	ggaggctgagtgggctacgt	94ºC10'35X(94ºC30''60ºC30''72ºC60'')72ºC7'4º	240	

		R10-RET	caatttcctcccttgttggga		
	Exon 11	M13F11-RET M13R11-RET	ccatgaggcagagcata ggcagaacacaggcctcgt	94ºC10'35X(94ºC30''60ºC30''72ºC60'')72ºC7'4º	474
	Exon 13	M13F13-RET M13R13-RET	ctctctgtctgaacttgggc tcaccctgcagctggcctta	94ºC10'35X(94ºC45''64ºC45''72ºC45'')72ºC7'4º	239
	Exon 14	M13F14-RET M13R14-RET	cacccccttactcattgggt ggtgagccatagcatggca	94ºC10'35X(94ºC30''60ºC30''72ºC60'')72ºC7'4º	487
	Exon 15	F15-RET R15-RET	gtctcaccaggccgctac gtgcacctgggatccc	94ºC10'35X(94ºC45''64ºC45''72ºC45'')72ºC7'4º	289
	Exon 16	M13F16-RET M13R16-RET	ccagcactgatgagggatgt ccactacatgtataagggtgt	94ºC10'35X(94ºC45''64ºC45''72ºC45'')72ºC7'4º	673

M13F tail: tgtaaaacgacggccagt; M13R tail: caggaaacagctatgacc

PCR conditions and used primers in the CNV detection by QMPSF

Gene	Amplicon name	Primers Name	Primers sequence (5' to 3')	Conditions	Amplicon size (pb)	Others (DMSO…)
SRD5A2	Exon 1	M13F1-SRD5A2 M13R1FAM-SRD5A2	gcagcggccaccggcgagg /56-FAM/agcagggcagtgcgctgcact	94°C5'30X(94°45''63.5°45''72°45'')72°7'4°	376	
	Exon 2	M13F2-SRD5A2 M13R2FAM-SRD5A2	tgaatcctaacctttcctccc /56-FAM/agctgggaagtaggtgagaa	94°5'25X(94°45''57°45''72°45'')72°7'4°	253	
	Exon 3	M13F3-SRD5A2 M13R3FAM-SRD5A2	tgtgaaaaaagcaccacaatct /56-FAM/agcggaagagtgagagtctgg	94°5'25X(94°45''57°45''72°45'')72°7'4°	226	
	Exon 4	M13F4-SRD5A2 M13R4FAM-SRD5A2	tgattgaccttccgattctt /56-FAM/agcagaagaagaaagctacgt	94°5'25X(94°45''57°45''72°45'')72°7'4°	250	
	Exon 5	M13F5-SRD5A2 M13R5FAM-SRD5A2	tcagccactgctccattatat /56-FAM/agcagggcagtgcgctgcact	94°5'25X(94°45''57°45''72°45'')72°7'4°	184	
AR	Exon 2	F2-AR M13R2FAM-AR	gcctatttctgccattca /56-FAM/agccctgggccctgaaaggt	94°5'24X(94°20''54°20''72°20'')72°10'4°	288	DMSO (10%)

Genetic markers and primer pairs used to amplify by PCR the surroundings of the *NR0B1* gene.

Name	Location	Primers sequence (5' to 3')	Product size	PCR conditions
DXS1218	X:29155323-29155589	/5' 6-FAM/AAGACTAAGATTGTTTCAGTTTGTT GTGCTTTTGTAATTTTTACCCA	255	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º
DXS8039	X:30218272-30218424	/5' 6-FAM/AAACTAACAAACCTCTAGCCAG TTCATGTCCATGTGAACAG	147	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º
DXS1083	X:29977924-29,978,270	/5' 6-FAM/ATCACACTAGCTACATTTTAAGTTC TCTGGAAGTCTAAGTCAATATCATC	131	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º
DXS992	X:30750317-30750519	/5' 6-FAM/AAGAATGGGACTCCATTTCA GCTTATCCACTGGGACAGAA	201	94ºC5'35X(94ºC30''55,7ºC30''72ºC30'')72ºC7'4º

PCR, Polymerase Chain Reaction

Primers used in the site-directed mutagenesis.

Gene	Primers Name	Primers sequence (5' to 3')
GATA4	c.677T Forward	tcgcctccagagcagggtggacatagc
	c.677T Reverse	gctatgtccaccctgctctggaggcga
LHCGR	c.757C Forward	cattcagaggctaattgccacgtcaccctattctctaaaaaaattg
	c.757C Reverse	caatttttttagagaatagggtgacgtggcaattagcctctgaatg
	c.1660T Forward	gctacattaaaatttatttttgcagtttgaaacccagaattaatggctacc
	c.1660T Reverse	tggtagccattaattctgggtttcaaactgcaaaataaattttaatgtag
NR5A1	c.88T>A Forward	cttgcagctctcactcgtgagcagtccgt
	c.88T>A Reverse	acggactgctcacgagtgagagctgcaag
	c.902G>A Forward	gcagctcgctccagtagttctgcagcagc
	c.902G>A Reverse	gctgctgcagaactactggagcgagctgc
	c.71A>T Forward	gtgtccggctacctctacggactgctc
	c.71A>T Reverse	gagcagtccgtagaggtagccggacac

Supplementary data 7. Preparation of the solutions used in the different methods.

1 Agarose gel electrophoresis procedure

1.1 TAE 1X buffer: Dilute 80mL of TAE 50X buffer in 4L of bidistilled water.

1.2 5% agarose gel in TAE1X: In a 500mL bottle, mix 6g of agarose in 400mL of TAE1X. Melt in the microwave.

2 Targeted gene sequencing panel procedure

2.1 Fresh Melt-Off Solution: Mix 280µL Tween Solution and 40µL NaOH 1M.

3 aCGH procedure

3.1 10X Blocking Agent: Add 1350µL Milli-Q water to the vial containing the 10x aCGH.

3.2 Blocking Agent. Leave at RT for 60 minutes.

4 Capillary Sanger sequencing

4.1 3M Sodium Acetate: Weigh 24.6g CH3COONa and mix with 80ml Milli-Q water.Adjust the volume to 100ml Milli-Q water and mix.

4.2 125mM EDTA solution: Weight 4.65g EDTA and mix with 80ml Milli-Q water. Add NaOH solution to adjust the solution pH to 8.0. Adjust the volume to 100ml Milli-Q water and mix.

5 FISH procedure

5.1 1X PBS: Add 5 tablets of PBS to 1L of MilliQ water and mix.

5.2 57mM KCL: Add 559g KCL to 100mL MilliQ water and mix. Preheat at 37° before use.

5.3 Carnoy solution: Mix 10mL acetic acid and 30mL methanol.

5.4 Supplemented RPMI medium: Mix 4.5mL RPMI culture medium, 500µL FBS and 50 µL phytohemagglutinin.

5.5 20X SSC solution: Mix 132g 20X SCC in 400ml Milli-Q water. Adjust pH to 5.3. Add Milli-Q water to a final volume of 500mL and filter through a 0.45µm filtration unit.

5.6 2X SSC/0.1% NP-40 wash solution: Mix 10mL 20X SCC solution 85mL Milli-Q water. Add 100µL NP-40 and mix. Adjust pH to 7.0. Add Milli-Q water to a final volume of 100mL.

5.7 Denaturation solution: Mix 49mL formamide, 7mL 20X SCC and 14mL Milli-Q water. Adjust pH to 7-8.0.

5.8 Preparation of 0.4X SCC/0.3% NP-40 Wash solution: Mix 2mL 20X SCC, 90mL Milli-Q water and 300µL NP-40. Adjust pH to 7-8.0. Add milli-Q water to a final volume of 100mL.

6 Cleaning of the slides for FISH procedure

- Pour 50mL acetone 100%, hydrochloric acid 1%, ethanol 70% and Milli-Q water in four different Coplin jars.
- Insert a slide in 50mL acetone 100%.
- Transfer the slide to hydrochloric acid 1% for 10 minutes.
- Transfer the slide to ethanol 70% for 10 minutes.
- Transfer the slide to Milli-Q water for 5 minutes.
- Dry the slide at RT and then refresh at 4°C for 15 minutes. Slides could be stored at 4°C.

7 Transformation of E. coli competent cells

7.1 LB liquid medium: In a 1L autoclave bottle, add 25g LB broth powder to 1L ultrapure water. Autoclave on liquid cycle (121°C) for 15 minutes. Cool to 60°C for 10 minutes before use. Store at RT until use.

7.2 LB agar plates: In a 1L autoclave bottle, add 40g LB agar powder to 1L ultrapure water. Autoclave on liquid cycle (121°C) for 15 minutes. Cool to 60°C for 10 minutes before use. Add 300µL ampicillin (100mg/ml) and mix. Pour out the media into petri dishes. Cool to RT and store them upside room at 4ºC.

7.3 Transformation Storage Solution (TSS) Buffer: In a 200mL bottle mix 5mL PEG 50%, 25mL DMSO, 0,5mL MgCl2 1M, 0,5mL MgSO4 and 41,5mL bidistilled water for a total volume of 50mL..

7.4 KCM 5x Buffer: In a 200mL bottle mix 25mL KCl 1M, 7,5mL $CaCl_2$ 1M, 12,5mL $MgCl_2$ 1M and 5mL bidistilled water for a total volume of 50mL.

8 Cell lines and culture

8.1 Supplemented MEM media (10% FBS, 1% L-glutamin, 1% Pen/strep): Remove 50mL of the MEM media bottle. Add 50mL FBS, 5mL L-glutamin and 5mL Penicillin-Streptomycin to the MEM bottle. Mix by inverting several times.

8.2 Supplemented DMEM media (10% FBS, 1% Pen/strep, 1% Sodium pyruvate): Remove 50mL of the DMEM media bottle. Add 50mL FBS, 5mL Penicillin-Streptomycin and 5mL sodium pyruvate. Mix by inverting several times.

9 Western-blot

9.1 Protease Inhibitor cocktail: In a 1.5mL Eppendorf dissolve 1 tablet in 1mL bidistilled water.

9.2 Protein Lysis Buffer 1X: In a 1L bottle mix 146mg EDTA, 1230mg Tris, 4350mg NaCl and 5g Triton X-100. Add bidistilled water for a total volume of 50mL.

9.3 MOPS 1X Buffer: Reconstitute the bag of powder in 1000 ml bidistilled water to make 1X running buffer.

9.4 Transfer Buffer: Reconstitute the bag of powder in 900 ml bidistilled water and 100mL ethanol.

9.5 PBS buffer: In a 1L bottle, prepare 800mL bidistilled water and add 8 g of NaCl, 200mg KCl, 1.44g Na2HPO4, 240 mg KH2PO4 and bidistilled water up to 1L

9.6 Blocking buffer: In a 100mL bottle mix 2,5g milk with 50mL PBS buffer.

9.7 Washing buffer: In a 200mL bottle mix 0.05mL Tween20 with 100mL PBS buffer.

9.8 Primary antibody dilution: In a 15mL falcon mix 5mL washing buffer with 0.05g BSA.

9.9 Monoclonal Anti-HA antibody produced in mouse: Add 1mL bidistilled water to a final concentration of 100µg/mL.

9.10 IRDye® 800CW Goat anti-Mouse IgG (H + L): Add 1mL bidistilled water to a final concentration of 100µg/mL.

9.11 IRDye® 680RD Donkey anti-Mouse IgG (H + L): Add 1mL bidistilled water to a final concentration of 100µg/mL.

9.12 0.5X Odyssey Blocking Buffer: In a 50mL falcon mix 10mL Odyssey Blocking Buffer 1X with 10mL bidistilled water.

9.13 Tris buffered saline with Tween® 20 (TBS-T) 1X Buffer: Dissolve 1 tablet in 500 mL bidistilled water.

Supplementary data 8. Coverage and localization of the regions amplified in the targeted gene panel.

	Chromosome			Bases				
Gene	**Number**	**Start**	**End**	**Total**	**Covered**	**Missed**	**Coverage (%)**	**Number of amplicons**
AMH.1	chr19	2249310	2249746	436	418	18	95.9	2
AMH.2	chr19	2250336	2250478	142	142	0	100	1
AMH.3	chr19	2250649	2250760	111	111	0	100	1
AMH.4	chr19	2250840	2251007	167	151	16	90.4	2
AMH.5	chr19	2251098	2252072	974	542	432	55.6	3
AMHR2	chr12	53817613	53825343	2413	2413	0	100	14
AR	chrX	66763848	66950486	11288	11288	0	100	43
ATRX	chrX	76760330	77041780	12970	12925	45	99.7	66
BMP15	chrX	50653709	50659666	1362	1362	0	100	6
CBX2	chr17	77751951	77761474	5246	5235	11	99.8	24
CYP11A1	chr15	74630077	74660106	2928	2928	0	100	17
CYP11B1.1	chr8	143960991	143961262	271	271	0	100	2
CYP11B1.2	chr8	143960441	143960603	162	162	0	100	2
CYP11B1.3	chr8	143958423	143958639	216	168	48	77.8	2
CYP11B1.4	chr8	143958093	143958301	208	208	0	100	2
CYP11B1.5	chr8	143957655	143957818	163	163	0	100	2
CYP11B1.6	chr8	143957122	143957294	172	172	0	100	2
CYP11B1.7	chr8	143956646	143956728	82	82	0	100	2
CYP11B1.8	chr8	143956365	143956570	205	205	0	100	3
CYP11B1.9	chr8	143954670	143955907	1237	1114	123	90.1	6
CYP17A1	chr10	104590262	104597315	2270	2270	0	100	12
CYP19A1	chr15	51500228	51630820	5064	5064	0	100	24
CYP21A2.1	chr6	32006042	32006401	359	359	0	100	3
CYP21A2.2	chr6	32006497	32006593	96	96	0	100	2
CYP21A2.3	chr6	32006858	32007025	167	167	0	100	2
CYP21A2.4	chr6	32007133	32007234	101	101	0	100	1
CYP21A2.5	chr6	32007315	32007424	109	109	0	100	2
CYP21A2.6	chr6	32007522	32007616	94	94	0	100	1
CYP21A2.7	chr6	32007782	32007988	206	206	0	100	1
CYP21A2.8	chr6	32008183	32008362	179	179	0	100	1
CYP21A2.9	chr6	32008445	32008549	104	104	0	100	2
CYP21A2.10	chr6	32008637	32009447	810	662	148	81.7	5
DHH	chr12	49483180	49488627	2084	2084	0	100	11
DMRT1	chr9	841664	969115	2472	2469	3	99.9	12
DMRT2	chr9	1050328	1057579	3135	3116	19	99.4	15
ESR1	chr6	152011605	152424433	7568	7568	0	100	32
ESR2	chr14	64693725	64805293	4519	4501	18	99.6	22
FGF9	chr13	22245189	22278665	4680	4525	155	96.7	19
FOXL2	chr3	138663040	138666007	2967	2863	104	96.5	13
FOXO3.1	chr6	108882069	108883032	963	528	435	54.8	3
FOXO3.2	chr6	108984658	108986092	1434	1434	0	100	6
FOXO3.3	chr6	109001030	109005971	4941	4941	0	100	18
FSHR	chr2	49189270	49381691	3274	3274	0	100	17
GATA4.1	chr8	11561713	11561813	100	100	0	100	1
GATA4.2	chr8	11565365	11566437	1072	673	399	62.8	3
GATA4.3	chr8	11606428	11606602	174	174	0	100	1
GATA4.4	chr8	11607620	11607750	130	130	0	100	1
GATA4.5	chr8	11612555	11612650	95	95	0	100	1

GATA4.6	chr8	11614440	11614596	156	156	0	100	1
GATA4.7	chr8	11615802	11617511	1709	1686	23	98.7	7
HARS2	chr5	140070985	140078928	3260	3260	0	100	17
HSD17B3	chr9	98997563	99064459	1684	1684	0	100	12
HSD17B4	chr5	118788112	118878055	4188	4188	0	100	30
HSD3B2	chr1	119957528	119965687	2099	2099	0	100	9
INHA	chr2	220436928	220440460	1531	1531	0	100	6
INSL3	chr19	17927296	17932408	1050	1050	0	100	8
KISS1	chr1	204159443	204165644	864	864	0	100	4
KISS1R	chr19	917316	921040	1875	1842	33	98.2	12
LHCGR	chr2	48913887	48982905	3643	3616	27	99.3	20
MAMLD1	chrX	149531519	149682473	5260	5238	22	99.6	24
MAP3K1.1	chr5	56111401	56111882	481	453	28	94.2	3
MAP3K1.2	chr5	56152427	56152577	150	150	0	100	1
MAP3K1.3	chr5	56155534	56155742	208	208	0	100	2
MAP3K1.4	chr5	56160561	56160761	200	197	3	98.5	1
MAP3K1.5	chr5	56161167	56161283	116	116	0	100	1
MAP3K1.6	chr5	56161656	56161804	148	148	0	100	2
MAP3K1.7	chr5	56167737	56167858	121	121	0	100	1
MAP3K1.8	chr5	56168468	56168549	81	81	0	100	2
MAP3K1.9	chr5	56168652	56168832	180	180	0	100	3
MAP3K1.10	chr5	56170859	56171137	278	278	0	100	2
MAP3K1.11	chr5	56174807	56174928	121	121	0	100	1
MAP3K1.12	chr5	56176538	56176629	91	91	0	100	1
MAP3K1.13	chr5	56176908	56177099	191	191	0	100	1
MAP3K1.14	chr5	56177397	56178693	1296	1296	0	100	5
MAP3K1.15	chr5	56179354	56179506	152	152	0	100	1
MAP3K1.16	chr5	56180491	56180653	162	162	0	100	1
MAP3K1.17	chr5	56181759	56181890	131	116	15	88.5	1
MAP3K1.18	chr5	56183205	56183347	142	142	0	100	1
MAP3K1.19	chr5	56184053	56184184	131	131	0	100	1
MAP3K1.20	chr5	56189358	56191979	2621	2621	0	100	10
NR0B1	chrX	30322513	30327520	1685	1685	0	100	7
NR5A1	chr9	127243489	127269724	3445	3417	28	99.2	20
POR.1	chr7	75544397	75544497	100	100	0	100	1
POR.2	chr7	75583307	75583498	191	191	0	100	1
POR.3	chr7	75601731	75601779	48	48	0	100	1
POR.4	chr7	75608767	75608899	132	132	0	100	2
POR.5	chr7	75609645	75609806	161	161	0	100	1
POR.6	chr7	75609807	75610365	558	558	0	100	3
POR.7	chr7	75610366	75610490	124	105	19	84.7	1
POR.8	chr7	75610829	75610926	97	97	0	100	1
POR.9	chr7	75611540	75611642	102	102	0	100	2
POR.10	chr7	75612838	75612961	123	123	0	100	2
POR.11	chr7	75613056	75613174	118	96	22	81.4	1
POR.12	chr7	75614095	75614375	280	280	0	100	3
POR.13	chr7	75614376	75614527	151	121	30	80.1	1
POR.14	chr7	75614897	75615168	271	271	0	100	2
POR.15	chr7	75615237	75615394	157	157	0	100	2
POR.16	chr7	75615473	75615559	86	86	0	100	2
POR.17	chr7	75615560	75615654	94	94	0	100	1
POR.18	chr7	75615655	75616173	518	489	29	94.4	3
PSMC3IP	chr17	40724302	40729874	2380	2380	0	100	14

RSPO1	chr1	38076925	38100620	3539	3535	4	99.9	17
RXFP2	chr13	32313653	32377034	3703	3703	0	100	21
SOX3	chrX	139585126	139587250	2124	2124	0	100	10
SOX9	chr17	70117135	70122585	4084	4084	0	100	18
SRD5A2	chr2	31749630	31806065	2695	2695	0	100	12
SRY	chrY	2654870	2655807	937	937	0	100	4
STAR	chr8	38000192	38008625	3045	3013	32	98.9	15
TSPYL1	chr6	116595996	116601305	5309	5223	86	98.4	20
WNT4	chr1	22443772	22469544	4155	3973	182	95.6	20
WT1	chr11	32409296	32457106	3864	3864	0	100	22
WWOX	chr16	78133301	79246589	3373	3373	0	100	19
ZFPM2	chr8	106331121	106816792	4880	4880	0	100	21
Mean							98.5	874

Chr, chromosome

Supplementary data 9. Targeted regions not covered according to the Ion Ampliseq Designer tool.

Gene	*Locus*	Missed bases (bp)	Targeted region in the gene
AMH	Chr19:2249310-2249746	18	Exon 1-Intron 1
AMH	Chr19:2250840-2251007	16	Exon 4
AMH	Chr19:2251098-2252072	432	Exon 5
ATRX	ChrX:76761410-76761454	45	Exon 35 NC
CBX2	Chr17:77751952-77751962	11	Exon 1 NC
CYP11B1	Chr8:143958423-143958639	48	Exon 3
CYP11B1	Chr8:143954670-143955907	123	Exon 9 NC
CYP21A2	Chr8:32008637-32009447	148	Exon 10
DMRT1	Chr9:842215-842217	3	Intron 1
DMRT2	Chr9:1052002-1052020	19	Exon 2
ESR2	Chr14:64805031-64805048	18	5'UTR
FGF9	Chr13:22245393-22245547	155	5'UTR
FOXL2	Chr3:138664143-138664246	104	Exon 1 NC
FOXO3	Chr6:108882069-108883032	435	Exon 1
GATA4	Chr8:11565365-11566437	399	Exon 2
GATA4	Chr8:11615802-11617511	23	Exon 7 NC
KISS1R	Chr19:917503-917518	16	Exon 1
KISS1R	Chr19:919984-919990	7	Exon 4
KISS1R	Chr19:920265-920274	10	Exon 5
LHCGR	Chr2:48925927-48925953	27	Intron 8-Exon 9
MAMLD1	ChrX:149680802-149680823	22	Exon 5
MAP3K1	Chr5:56111401-56111882	28	Exon 1
MAP3K1	Chr5:56160561-56160761	3	Exon 4-intron 4
MAP3K1	Chr5:56181759-56181890	15	Exon 17-intron 17
NR5A1	Chr9:127243914-127243941	28	Exon 9 NC
POR	Chr7:75610366-75610490	19	Exon 6
POR	Chr7:75613056-75613174	22	Exon 10-intron 10
POR	Chr7:75614376-75614527	30	Exon 12
POR	Chr7:75615655-75616173	29	Exon 16
RSPO1	Chr1:38100365-38100368	4	Exon 1 NC
STAR	Chr8:38000779-38000789+ Chr8:38001112-38001132	32	3'UTR
TSPYL1	Chr6:116596977-116597001+ Chr6:116599093-116599153	86	Exon 1 NC, 3'UTR
WNT4	Chr1:22469363-22469544	182	5'UTR-exon 1

Chr, chromosome; NC, non-coding; UTR, untranslated region

Supplementary data 10. Example of the quality values achieved after the NGS.

Key quality metrics achieved in the patients harbouring a genetic variant.

Case	Mapped reads	On target (%)	Coverage depth (Average)	Uniformity (%)	Base coverage at 20x (%)	Case	Mapped reads	On target (%)	Coverage depth (Average)	Uniformity (%)	Base coverage at 20x (%)
GN0011	423106	93.1	374.4	84.4	94.2	GN0158	226113	95.3	204.1	90.0	93.5
GN0020	348581	95.0	306.0	89.4	94.6	GN0163	260668	95.1	243.1	89.2	93.7
GN0028	201246	93.6	195.0	89.6	93.1	GN0164	282266	94.9	260.5	89.8	94.1
GN0034	176079	94.1	172.3	87.6	89.8	GN0171	243138	92.6	232.5	89.7	93.3
GN0046	258515	94.5	253.7	88.8	93.6	GN0182	255763	95.0	238.4	89.6	93.7
GN0051	183729	95.8	169.7	90.5	93.8	GN0185	321454	95.9	302.1	89.8	94.4
GN0070	227562	94.8	198.7	89.6	93.4	GN0186	292495	95.0	274.0	89.9	94.1
GN0075	345109	95.4	304.4	90.4	94.9	GN0187	243457	95.1	227.6	89.9	93.5
GN0090	235797	95.1	216.5	90.7	94.6	GN0189	239962	95.2	222.1	89.7	93.4
GN0096	197700	94.4	178.5	90.2	93.6	GN0194	265677	95.2	249.2	89.8	93.8
GN0118	160389	95.1	149.0	90.6	92.9	GN0198	220407	95.8	202.7	87.2	91.3
GN0142	385663	94.7	338.6	90.1	94.9	GN0199	453904	95.3	476.4	88.5	94.7
GN0147	297298	94.9	264.2	89.1	93.8	GN0203	548956	95.2	575.7	89.6	94.7
GN0150	223689	94.7	205.4	87.9	94.1	GN0207	416516	95.6	434.7	88.4	94.2
GN0154	316588	92.7	304.8	59.6	94.5	OT0567	359366	96.0	370.3	89.8	95.3
GN0155	133276	93.7	118.5	87.9	89.8	RE0045	113944	95.0	104.5	89.1	89.6
GN0156	214614	94.9	197.9	89.3	92.9	**Mean**	273756.7	94.7	258.3	88.4	93.5
GN0157	234703	93.9	216.5	90.2	93.1						

Quality metrics of the positive controls included in the panel.

Case	Run	Mapped reads	On target (%)	Coverage depth (Average)	Uniformity (%)	Base coverage at 20x (%)
GN0012	Run 8	339833	94.9	316.7	89.2	94.3
GN0038	Run 2	160825	94.3	145.6	90.5	92.6
GN0038	Run 5	411785	95.2	404.4	89.6	94.9
GN0041	Run 3	231384	95.0	201.9	87.4	92.6
GN0041	Run 6	284439	91.7	270.2	89.9	94.1
GN0042	Run 4	291990	92.7	280.3	90.0	94.6
GN0109	Run 3	166327	95.9	148.8	89.6	91.9
GN0111	Run 9	315034	95.6	328.6	88.4	93.9
GN0119	Run 10	376231	95.7	406.6	90.0	94.8
GN0123	Run 10	452078	95.1	485.7	89.7	95.3
GN0141	Run 1	417085	94.9	384.9	88.6	95.2
OT0327	Run 4	256903	94.1	250.2	89.8	93.5
OT0327	Run 5	364178	95.0	363.1	89.0	94.3
POL0285	Run 6	345955	84.8	302.1	89.1	93.5
Sane controls	Run 1	218436	94.8	201.9	91.2	94.9
Mean		308832	93.9	299.4	89.5	94.0

Supplementary data 11. Amplicons not covered in most of the samples.

Amplicon	Gene	*Locus*	Targeted region in the gene
AMPL7155526017	*AR*	chrX: 66764106-66764373	5'UTR
AMPL7154768614	*AR*	chrX: 66766558-66766723	Exon 1, intron 1
AMPL7155525439	*AR*	chrX: 66948258-66948258	Exon 8 NC
AMPL7155525424	*AR*	chrX:66944026-66944026	Exon 8 NC
AMPL7154768613	*AR*	chrX:66766308-66766568	Exon 1
AMPL7155435957	*ATRX*	chrX:76763526-76763843	Exon 35, exon 35 NC
AMPL7156609654	*ATRX*	chrX: 76937989-76938290	Exon 9
AMPL7154463818	*ATRX*	chrX:76845183-76845498	Intron 26 to intron 27
AMPL7154592293	*CBX2*	chr17:77752094-77752403	Exon 1 to intron 2
AMPL7158539628	*CYP11A1*	chr15:74657976-74658296	Intron 1
AMPL7154501291	*CYP11B1*	chr8:143958112-143958434	Intron 3, exon 4
AMPL7156017306	*CYP21A2*	chr6:32005824-32006152	5'UTR, exon 1 NC
AMPL7158500645	*DHH*	chr12:49483672-49483813	Exon 3
AMPL7155128017	*ESR2*	chr14:64749208-64749531	Exon 2, intron 2
AMPL7156599351	*ESR2*	chr14:64805252-64805452	5'UTR
AMPL7154373786	*FOXO3*	chr6:108881843-108882156	5'UTR, exon 1 NC
AMPL7158612785	*FOXO3*	chr6:109005657-109005971	Exon 3 NC
AMPL7154859626	*FSHR*	chr2:49190282-49190576	Exon 10
AMPL7157052817	*KISS1*	chr1:204159383-204159693	Exon 3, 3'UTR
AMPL7157261487	*KISS1R*	chr19:918338-918653	Intron 1, exon 2
AMPL7157052823	*KISS1R*	chr19:920817-920968	Exon 5 NC

Amplicon	Gene	*Locus*	Targeted region in the gene
AMPL7155127628	*LHCGR*	chr2:48914044-48914364	Exon 1 NC
AMPL7153405652	*MAMLD1*	chrX:149671675-149671905	Intron 3
AMPL7158500630	*MAMLD1*	chrX:149681097-149681375	Exon 5
AMPL7154592236	*MAP3K1*	chr5:56111784-56112084	Exon 1, intron 1
AMPL7154592239	*MAP3K1*	chr5:56111315-56111589	5'UTR, exon 1
AMPL7153033277	*MAP3K1*	chr5:56168711-56168877	Exon 9, intron 9
AMPL7154591595	*NR5A1*	chr9:127265626-127265961	Intron 1, exon 2
AMPL7157234687	*RSPO1*	chr1:38077119-38077412	Exon 7 NC, 3'UTR
AMPL7157234693	*RSPO1*	chr1:38097814-38098128	Intron 1 to intron 2
AMPL7154809599	*SOX3*	chrX:139585781-139586094	Exon 1, exon 1 NC
AMPL7154809601	*SOX3*	chrX:139586276-139586598	Exon 1
AMPL7158500104	*SOX9*	chr17: 70116917-70117240	5'UTR, exon 1 NC
AMPL7154592302	*SOX9*	chr17:70120012-70120345	Exon 3
AMPL7157408783	*SOX9*	chr17:70119859-70120179	Exon 3
AMPL7158539633	*TSPYL1*	chr6:116596344-116596668	3'UTR
AMPL7158539637	*TSPYL1*	chr6:116597283-116597577	3'UTR
AMPL7153146546	*WNT4*	chr1:22469178-22469362	Exon 1, intron 1
AMPL7153218849	*WT1*	chr11:32456878-32457093	Exon 1 NC, exon 1
AMPL7154427779	*WT1*	chr11:32456429-32456759	Exon 1
AMPL7156733712	*ZFPM2*	chr8:106813199-106813518	Intron 7, exon 8

Chr, chromosome; NC, non-coding; UTR, untranslated region

Supplementary data 12. Regions amplified by PCR after NGS.

Gene	*Locus*	Targeted region in the gene
AMH	Chr19:2249310-2249746	Exon 1
AMH	Chr19:2250840-2251007	Exon 4
AMH	Chr19:2251098-2252072	Exon 5
AR	chrX: 66766558-66766723	Exon 1
AR	chrX:66766308-66766568	Exon 1
ATRX	chrX:76763526-76763843	Exon 35
ATRX	chrX: 76937989-76938290	Exon 9
ATRX	chrX:76845183-76845498	Exon 27
CBX2	chr17:77752094-77752403	Exon 1, exon 2
CYP11B1	Chr8:143958423-143958639	Exon 3
CYP11B1	chr8:143958112-143958434	Exon 4
CYP21A2	Chr8:32008637-32009447	Exon 10
DHH	chr12:49483672-49483813	Exon 3
DMRT2	Chr9:1052002-1052020	Exon 2
ESR2	chr14:64749208-64749531	Exon 2
FOXO3	Chr6:108882069-108883032	Exon 1
FSHR	chr2:49190282-49190576	Exon 10
GATA4	Chr8:11565365-11566437	Exon 2
KISS1	chr1:204159383-204159693	Exon 3
KISS1R	Chr19:917503-917518	Exon 1
KISS1R	Chr19:919984-919990	Exon 4
KISS1R	Chr19:920265-920274	Exon 5
KISS1R	chr19:918338-918653	Exon 2

Gene	*Locus*	Targeted region in the gene
LHCGR	Chr2:48925927-48925953	Exon 9
MAMLD1	ChrX:149680802-149680823	Exon 5
MAMLD1	chrX:149681097-149681375	Exon 5
MAP3K1	Chr5:56111401-56111882	Exon 1
MAP3K1	Chr5:56160561-56160761	Exon 4
MAP3K1	Chr5:56181759-56181890	Exon 17
MAP3K1	chr5:56111784-56112084	Exon 1
MAP3K1	chr5:56111315-56111589	Exon 1
MAP3K1	chr5:56168711-56168877	Exon 9
NR5A1	chr9:127265626-127265961	Exon 2
POR	Chr7:75610366-75610490	Exon 6
POR	Chr7:75613056-75613174	Exon 10
POR	Chr7:75614376-75614527	Exon 12
POR	Chr7:75615655-75616173	Exon 16
SOX3	chrX:139585781-139586094	Exon 1
SOX3	chrX:139586276-139586598	Exon 1
SOX9	chr17:70120012-70120345	Exon 3
SOX9	chr17:70119859-70120179	Exon 3
WNT4	Chr1:22469363-22469544	Exon 1
WNT4	chr1:22469178-22469362	Exon 1
WT1	chr11:32456878-32457093	Exon 1
WT1	chr11:32456429-32456759	Exon 1
ZFPM2	chr8:106813199-106813518	Exon 8

Chr, chromosome; NC, non-coding; UTR, untranslated region

Supplementary data 13. Detected number of variants detected by the Ion Reporter software.

	SNVs and indels					CNVs and others			
Case	**Variant calling**	**p-value < 0.001, MAF <0.05**	**Variants present in NC**	**Total**	**Clinical feature**	**Variant calling**	**MAPD**	**Confident germline CNV**	**Clinical feature**
GN0001	287	54	15	39	0	301	0.131	1	0
GN0003	270	40	15	25	0	276	0.114	0	0
GN0011	298	52	10	42	1	341	0.251	5	0
GN0012	255	46	8	38	0	284	0.256	6	0
GN0013	264	47	13	34	0	272	0.226	0	0
GN0014	290	40	15	25	0	312	0.421	0	0
GN0017	270	50	18	32	0	429	0.355	36	0
GN0020	245	34	10	24	1	261	0.085	0	0
GN0025	265	54	15	39	1	446	0.427	55	0
GN0026	266	42	10	32	0	285	0.236	1	0
GN0027	240	37	9	28	0	256	0.148	1	0
GN0028	232	40	13	27	2	250	0.132	0	0
GN0029	235	38	11	27	0	240	0.107	0	0
GN0031	289	40	11	29	0	299	0.149	0	0
GN0033	284	41	10	31	0	313	0.257	4	0
GN0034	217	35	10	25	0	237	0.216	1	1
GN0038	259	44	14	30	1	261	0.244	1	0
	252	44	11	33	1	261	0.119	0	0
GN0039	246	48	13	35	0	260	0.13	0	0
GN0041	248	39	9	30	0	258	0.132	1	0
	249	40	12	28	0	263	0.194	0	0
GN0042	282	54	13	41	2	291	0.134	0	0
GN0043	253	37	13	24	0	484	0.647	14	0
GN0046	255	42	13	29	1	276	0.157	1	0
GN0050	287	49	17	32	0	301	0.207	1	0
GN0051	271	69	12	57	0	283	0.228	0	0
GN0056	237	42	12	30	0	338	0.402	32	0
GN0059	263	49	10	39	0	278	0.149	1	0
GN0064	246	52	15	37	0	304	0.307	10	0
GN0066	546	35	10	25	0	262	0.088	0	0
GN0070	315	63	10	53	0	326	0.113	0	0
GN0075	266	33	13	20	1	280	0.114	0	0
GN0084	266	45	10	35	0	277	0.147	0	0
GN0090	335	100	11	89	0	346	0.236	1	0
GN0096	325	86	11	75	0	333	0.202	0	0
GN0100	262	38	12	26	0	266	0.27	1	0
GN0103	257	44	10	34	0	264	0.207	0	0
GN0104	253	45	11	34	0	265	0.135	0	0
GN0108	266	53	15	38	0	269	0.242	0	0
GN0109	274	39	8	31	1	258	0.104	0	0
GN0111	268	44	13	31	1	295	0.232	3	0
GN0114	235	46	12	34	0	238	0.217	1	0

GN0118	294	74	18	56	0	295	0.157	0	0
GN0119	269	68	8	60	0	319	0.199	7	0
GN0122	238	41	13	28	0	257	0.207	0	0
GN0123	250	45	13	32	1	284	0.189	3	0
GN0124	245	39	8	31	0	255	0.179	0	0
GN0138	280	47	17	30	0	291	0.19	0	0
GN0141	269	43	17	26	1	310	0.323	8	0
GN0142	262	52	11	41	1	277	0.087	0	0
GN0144	272	52	11	41	0	300	0.255	7	0
GN0145	271	45	9	36	0	290	0.123	2	0
GN0147	262	46	10	36	1	282	0.15	1	0
GN0148	284	54	15	39	0	494	0.433	80	0
GN0150	258	49	16	33	1	310	0.409	9	0
GN0151	239	43	11	32	0	251	0.143	1	0
GN0152	283	41	11	30	0	294	0.141	2	0
GN0154	253	40	13	27	1	264	0.124	0	0
GN0155	268	52	12	40	1	266	0.224	0	0
GN0156	284	50	10	40	1	309	0.241	4	0
GN0157	296	74	10	64	1	320	0.248	3	0
GN0158	338	86	14	72	0	334	0.266	2	0
GN0160	257	46	12	34	0	267	0.131	0	0
GN0162	281	37	11	26	0	304	0.223	4	0
GN0163	234	33	7	26	0	264	0.259	6	0
GN0164	234	42	8	34	1	268	0.265	3	0
GN0167	248	39	7	32	0	270	0.251	5	0
GN0169	272	45	11	34	0	283	0.21	1	0
GN0171	260	47	13	34	2	270	0.213	0	0
GN0173	268	35	7	28	0	286	0.244	5	0
GN0174	255	44	13	31	0	269	0.206	0	0
GN0175	275	47	9	38	0	323	0.212	4	0
GN0176	255	43	12	31	0	266	0.223	0	0
GN0178	256	42	17	25	0	261	0.222	0	0
GN0179	280	51	10	41	0	297	0.205	1	0
GN0182	298	42	10	32	0	328	0.249	5	0
GN0183	252	39	9	30	0	286	0.173	6	0
GN0185	242	40	10	30	1	265	0.24	3	0
GN0186	247	50	11	39	1	270	0.258	4	0
GN0187	249	43	10	33	0	261	0.243	2	1
GN0189	265	48	10	38	1	285	0.257	3	0
GN0190	288	38	10	28	0	313	0.25	4	0
GN0191	240	32	10	22	0	273	0.244	4	0
GN0192	268	43	11	32	0	308	0.196	7	0
GN0194	258	46	13	33	1	281	0.259	5	0
GN0195	279	34	10	24	0	316	0.257	6	0
GN0196	290	45	10	35	0	214	0.172	9	0
GN0198	265	61	13	48	2	286	0.336	0	0
GN0199	284	44	10	34	0	323	0.205	7	0
GN0200	259	37	10	27	0	293	0.19	8	0

GN0201	261	38	9	29	0	296	0.185	4	0
GN0202	226	31	9	22	0	253	0.203	3	0
GN0203	326	61	12	49	1	355	0.18	4	0
GN0204	243	37	11	26	0	283	0.234	9	0
GN0205	273	43	11	32	0	322	0.198	7	0
GN0207	283	51	13	38	1	339	0.23	7	0
OT0327	264	38	11	27	0	270	0.106	1	0
	272	47	10	37	0	290	0.149	1	0
OT0567	275	54	8	46	0	356	0.278	12	1
POL0285	242	34	12	22	0	264	0.189	1	1
RE0045	274	50	15	35	1	276	0.16	0	0
NC	308			0		310	0.158		0
Mean	269.0	46.4	11.5	34.5	0.3	292.9	0.21	4.5	0.04
Total	27444	4688	1165	3523	34	29885		457	4

NC, negative control.

Supplementary data 14. Alignment of multiple protein sequences (A, SRY; B, SF1; C, WT1; D, GATA4; E, WWOX; F, ESR1 and ESR2; G, DMRT2; H, ZFPM2; I, MAMLD1; .J, AR; K, LHCGR; L, STAR; M, AMH) across species. The altered amino acids in the different proteins are given in bold.

A

aa position at SRY							**131**						
Hsapiens	P	N	Y	K	Y	R	**P**	R	R	K	A	K	M
Ptroglodytes	P	N	Y	K	Y	R	**P**	R	R	K	A	N	M
Mmulatta	No homologue												
Fcatus	No homologue												
Mmusculus	No homologue												
Ggallus	P	D	Y	K	Y	R	**P**	R	R	K	T	K	T
Trubripes	No homologue												
Drerio	P	D	Y	K	Y	R	**P**	R	R	K	T	KT	L
Dmelanogaster	P	D	Y	K	Y	R	**P**	R	R	K	T	K	T
Xtropicalis	P	D	Y	K	Y	R	**P**	R	R	K	T	KT	L

B

aa position at SF1							**24**					**30**											**301**								
Hsapiens	D	K	V	S	G	Y	**H**	Y	G	L	L	T	**C**	S	C	K	G	F	…	M	T	L	L	Q	N	**C**	W	S	E	L	L	V
Ptroglodytes	No homologue																		…	No homologue												
Mmulatta	D	K	V	S	G	Y	**H**	Y	G	L	L	T	**C**	S	C	K	G	F	…							**C**	W	S	E	L	L	V
Fcatus	No homologue																		…	M	T	L	L	Q	N	**C**	W	S	E	L	L	V
Mmusculus	D	K	V	S	G	Y	**H**	Y	G	L	L	T	**C**	S	C	K	G	F	…						N	**C**	W	S	E	L	L	V
Ggallus	No homologue																		…	M	K	L	L	Q	N	**C**	W	S	E	L	L	V
Trubripes	No homologue																		…	No homologue												
Drerio	D	K	V	S	G	Y	**H**	Y	G	L	L	T	**C**	S	C	K	G	F	…	M	R	L	L	H	N	**C**	W	S	E	L	L	L
Dmelanogaster	D	K	V	S	G	Y	**H**	Y	G	L	L	T	**C**	S	C	K	G	F	…	M	K	L	L	Q	H	**S**	W	S	D	M	L	V
Xtropicalis	D	K	V	S	G	Y	**H**	Y	G	L	L	T	**C**	S	C	K	G	F	…	No homologue												

C

aa position at WT1								**75**											**182**											
Hsapiens	Q	Q	G	S	L	G	**E**	Q	Q	Y	S	V	P	…	F	K	H	E	D	P	**M**	G	Q	Q	G	S	L
Ptroglodytes	Q	Q	G	S	L	G	**E**	Q	Q	Y	S	V	P	…	F	K	H	E	D	P	**M**	G	Q	Q	G	S	L
Mmulatta	Q	Q	G	S	L	G	**E**	Q	Q	Y	S	V	P	…	F	K	H	E	D	P	**M**	G	Q	Q	G	S	L
Fcatus	No homologue													…	No homologue																
Mmusculus	Q	Q	G	S	L	G	**E**	Q	Q	Y	S	V	P	…	F	K	H	E	D	P	**M**	G	Q	Q	G	S	L
Ggallus	Q	Q	P	S	L	G	**D**	Q	Q	Y	S	V	P	…	F	K	H	E	D	P	**M**	S	Q	Q	P	S	L
Trubripes	P	P	N	N	I	V	**E**	Q	Q	F	P	A	P	…	F	K	H	E	D	T	**L**	S	P	P	N	N	I
Drerio	Q	Q	S	N	M	G	**D**	Q	Q	Y	P	V	P	…	F	K	H	E	D	S	**I**	T	Q	Q	S	N	M
Dmelanogaster	No homologue													…	No homologue																
Xtropicalis	Q	Q	T	S	L	G	**E**	Q	Q	Y	S	V	P	…	F	K	H	E	D	P	**M**	G	Q	Q	T	S	L

D

aa position at GATA4	226
Hsapiens	C G A M S T **P** L W R R D G
Ptroglodytes	C G A M S T **P** L W R R D G
Mmulatta	C G A M S T **P** L W R R D G
Fcatus	C G A M S T **P** L W R R D G
Mmusculus	C G A M S T **P** L W R R D G
Ggallus	C G A M S T **P** L W R R D G
Trubripes	C G A M S T **P** L W R R D G
Drerio	C G A M S T **P** L W R R D G
Dmelanogaster	C G A I S T **P** L W R R D G
Xtropicalis	C G A M S T **P** L W R R D G

E

aa position at WWOX	62	366
Hsapiens	A G D L P Y **G** W E Q E T D ...	V Y C A A V **P** E L E G L G
Ptroglodytes	Y **G** W E Q E T D ...	Y C A A A **P** E L E G L G
Mmulatta	Y **G** W E Q E T D ...	No homologue
Fcatus	No homologue ...	V Y C A A A **P** E L E G L G
Mmusculus	Y **G** W E Q E T D ...	Y C A V A **P** E L E G L G
Ggallus	Y **G** W E Q E T D ...	Y C A T A **A** E L E G L G
Trubripes	No homologue ...	No homologue
Drerio	Y **G** W E Q E T D ...	No homologue
Dmelanogaster	T G E L P L **G** W E K Y Y D ...	T A **N** E L T G L S
Xtropicalis	No homologue ...	No homologue

F

aa position at ESR1	594	aa position at ESR2	221
Hsapiens	A E G F P A **T** V *	**Hsapiens**	V K C G S R **R** E R C G Y R
Ptroglodytes	A E G F P A **T** V	**Ptroglodytes**	V K C G S R **R** E R C G Y R
Mmulatta	A E G F P A **T** V	**Mmulatta**	V K C G S R **R** E R C G Y R
Fcatus	No homologue	**Fcatus**	V K C G S R **R** E R C G Y R
Mmusculus	A E G F P N **T** I	**Mmusculus**	V K C G S R **R** E R C G Y R
Ggallus	No homologue	**Ggallus**	M K C G S R **R** E R C
Trubripes	No homologue	**Trubripes**	M K C G V R **R** E R C G Y R
Drerio	No homologue	**Drerio**	M K C G L R **R** D R S S Y Q
Dmelanogaster	No homologue	**Dmelanogaster**	L K E G V RLD **R** V R G G R Q
Xtropicalis	E V S L Q S **T** V	**Xtropicalis**	M K C G T R **R** E R C G Y R

G

aa position at DMRT2	536												
Hsapiens	Q	V	G	T	K	L	**S**	V	N	E	P	L	S
Ptroglodytes	Q	V	G	T	K	L	**S**	V	N	E	P		
Mmulatta	Q	V	G	T	K	L	**S**	V	N	E	P		
Fcatus	No homologue												
Mmusculus	Q	V	G	T	K	L	**S**	A	N	E	P		
Ggallus	No homologue												
Trubripes	P	V	G	S	K	M	**V**	P					
Drerio	P	R	C	S	K	V	**T**	A	G	E	S	L	S
Dmelanogaster	No homologue												
Xtropicalis	E	S	G	T	K	L	**T**	V	S	E	P	L	P

H

aa position at ZFPM2	1026												
Hsapiens	S		S	N	G	C	A	**A**	L	K	K	D	S	L
Ptroglodytes			S	N	G	C	A	**A**	L	K	K	D	S	L
Mmulatta	S		S	N	G	C	A	**A**	L	K	K	D	S	L
Fcatus	S		S	N	G	C	A	**P**	P	K	K	D	P	L
Mmusculus			S	N	G	C	A	**V**	P	K	K	D	S	L
Ggallus	P		S	N	G	C	A	**V**	Q	K	K	E	S	L
Trubripes	QQPP		A	N	G	C	P	**H**	P	G	K	E	S	L
Drerio	A		A	N	G	C	P	**H**	P	N	K	D	P	L
Dmelanogaster	No homologue												
Xtropicalis	P		S	N	R	C	S	**L**	M	K	K	E	S	L

I

aa position at MAMLD1	670												
Hsapiens	H	L	M	P	P	R	**T**	G	L	L	Q	N	N
Ptroglodytes	L	L	S	Q	Q	Q	**Q**	Q	Q	Q	Q	Q	Q
Mmulatta	No homologue												
Fcatus	P	L	S	Q	Q	Q	**Q**	P	P	S	Q	Q	P
Mmusculus	S	I	F	K	P								
Ggallus	H	M	L	P	S	N	**P**	A	M	L	Q	S	T
Trubripes	No homologue												
Drerio	No homologue												
Dmelanogaster	No homologue												
Xtropicalis	R	M	I	A	S	P	**Q**	G					

J

aa position at AR	881
Hsapiens	H Q F T F D **L** L I K S H M
Ptroglodytes	No homologue
Mmulatta	H Q F T F D **L** L I K S H M
Fcatus	H Q F T F D **L** L I K S H M
Mmusculus	H Q F T F D **L** L I K
Ggallus	H Q F T F D **L** L I K A H M
Trubripes	H Q F T F D **L** F V Q A Q S
Drerio	H Q F T F D **L** F V Q A R S
Dmelanogaster	No homologue
Xtropicalis	H Q F T F D **L** F V K A Q M

K

aa position at LHCGR	190	253
Hsapiens	N G F E E V **Q** S H A F N G … R L I A T S **S** Y S L K K L
Ptroglodytes	N G F E E V **Q** S H A F N G … R L I A T S **S** Y S L K K L
Mmulatta	N G F E E V **Q** S H A F N G … T L I A T S **S** Y S L K K L
Fcatus	N G F E E I **Q** S H A F N G … T L I A T S **S** Y S L K K L
Mmusculus	E E V **Q** S H A F N G … T L I A T S **S** Y S L K T L
Ggallus	No homologue … No homologue
Trubripes	N G F R T I **R** S H A F N G … T L K A T L **T** F A L K T L
Drerio	N G F Q E I **E** S H A F N G … M L T A R S **A** F A L K K L
Dmelanogaster	N E I S Y V **D** D S A F F G … A L Y I Q N **T** H T L K T I
Xtropicalis	N G F E D I **Q** S Y A F N G … V L V A E Y **S** Y Y L K Q L

L

aa position at STAR	17	121
Hsapiens	G S S Y R H **M** R N M K G L … D V G K V F **R** L E V V V D
Ptroglodytes	G S S Y R H **M** R N M K G L … **R** L E V V V D
Mmulatta	G S S Y R H **M** R N M K G L … **Q** L E V V V D
Fcatus	G S S Y R H **V** R N M K G L … **R** L E V V V D
Mmusculus	G S S Y R H **M** R N M K G L … **R** L E V V V D
Ggallus	A I S Y Q H **L** R N V T G L … **R** L E V V V D
Trubripes	G I S Y R H **M** R N M T G L … **K** L E V V L E
Drerio	G I S Y R H **M** R N M T G L … **K** L E V T L E
Dmelanogaster	No homologue
Xtropicalis	G I S Y R H **L** R N M T G L … **K** L E A V V E

M

aa position at AMH							**143**						
Hsapiens	E	V	T	W	E	P	**T**	P	S	L	R	F	Q
mutated	E	V	T	W	E	P	**I**	P	S	L	R	F	Q
Ptroglodytes	No homologue												
Mmulatta	No homologue												
Fcatus	E	V	T	W	E	P	**T**	P	S	L	K	F	Q
Mmusculus	E	V	I	W	E	P	**E**	L	L	L	K	F	Q
Ggallus	E	GR	S	W	R	V	**V**	P	T	L	SMGLS	W	R
Drerio	No homologue												
Dmelanogaster	No homologue												
Xtropicalis	No homologue												

Supplementary data 15. Genetic tests in the relatives of the index case.

Case	Gene and variants	Zygosity	Family studies
GN0004	SRY; 46,XX (*SRY*+)	-	N/A
GN0007	*SRY*; c.289C>T; p.Gln97Ter	Hemi	Affected sister: hemi, father: mosaicism. Other relatives: wt.
GN0009	*WT1*; c.1447+5G>A	Het	Mother: wt.
GN0011	*MAP3K1*; c.2291T>G; p.Leu764Arg	Het	Parents and brother: wt.
GN0012	*NR5A1*; c.437G>C; p.Gly146Ala	Het	Mother: wt.
GN0018	*AR;* c.2323C>T; p.Arg775Cys	Hemi	N/A
GN0020	*WWOX*; c.1096C>G; p.Pro366Ala	Het	N/A
GN0023	*SRY*; 46,XX (SRY+)	-	N/A
GN0024	*AR*; c.2086G>A; p.Asp69Asn	Hemi	N/A
GN0028	*NR5A1*; c.88T>A; p.Cys30Ser	Het	Parents and brother: wt.
	STAR; c.361C>T; p.Arg121Trp	Het	Father: het. Mother and brother: wt.
GN0034	*LHCGR*; arr [hg19] 2p16.3(48,905,663-48,983,208)x0	Hom	N/A
GN0035	*AR*; c.2522G>A; Arg841His	Hemi	N/A
GN0037	*AR;* c.2178C>G; p.Phe726Leu	Hemi	Mother: het. Father and sister: wt.
GN0038	*HSD17B3*; c.845C>T; p.Pro282Leu	Hom	Mother: het.
GN0041	*AR*; c.(1616+1_1617-1)_(1768+1_1767-1)del; p.(Arg539_Asp305del)	Hemi	Mother: het, affected sister: hemi. Father: wt.
GN0042	*NR5A1*; c.614_615insC; p.Gln206ThrfsTer20	Het	Mother: wt.
	AMH; c.428C>T; p.Thr143Ile	Het	Mother: wt.
GN0046	*SRD5A2*; c.377A>G; p.Gln126Arg	Het	Father: het. Mother: wt.
	SRD5A2; c.(-1+1_1-1)_(281+1_280-1)del; p.(Met1_Arg94del).	Het	Parents: wt.
GN0051	*NR5A1;* c.437G>C; p.Gly146Ala	Het	N/A
GN0054	*SRY*; 46,XX (SRY+)	-	N/A
GN0055	*AR;* c.2710G>A; p.Val904Met	Hemi	Mother: mosaicism.
GN0068	*LHCGR;* c.1713G>T; p.Met571Ile	Het	Affected father: het. Mother and paternal aunts: wt.
GN0070	*NR5A1*; c.437G>C; p.Gly146Ala	Het	N/A
GN0075	*NR5A1*; c.250C>T; p.Arg84Cys	Het	Mother, brother, grandfather, maternal aunt and uncle: het. Father, grandmother and sister: wt.
GN0076	*AR;* c.2566C>T; p.Arg856Cys	Hemi	Mother: het.
GN0078	*NR0B1;* c.913C>T; p.Gln305Ter	Hemi	Cousin: hemi, aunt: het. Uncle: wt.

GN0080	*AR*; c.298insC; p.His100ProfsTer3	Hemi	Mother: het. Father and sister: wt.
GN0088	*LHCGR;* c.1193T>C; p.Met398Thr	Het	Parents: wt.
GN0090	*NR5A1*; c.437G>C; p.Gly146Ala	Hom	N/A
GN0091	*NR0B1*; c.291delC; p.Glu98ArgfsTer166	Hemi	Mother: het. Father, maternal aunt and grandmother: wt.
GN0096	*NR5A1*; c.437G>C; p.Gly146Ala	Hom	N/A
GN0101	*NR0B1*; g.(?_30327014)_(30361290_?)del	Hemi	N/A
GN0109	*NR5A1*; c.910_913delGAGC; p.Glu304CysfsTer26	Het	Mother: het. Sister: wt.
GN0111	*NR5A1*; c.902G>A; p.Cys301Tyr	Het	Mother: mosaicism. Father and Brother: wt.
GN0112	*AR;* c.827delC;p.Pro276HisfsTer20	Hemi	Mother: het. Father: wt.
GN0118	*NR5A1*; c.437G>C; p.Gly146Ala	Het	N/A
GN0119	*NR5A1;* c.437G>C; p.Gly146Ala	Hom	N/A
GN0123	*NR5A1*; c.71A>T; p.His24Leu	Het	N/A
GN0125	*AR*; c.865G>T; p.Glu289Ter	Hemi	N/A
GN0132	*WT1*; c.1447+4C>T	Het	N/A
GN0133	*SRY*; 46,XX (SRY+)	-	N/A
GN0139	*AR;* c.2642T>G; p.Leu881Arg	Hemi	Mother: het.
GN0141	*SRY*; c.391C>T; p.Pro131Ser	Hemi	N/A
GN0142	*DMRT2*; c.1607C>T; p.Ser536Leu	Het	Mother: het. Other relatives: wt
GN0146	*AR*; c.2473C>A; p.Gln825Lys	Hemi	N/A
GN0147	*LHCGR*; c.757T>C; p.Ser253Pro	Hom	N/A
	NR5A1; c.437G>C; p.Gly146Ala	Het	
GN0150	*WT1*; c.223G>A;p.Glu75Lys	Het	Mother: het.
GN0153	*NR0B1*; c.528C>G; p.Tyr176Ter	Hemi	Mother and grandmother: het. Father and aunt: wt.
GN0154	*MAMLD1*; c.2009C>T; p.Thr670Ile	Het	N/A
GN0155	*ZFPM2*; c.3077C>T, p.Ala1026Val	Het	N/A
GN0156	*WT1*; c.545T>A; p.Met182Lys	Het	N/A
	NR5A1; c.437G>C; p.Gly146Ala	Hom	
GN0157	*LHCGR*; c.568C>A; p.Gln190Lys	Het	Father: het. Mother: wt
	NR5A1; c.437G>C; p.Gly146Ala	Het	Father: homo. Mother: wt.
GN0158	*NR5A1*; c.437G>C; p.Gly146Ala	Hom	N/A
GN0159	*SRY*; 46,XX.ish der(X)t(X;Y)(p22.3;p11.3)(SRY)	-	Parents: wt.
GN0163	*NR5A1*; c.437G>C; p.Gly146Ala	Het	N/A

GN0164	*AR*; c.2270A>G; p.Asn757Ser	Hemi	N/A
GN0171	*GATA4*; c.677C>T; p.Pro226Leu	Het	Mother: het.
	LHCGR; c.1660C>T; p.Arg554Ter	Het	Mother: wt.
GN0177	*AR*; c.1301C>T; p.Ser434Phe	Hemi	N/A
GN0182	*NR5A1*; c.437G>C; p.Gly146Ala	Het	N/A
GN0185	*STAR*; c.50T>G; p.Met17Arg	Het	Father: het. Mother: wt.
GN0186	*SRD5A2*; c.271T>G; Tyr91Asp	Hom	Parents: het.
GN0187	*SRY*; 46,XX.ish der(X)t(X;Y)(p22.3;p11.3)(SRY)	-	N/A
GN0189	*AR*; c.2567G>A; p.Arg856His	Hemi	Mother: het. Father: wt.
GN0194	*AR;* c.2323C>T; p.Arg775Cys	Hemi	N/A
	NR5A1; c.437G>C; p.Gly146Ala	Het	
GN0198	*HSD17B4;* c.524delC; p.Ala175GlufsTer26	Het	Mother: het. Father: wt.
	ESR1; c.1781C>T;p.Thr594Met	Het	Mother: het. Father: wt.
GN0199	*NR5A1*; c.437G>C; p.Gly146Ala	Het	Mother: het. Father: wt.
GN0203	*WWOX*; c.184G>A; p.Gly62Arg	Het	N/A
GN0207	*ESR2*; c.661A>G; p.Arg221Gly	Het	N/A
OT0567	*WT1*; c.(1099-?_1551+?)del; p.(Asp367?_Leu517?)del	Het	Parents: wt.
POL0274	*NR0B1*; c.871T>A; p.Trp291Arg	Hemi	Mother: het. Father: wt.
POL0285	*NR0B1*; g.(?_29978097)_(30361290_?)del	Hemi	N/A
POL0301	*NR0B1*; c.712_713delAC; p.Thr238LeufsTer60	Hemi	Mother: het. Father: wt.
RE0045	*CYP17A1*; c.1246C>T; p.Arg416Cys	Hom	Affected sister: homo. Mother, sister and niece: het.

Hemi: hemizygosis; Het: heterozygosis; Homo: homozygosis; N/A: not available; wt: wild type.

Supplementary data 16. Article originated from this work

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 04 April 2018
doi: 10.3389/fendo.2018.00142

GATA4 Variants in Individuals With a 46,XY Disorder of Sex Development (DSD) May or May Not Be Associated With Cardiac Defects Depending on Second Hits in Other DSD Genes

*Idoia Martinez de LaPiscina[1,2,3†], Carmen de Mingo[4†], Stefan Riedl[5], Amaia Rodriguez[6], Amit V. Pandey[2,3], Mónica Fernández-Cancio[7], Nuria Camats[7], Andrew Sinclair[8], Luis Castaño[1,6], Laura Audí[7] and Christa E. Flück[2,3]**

[1] Endocrinology and Diabetes Research Group, BioCruces Health Research Institute, Cruces University Hospital, CIBERDEM, CIBERER, UPV-EHU, Barakaldo, Spain, [2] Pediatric Endocrinology, Diabetology and Metabolism, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland, [3] Pediatric Endocrinology, Diabetology and Metabolism, Department of BioMedical Research, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland, [4] Pediatric Endocrinology, La Fe Pediatric University Hospital, Valencia, Spain, [5] Division of Pediatric Pulmology, Allergology, and Endocrinology, St. Anna Children's Hospital, Department of Pediatrics, Medical University of Vienna, Vienna, Austria, [6] Pediatric Endocrinology Section, Cruces University Hospital, BioCruces Health Research Institute, CIBERDEM, CIBERER, UPV/EHU, Barakaldo, Spain, [7] Growth and Development Research, Pediatric Endocrinology Unit, Vall d'Hebron Research Institute (VHIR), CIBERER, Instituto de Salud Carlos III, Barcelona, Spain, [8] Department of Paediatrics, Murdoch Children's Research Institute, University of Melbourne, The Royal Children's Hospital, Melbourne, VIC, Australia*

Disorders of sex development (DSD) consist of a wide range of conditions involving numerous genes. Nevertheless, about half of 46,XY individuals remain genetically unsolved. *GATA4* gene variants, mainly related to congenital heart defects (CHD), have also been recently associated with 46,XY DSD. In this study, we characterized three individuals presenting with 46,XY DSD with or without CHD and *GATA4* variants in order to understand the phenotypical variability. We studied one patient presenting CHD and 46,XY gonadal dysgenesis, and two patients with a history of genetically unsolved 46,XY DSD, also known as male primary hypogonadism. Mutation analysis was carried out by candidate gene approach or targeted gene panel sequencing. Functional activity of GATA4 variants was tested *in vitro* on the CYP17 promoter involved in sex development using JEG3 cells. We found two novel and one previously described *GATA4* variants located in the N-terminal zinc finger domain of the protein. Cys238Arg variant lost transcriptional activity on the CYP17 promoter reporter, while Trp228Cys and Pro226Leu behaved similar to wild type. These results were in line with bioinformatics simulation studies. Additional DSD variations, in the *LRP4* and *LHCGR* genes, respectively, were identified in the two 46,XY individuals without CHD. Overall, our study shows that human *GATA4* mutations identified in patients with 46,XY DSD may or may not be associated with CHD. Possible explanations for phenotypical variability may comprise incomplete penetrance, variable sensitivity of partner genes, and oligogenic mechanisms.

Keywords: disorder of sexual development, DSD, GATA4, congenital heart defects, oligogenic, 46,XY DSD

OPEN ACCESS

Edited by:
Andrea Enzo Scaramuzza, Istituti Ospitalieri di Cremona, Italy

Reviewed by:
Alexandra Francois Roy Stewart, University of Ottawa, Canada
Marek Niedziela, Poznan University of Medical Sciences, Poland
Rodolfo A. Rey, Hospital General de Niños Ricardo Gutiérrez, Argentina

***Correspondence:**
Christa E. Flück
christa.flueck@dbmr.unibe.ch

†Co-first authors for equal contribution.

Specialty section:
This article was submitted to Pediatric Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *21 December 2017*
Accepted: *16 March 2018*
Published: *04 April 2018*

Citation:
Martinez de LaPiscina I, de Mingo C, Riedl S, Rodriguez A, Pandey AV, Fernández-Cancio M, Camats N, Sinclair A, Castaño L, Audí L and Flück CE (2018) GATA4 Variants in Individuals With a 46,XY Disorder of Sex Development (DSD) May or May Not Be Associated With Cardiac Defects Depending on Second Hits in Other DSD Genes. Front. Endocrinol. 9:142. doi: 10.3389/fendo.2018.00142

INTRODUCTION

Disorders of sex development (DSD) are defined as congenital conditions, in which development of chromosomal, gonadal, or anatomical sex is atypical (1). 46,XY DSD includes disorders in male gonad determination and differentiation, androgen biosynthesis or action, and anti-Müllerian hormone (AMH) synthesis or action (1). In the last two decades, numerous genes have been found to cause 46,XY DSD (2). In about 40% of 46,XY DSD individuals, the underlying genetic cause still remains unsolved. GATA4 has been more recently implicated in 46,XY DSD.

The GATA family of transcription factors consists of six members, three expressed in hematopoietic stem cells (GATA1–3) and three in tissues derived from mesoderm and endoderm, including heart, gonad, lung, liver, and gut (GATA4–6). GATA factors regulate tissue-specific gene expression either alone or in cooperation with other factors (3, 4). GATA proteins have two zinc fingers (ZNI and ZNII), which are highly conserved while the sequences of the amino-terminal and carboxyl-terminal domains exhibit lower similarities (5, 6). While the C-terminal zinc finger region is required for the DNA recognition and binding, and the N-terminal zinc finger region contributes to the stability, both zinc fingers are necessary for protein–protein interactions with other transcription factors (4, 7).

The human *GATA4* gene on chromosome 8p23.1 encodes an essential transcription factor for the developing gonad (8) and heart (9, 10). Gata4-null mice die due to severe abnormalities in heart tube formation and ventral morphogenesis (9, 10). Gata4 expression and function seems required for normal testicular and genital development (11–14). Human GATA4 interacts with several proteins, including NR5A1, WT1, and FOG2 to regulate the expression of sex determining genes *SRY*, *SOX9*, and *AMH*, as well as genes involved in sex differentiation such as *StAR*, *CYP17A1*, *CYP19A1*, *INHA*, and *HSD3B2* (15–18).

In humans, *GATA4* haploinsufficiency has been described in patients with different forms of congenital heart defects (CHD) since 1999 (OMIM_600576). So far more than 120 *GATA4* gene variants have been associated with cardiac defects (HGMD® professional 2017, accessed October 2017), while only four studies reported mutations related to a 46,XY DSD phenotype (19–22).

Overall, *GATA4* gene variations seem more likely to result in CHD than DSD, but the reason for this is unclear. And the fact that some individuals with GATA4 haploinsufficiency have a 46,XY DSD phenotype only remains similarly unexplained.

We, therefore, characterized three additional individuals presenting with 46,XY DSD and two novel and one previously described *GATA4* variants. They presented with or without CHD, but additional variants in two other genes were found that likely contributed to the DSD phenotype.

MATERIALS AND METHODS

Ethical Approval

Written informed consent was obtained from all subjects and their family members at the respective hospitals involved. The study was approved by the local ethical committees, namely the ethical boards of the La Fe University Hospital, Valencia, Spain, the St. Anna Children's Hospital, Vienna, Austria, as well as the ethics committee for clinical research of Euskadi (CEIC-E), Spain.

Genetic Analysis

Genomic DNA was isolated from peripheral blood leukocytes and analyzed for genetic alterations causing DSD using different approaches.

Case 1 was first studied by a candidate gene approach. *GATA4* gene segments corresponding to the 5′ UTR region, the 6 coding exons and their flanking intronic sequences were amplified by PCR using specific primers [Table S1 in Supplementary Material; (21)]. The PCR products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit on an automated ABI PRISM 3100 Genetic Analyzer (Applied biosystems, Foster City, CA, USA). Obtained sequences were analyzed and compared with the wild-type (wt) published reference sequence (RefSeq NM_002052) using SeqScape Software v. 2.5 (Applied Biosystems). The DNA sample was in addition studied on a sequencing panel (TruSight One Sequencing Panel, Illumina, San Diego, CA, USA) containing 4,813 disease-associated genes, including 94 well-known candidate genes for DSD 46,XY and 46,XX.

Case 2 was examined on the target gene panel for DSD by Haloplex technology (Agilent, Santa Clara, CA, USA). This system allowed to simultaneously sequence 64 diagnostic genes for DSD and 967 candidate genes (19).

Case 3 was analyzed by a customized Ion Ampliseq panel (ThermoFisher Scientific, Waltham, MA, USA) comprising coding and flanking regions of 48 DSD genes. Analysis was performed according to the manufacturer's instructions.

Of note, GATA4 variants identified by panel analysis were verified by Sanger sequencing. Parents and affected family members were tested to establish the mode of inheritance.

Identified *GATA4* sequence variants were then evaluated for their possible functional significance *in silico* using prediction software programs, including SIFT,[1] Provean,[2] PolyPhen-2,[3] MutationTaster,[4] MutPred,[5] and SNPs&GO.[6]

In all three cases, the involved clinician informed the patient about the genetic results, including information on pathogenic, probably pathogenic, and uncertain findings with regard to the DSD/genital phenotype. Incidental findings were not made.

In Vitro Functional Studies

Human placental JEG3 cells (300222[7]), human adrenal NCI-H295R cells (ATCC CRL-2128[8]), and non-steroidogenic human embryonic kidney HEK293 cells (ATCC CRL-1573) were cultured as described (18, 23) and used for functional assays.

[1] http://sift.jcvi.org/ (Accessed: March, 2017).
[2] http://provean.jcvi.org/index.php (Accessed: March, 2017).
[3] http://genetics.bwh.harvard.edu/pph2/ (Accessed: March, 2017).
[4] http://www.mutationtaster.org/ (Accessed: March, 2017).
[5] http://mutpred.mutdb.org/ (Accessed: March, 2017).
[6] https://snps-and-go.biocomp.unibo.it/snps-and-go/ (Accessed: March, 2017).
[7] http://www.dsmz.de/ (Accessed: April, 2017).
[8] http://www.lgcstandards-atcc.org (Accessed: April, 2017).

Promoter luciferase reporter vector −227CYP17A1_Δluc was available from previous work (18). Promoter reporter vectors pGL3AMH and pGL3SRY were cloned by PCR amplifying the fragments of −1,067/+32 *SRY* and −951/+115 *AMH* from human genomic wt DNA and inserting the fragments into pGL3 basic (Promega, Dübendorf, Switzerland). Vectors were verified by direct sequencing.

Human wt GATA4 cDNA was inserted into the mammalian pCMV expression vector. GATA4 mutations c.684A, c.712C, and c.677T were generated by site-directed mutagenesis using specific primers and the QuickChange protocol by Stratagene (Agilent technologies Inc., Santa Clara, CA, USA). We also constructed the previously described c.661A GATA4 variant for comparison (21). Mutant plasmids were confirmed by direct sequencing.

For functional studies, cells were cultured on 12-well plates and transiently transfected with wt or mutant GATA4 vectors and a CYP17 promoter reporter using a calcium phosphate transfection protocol (Invitrogen, ThermoFisher Scientific). 48 h after transfection, cells were washed and lysed, before luciferase activity was measured with the Dual-Luciferase® Reporter (DLR™) Assay system (Promega AG, Wallisellen, Switzerland) on a Veritas microplate Luminometer reader (Turner Bio systems Luminometer and Software by Promega). Specific *Firefly* luciferase readings were standardized against *Renilla* control readings and results expressed as relative luciferase units. Results are shown as the mean ± SEM of four independent experiments, all performed in duplicate. Data were statistically analyzed using the Student's *t* test. A p-value ≤ 0.05 was considered significant.

Protein Structure Analysis

Since the structure of GATA4 protein is not known, a hybrid homology model of the protein was made using YASARA (24) and AMBER. Three rounds of PSI-BLAST were performed with the GATA4 amino acid sequence (NP_002043) to build a position-specific scoring matrix, and then a search of protein structure database was conducted to find the templates for model generation. Structures of GATA3 (PDB # 4HC9 and 4HC7) and GATA1 (PDB # 3VD6 and 3VEK) were found to be best templates for making the GATA4 protein model. Using each template separately several models were made. Three loops from amino acid sequences NLDMFDDFSE, KNLNKSKTPA, and LIKPQRRLSASRRVGL needed to be modeled separately and joined in the structure. Side chains of modeled amino acids were optimized by screening of rotamer libraries and molecular dynamics (MD) simulations. Best quality parts from different models were then combined to generate a hybrid model, which had increased coverage and accuracy compared to individual models. *In silico* mutagenesis was performed to create the GATA4 mutations described in this report (25), which were then further refined by brief MD simulations of 500 ps. All models were then subjected 500 steps of steepest descent and simulated annealing minimizations using AMBER14 force filed and TIP3P water model (26, 27), and then subjected to 25 ns of explicit solvent MD simulations at 310 K. Structural analysis was performed after the MD simulation systems were stable. All illustrations were prepared with PyMOL[9] and rendered as ray-traced images using PovRay.[10]

RESULTS

Cases Reports of Three 46,XY DSD Patients Harboring GATA4 Gene Variations

Case 1

This "female" newborn was referred after birth because of ambiguous genitalia (**Table 1**). She presented with a small vulva, slight clitoral hypertrophy, labia were fused in posterior raphe and gonads were palpable in inguinal canal. Karyotype was 46,XY. Ultrasound and pelvic MRI showed a rudimentary uterus. Biochemical and hormonal studies at presentation and during follow-up revealed more or less normal values (compared to male age-controls), only undetectable and low AMH was remarkable (**Table 2**). She was also seen by a pediatric cardiologist for a heart murmur and found to have a complex CHD, which was not detected in prenatal screening (**Table 1**). At 40 days of life, surgical closure of the ventricular septal defect was performed. It was then noted that she also had congenital compression of the left bronchus leading to an asymmetry in the caliber of the pulmonary branches with a hypoplastic left branch. During follow-up over 3 years, she showed some signs of developmental delay. She is reared as a girl.

Case 2

Patient 2 presented micropenis with hypospadias and cryptorchidism at birth (**Table 1**). His karyotype was 46,XY. At 3 months of age, his gonadotropins were elevated at a testosterone of 0.92 ng/ml (**Table 2**). At 7 months, spontaneous descensus of the left testis occurred, whereas the right testis was found inguinal. Stimulation test with hCG resulted in a normal rise of testosterone. He was treated for micropenis with testosterone (3 mg × 25 mg monthly), which increased phallic size from 2.0 to 3.5 cm. At 17 months, right-sided orchidopexy was performed including biopsy. Histology revealed a sertoli-cell only phenotype; spermatogonia were not visible. Tubules appeared immature exhibiting a compressed, lumen-less structure. At follow-up at 10 years of age, testicular sonography revealed multiple calcifications with a testis size of 0.5 ml on the right and 0.8 ml on the left. The boy was still prepubertal with a small phallus of 3 cm and a hypoplastic scrotum. Biochemically, luteinizing hormone (LH), follicle stimulating hormone (FSH), and testosterone were already on the rise. AMH and adrenal steroids were in the normal, age- and sex-appropriate range. Apart from the development of his external genitalia, the boy has been healthy and psychomotor development is normal. He has no CHD.

Family history is remarkable for the following: the boy's mother is healthy with no fertility problems, but was diagnosed with a mitral valve prolapse in adulthood requiring no medical intervention. According to cardiologists, mitral valve prolapse is often seen with aging and does not qualify as CHD. The maternal uncle

[9]www.pymol.org (Accessed: October, 2017).
[10]www.povray.org (Accessed: October, 2017).

TABLE 1 | Clinical and genetic characterization of three 46,XY DSD patients harboring variants in the *GATA4* gene.

Case	Gender assigned	Phenotype					Genotype		
		Genital anatomy at initial presentation	Müllerian ducts	Gonadectomy	CHD	Syndromic features	Karyotype	*GATA4* gene variant	Other DSD genes
1	F	Clitoral hypertrophy, fused labia with posterior raphe, gonads palpable in inguinal canal	Rudimentary uterus	No. But, planned with cardiac surgery	Yes	Developmental delay: sitting 12 months, crawling 18 months, walking 24 months; at 3 years no language skills, signs of autism	46,XY	c.712T>C; p.Cys238Arg	None
2	M	Micropenis, hypospadias, bilateral cryptorchidism	No	No	No	No	46,XY	c.684G>C; p.Trp228Cys	*LRP4* (c.5680C>G; p.Ser1887Cys)
3	M	Micropenis, bilateral cryptorchidism (inguinal)	No	No	No	Severe obesity	46,XY	c.677C>T; p.Pro226Leu	*LHCGR* (c.1880C>T; p.Arg554Stop)

All sequence information about GATA4 gene is based on NM_002052. LRP4 and LHCGR information is based on NM_002334 and NM_000233, respectively.
CHD, congenital heart defects; LHCGR, luteinizing hormone/choriogonadotropin receptor; LRP4, LDL receptor-related protein 4.

TABLE 2 | Biochemical characterization of three 46,XY DSD patients harboring variants in the *GATA4* gene.

Case	Adrenal function							Gonadal function					
	Age at evaluation	ACTH (pg/ml)	Cortisol (μg/dl)	11-Deoxycortisol (ng/ml)	17OHP (ng/ml)	DHEA-S (ng/ml)	Androstenedione (ng/ml)	T (ng/ml)		FSH (U/l)	LH (U/l)	Estradiol (pg/ml)	AMH (ng/ml)
		Basal	Basal	Basal	Basal	Basal	Basal	Basal	After hCG	Basal	Basal	Basal	Basal
1	2 days	**243.0**	11.3	11.4	4.4	**1,660**		1.30		0.2	0.1	59	**<0.14**
	7 days	36.7	**1.9**										
	14 days									**19.3**	3.4	42	
	3.5 years					123	<0.3	<0.02		**14.1**	<0.1	<5.0	**4.4**
2	3 months							0.92		**36**	**12.3**		
	7 months							0.11	1.87				
	10 years	16	8.6		0.31	2.41	0.55	0.21		**14.9**	**0.9**	13	14.4
3	11 years							**1.30**		**63.3**	**14**		
	14 years							**1.30**		**36.7**	**18.5**		
	15 years							5.9[a]		1.5[a]	**<0.5**[a]		
	21 years							11.8[a]		9.5[a]	4.0[a]		

Laboratory test values outside the normal range of age and chromosomal sex are given in bold.
ACTH, adrenocorticotropic hormone; 17OHP, 17-hydroxyprogesterone; DHEA-S, dehydroepiandrosterone sulfate; T, testosterone; FSH, follicle stimulating hormone; LH, luteinizing hormone; AMH, anti-Müllerian hormone.
[a]*Under testosterone replacement treatment.*

had a similar history as the patient with micropenis and bilateral cryptorchidism at birth. Orchidopexy and testosterone treatment for micropenis were performed in early childhood. Pubertal onset occurred spontaneously and no testosterone treatment was required. At 18 years of age, he had markedly elevated gonadotropins (LH 25.8 IU/l; FSH 53.8 IU/l) in the presence of a low-normal testosterone. Testicular volume was 3 ml. Azoospermia was ascertained at 23 years and subsequent testicular biopsy revealed sertoli-cell only syndrome. During the following years, testosterone levels decreased continuously and AMH became undetectable. He is otherwise healthy, without signs of a CHD.

Case 3

This patient was noted at birth to have micropenis and bilateral cryptorchidism, but no action was taken (**Table 1**). At 11 years of age he still presented micropenis (3.5 cm), testes were non-palpable and severe obesity (BMI 36.6, +6.4 SD) was observed. His gonadotropins were elevated with low testosterone. Karyotype was 46,XY. MRI showed both testes in inguinal canal. At 14.3 years, replacement testosterone treatment was started with gradual increasing doses. After treatment, LH suppression and normalization of FSH was observed. At 18 years, azoospermia was found. At recent follow-up at 21 years of age, he presented with severe obesity (BMI 44.9, +6.8 SD) and short stature (156.1 cm). Penis (8 cm) was buried in subcutaneous fat and left testis was in scrotum (0.5 ml). Biochemically, he presented normal gonadotropins with a testosterone of 11.8 ng/ml under treatment. Ultrasound showed an ovoid structure (17 mm × 16 mm) in right inguinal canal corresponding to atrophic testis. Currently, laparoscopy is planned to investigate the right testis and biopsy both to assess the malignancy risk. Echocardiography revealed no CHD.

Genetic Characterization of Three Novel *GATA4* Gene Variants

We found novel *GATA4* gene variants in three patients presenting with 46,XY DSD at birth.

In case 1, direct analysis by Sanger sequencing revealed a T to C change in heterozygosis at position 712 (c.712T>C; p.Cys238Arg) in exon 3 of GATA4. Both parents were studied and did not carry the variant, thus this GATA4 variant is a *de novo* change in the patient.

In case 2, DSD targeted gene panel sequencing identified a heterozygous GATA4 c.684G>C nucleotide change, which is predicted to result in a p.Trp228Cys alteration. This variant has been previously listed in a publication by Eggers et al. reporting on DSD panel diagnostics (19). Remarkably, the patients' mother presenting with a mitral valve prolapse, and his maternal uncle having a similar history of micropenis and cryptorchidism, were both carriers of the same heterozygous GATA4 variant.

In case 3, a heterozygous C to T transition located at c.677 of the coding GATA4 DNA sequence predicted to replace a proline with leucine at position p.226 (c.677C>T; p.Pro226Leu) was found by DSD targeted panel sequencing. The same heterozygous variant was also found in the healthy mother.

Looking at these GATA4 variants, it is remarkable that all are located in the N-terminal zinc finger domain of the GATA4 protein (**Figure 1**), where the previously described variant p.Gly221Arg (found in a patient with 46,XY DSD) is also located (21). Prediction software programs and comparison of GATA4 proteins across species revealed that Cys238, Trp228, and Pro226, and the surrounding region are highly conserved (**Figure 2**). Therefore, the identified GATA4 gene variants were predicted as potentially pathogenic.

Additional (DSD) Gene Variations Identified by Next Generation Sequencing (NGS)

In case 1, additional targeted exome sequencing of 4,813 disease-causing genes, containing 94 genes related to DSD, did not reveal any further sequence variant. In cases 2 and 3, without CHD, DSD targeted gene sequencing identified two additional gene variations. Case 2 revealed a novel, heterozygous C to G change at position 5660 (c.5560C>G; p.Ser1887Cys) of the LDL receptor-related protein 4 (*LRP4*) gene (OMIM 604270). This gene variant was also found in the mother and the maternal uncle. In case 3, a heterozygous change of cytosine to thymine was found at location 1660 of the *LHCGR* gene (OMIM 152790). This mutation changes codon 554 from arginine to a stop codon (c.1160C>T; p.Arg554Stop) and has been previously reported (28). The mother did not carry this luteinizing hormone/chorionadotropin receptor (LHCGR) mutation.

In Vitro Functional Studies of Novel GATA4 Variants

To test the transcriptional activity of identified GATA4 variants, we constructed mammalian expression vectors of wt and mutant GATA4 and tested them on three different promoters that have been described being regulated by GATA4, namely the AMH, SRY, and CYP17 promoters. For these studies, we used different

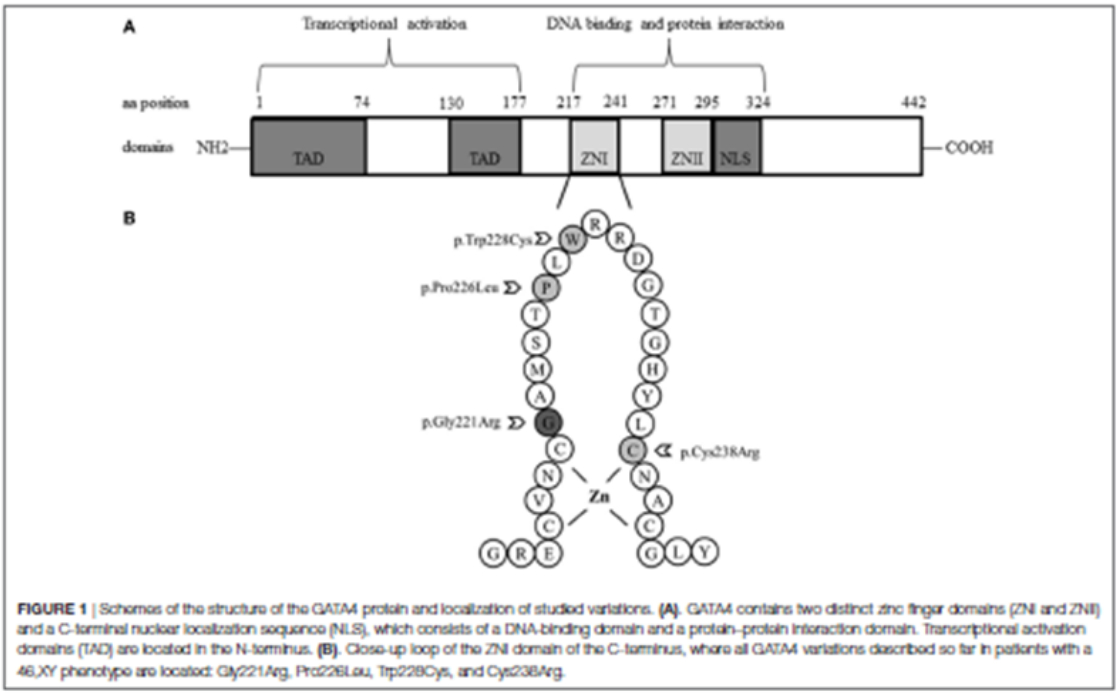

FIGURE 1 | Schemas of the structure of the GATA4 protein and localization of studied variations. **(A)**. GATA4 contains two distinct zinc finger domains (ZNI and ZNII) and a C-terminal nuclear localization sequence (NLS), which consists of a DNA-binding domain and a protein–protein interaction domain. Transcriptional activation domains (TAD) are located in the N-terminus. **(B)**. Close-up loop of the ZNI domain of the C-terminus, where all GATA4 variations described so far in patients with a 46,XY phenotype are located: Gly221Arg, Pro226Leu, Trp228Cys, and Cys238Arg.

FIGURE 2 | Multiple alignment of parts of the GATA4 protein sequences across species. Localization of the newly identified human aa variants are given in red and seem highly conserved across species. The localization of the so far only GATA4 mutation reported to cause a 46,XY DSD phenotype (p.Gly221Arg) is given in bold and its localization is also highly conserved.

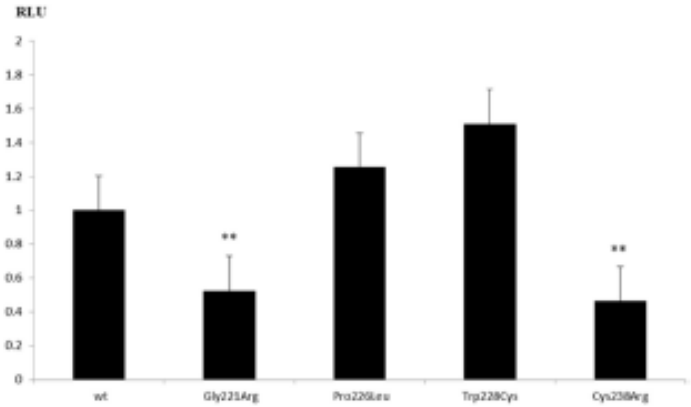

FIGURE 3 | Transcriptional activity of GATA4 variants on the *CYP17A1* promoter. Human placental JEG3 cells were transfected with a CYP17-promoter luciferase reporter construct, and the activity of wild-type (wt) and mutant GATA4 to trans-activate the promoter was tested using the Promega Dual Luciferase readout system. Results are shown as the mean ± SEM of four independent experiments, all performed in duplicate. **p-Value ≤ 0.01.

cell systems (HEK293, NCI-H295R, and JEG3), but found that only JEG3 cells transfected with the CYP17 promoter revealed consistent results for comparing wt to mutant GATA4. We found that GATA4 variant Cys238Arg lost transcriptional activity (**Figure 3**) similar to the previously described Gly221Arg mutant (21). By contrast, GATA4 variants Trp228Cys and Pro226Leu activated the CYP17 promoter similar to wt.

In Silico GATA4 Protein Structure-Function Analysis

Using the structures of GATA1 and GATA3 available in the protein structure database, we created a protein structure model of GATA4. GATA4 was modeled with dual zinc atoms, which were coordinated with cysteines (**Figure 4**). One Zn atom is bound to cysteines 217, 220, 238, and 241, while the second Zn atom is bound to cysteines 272, 275, 293, and 296. Mutation of any of these cysteine molecules as well as neighboring residues is expected to result in loss of Zn binding. Among the mutations studied in this report, we found the mutations Gly221Arg and Cys238Arg to cause loss of Zn binding leading to an unstable protein with predicted loss of function (**Figure 4**). The mutations Pro226Leu and Trp228Cys are involved in interaction with DNA and their effects were predicted to be variable. Since the core structure of the GATA4 protein is not expected to be affected by these two mutations, it is the interactions with different DNA fragments that is predicted to be changed. As we have observed in case of SF1 mutations, the mutations causing a change in the interaction with DNA can have variable effects (23). While interaction and activation of one partner sequence can be affected and may cause full or partial loss of activation, the newly created nucleotide recognition element due to mutations could interact and activate another promoter sequence, resulting in altered gene expression. In case of Gly221Arg, loss of interaction with FOG2 protein has been reported. Structural analysis of this mutation shows a likelihood of unstable protein due to loss of Zn binding and therefore, an alteration in the interaction with GATA4 binding partners is expected.

DISCUSSION

We characterized three human GATA4 sequence variations found in three individuals with a 46,XY DSD phenotype with and without CHD. These variants were all missense changes and were present in heterozygous form. NGS revealed additional likely

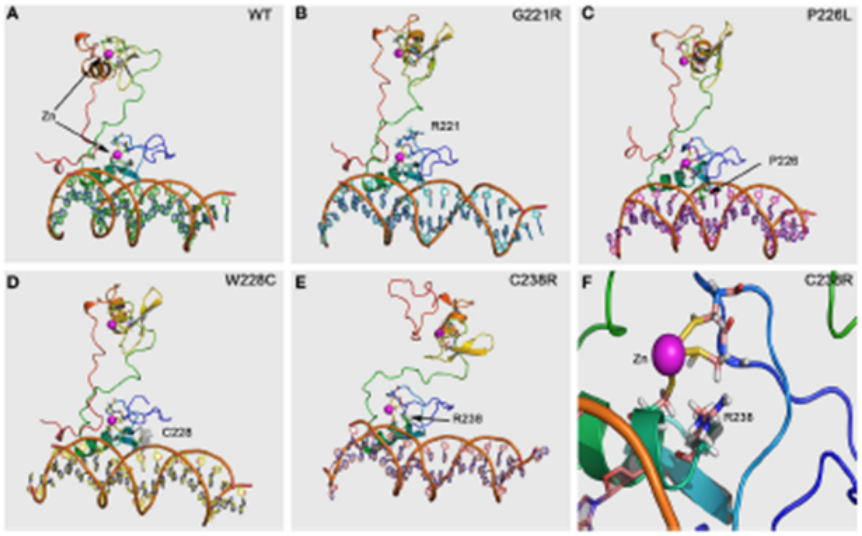

FIGURE 4 | GATA4 protein model and prediction of effects by mutations. The GATA4 protein model was made using the structures of GATA1 and GATA3 available in the protein structure database. After performing a structure-based search using Phi-BLAST, two different structures of GATA1 and GATA3 were selected and used for model building as described in Section "Materials and Methods." **(A)** Structural of WT GATA4 showing the Zn coordination and binding to DNA. **(B)** The Gly221Arg mutation is predicted to cause disruption of Zn binding and protein instability, resulting in a loss of function mutation. **(C)** The mutation Pro226Leu is in the nucleotide interaction interface and was predicted to alter the recognition of promotor sequences for binding to GATA4. This can lead to variable effects with some promotors showing an effect while other could be affected. **(D)** The mutation Trp228Cys was also predicted to cause variable effects due to its proximity to DNA-binding interface of GATA4. Inclusion of extra cysteine was not predicted to cause the disruption of Zn binding as Cys228 in the mutant was far away from the Zn binding site. **(E)** The mutation Cys238Arg was predicted to cause loss of Zn binding, leading to an instable and non-functional GATA4 protein. **(F)** A close-up of the Cys238Arg mutation showing the loss of Zn coordination to cysteine due to introduction of the arginine, which could not bind the Zn molecule to stabilize the protein. A severe loss of function effect was predicted for the Cys238Arg mutation.

disease causing variants in two other genes (so-called second hits), namely *LRP4* or *LHCGR*, in two patients with a 46,XY DSD phenotype without CHD.

In both mice and men, GATA4 is expressed in the somatic cell population of the developing heart and gonads and plays an essential role in their development (8–10). Human GATA4 mutations associated with cardiac malformations are described in numerous patients (Table S2 in Supplementary Material) and are located throughout the gene in the coding and non-coding sequence including 3′- and 5′-UTR (29–33). They comprise single nucleotide variants as well as deletions and duplications, which lead to missense/nonsense mutations or splicing errors and thus altered proteins or protein expression. Studies of the pathogenesis of CHD have reported that mutations in the GATA4 gene decrease the ability to *trans*activate target genes or fail to interact with proteins involved in heart development (34, 35). Moreover, it has been recently demonstrated that common variants in a region of GATA4 3′UTR contribute to the risk of CHD by changing posttranscriptional gene regulation at the level of miRNA (36). Altogether, these studies show that *GATA4* gene mutations contribute to the susceptibility of CHD. A list of these numerous GATA4 mutations discovered in patients with CHD so far is provided in Table S2 in Supplementary Material. For a more comprehensive review of the role of GATA4 in CHD we refer to Ref. (31).

By contrast, the few GATA4 missense mutations found in 46,XY DSD individuals with or without CHD are all located in the N-terminal zinc finger domain, which is responsible for DNA binding and interaction with cofactors (4, 7).

Functional characterization of GATA4 variants with respect to the 46,XY DSD phenotype has only been performed for the p.Gly221Arg mutation so far (21). *In vitro* studies revealed that p.Gly221Arg lacked DNA binding, had impaired transactivation activity on the AMH promoter, and failed to bind cofactor FOG2. Functional testing of three GATA4 variants identified in 46,XY DSD individuals of our study showed similarly disruptive effect for the missense mutation p.Cys238Arg, but no effect on transactivation activity on the CYP17 promoter for GATA4 variants p.Pro226Leu and pTrp228Cys. While all these variants are conserved across species (**Figure 2**) and located in the N-terminal zinc finger domain of GATA4 (**Figure 1**), only Gly221 and Cys238 are close to Zn binding sites. The Gly221 is not directly involved in Zn binding but is situated next to Cys220 which binds the Zn atom, and therefore, the mutation Gly221Arg will disrupt the Zn binding, leading to a non-functional GATA4. The Cys238 binds Zn and its mutation to arginine leads to loss of Zn binding (**Figure 4**).

GATA4 regulates the expression of multiple genes coding for hormones or components of the steroidogenic pathway during testis development and function. In Gata4^{ki} mice with p.Val217Gly mutation interaction of Gata4 with cofactor Fog is abrogated, and consequently animals display anomalies of testis development (12–14). Moreover, GATA4 functionally interacts with NR5A1 in Sertoli cell cultures to positively regulate the expression of AMH, and therefore, it has been reported that mutations in NR5A1 may cause 46,XY DSD due to lack of interaction with GATA (15). No gonadal involvement is mostly detected in families with *GATA4* mutations and isolated CHD, possibly because some of the variants retain some DNA-binding activity and exhibit different degrees of transcriptional activation on gonadal promoters and thus, remain able to synergize with NR5A1. In the present study, the p.Cys238Arg mutation was found in a patient with a complex CHD, genital ambiguity, and persistent Müllerian ducts, which led to female gender assignment. We propose that cysteine to arginine change in position 238 of GATA4 lacks activity to bind DNA reducing the transactivation of AMH critically.

By contrast, variants p.Pro226Leu and pTrp228Cys found in cases 2 and 3 did not affect CYP17 promoter activity. These individuals had a less severe 46,XY DSD phenotype, were raised as males, and had no evidence of heart anomalies. White et al. (22) described a 35 kb deletion immediately 3′ UTR of the *GATA4* gene in a patient presenting with complete gonadal dysgenesis (GD). Additionally, Harrison et al. (20) screened patients with 46,XY DSD/GD with array CGH and found an infant presenting a deletion of 0.22 Mb upstream *GATA4* in chromosome 8p23.1. Same deletion was found in his healthy mother and his maternal grandmother, who had CHD. The authors proposed that the identified rearrangements obviously do not affect the coding sequence of GATA4, and may, therefore, not manifest with CHD, but rather disrupt regulatory elements controlling gene expression essential in the developing gonad (22). Thus, the phenotypic variability could be explained by genetic modifiers (37, 38). In addition, GATA4 regulates multiple (gonadal) promoters to variable degrees, rendering some promoters more sensitive to haploinsufficiency than others (15–18, 21, 39). Structural changes caused by both the p.Pro226Leu and pTrp228Cys variations were not predicted to be disruptive and core GATA4 structure was not altered. Since the changes were in the DNA interaction sites, it is expected that both p.Pro226Leu and pTrp228Cys mutations could have altered binding and activation of some of GATA4 interaction partners and could also bind to other promoters and potentially change the transcription of several other genes.

In fact, we found segregating genetic variants besides GATA4 in cases 2 and 3 using NGS. In one 46,XY DSD subject without CHD, a heterozygote variant in *LRP4* gene was found. Mutations in *LRP4* have been related to the Cenani–Lenz syndactyly syndrome and disruption of canonical WNT/beta-catenin signaling (OMIM 604270), which is not only important in bone formation but also in sexual development (40).

In our other 46,XY DSD patient without CHD, a heterozygote mutation in *LHCGR* gene was found together with the GATA4 variant. The same inactivating *LHCGR* mutation was previously reported in 46,XY DSD and 46,XX primary amenorrhea (28). LHCGR is essential for fetal differentiation of the neutral external genitalia into the male phenotype. Inactivating, homozygous, or compound heterozygous mutations of *LHCGR* cause resistance to hormonal stimulation and varying degree of Leydig cell hypoplasia in 46,XY subjects (41–43). Looking at the 5′UTR of *LHCGR*, several GATA sites are present suggesting that GATA4 may regulate this gene. Therefore, combined heterozygocity for *GATA4* and *LHCGR* variants in our patient may explain the 46,XY DSD phenotype.

Finally, phenotypical variability with same heterozygous *GATA4* mutation (p.Gly221Arg) observed within same family manifesting with either 46,XY DSD or CHD only, indicates that there might be incomplete penetrance (21). Interestingly, the same observation was made in mice heterozygote for a *Gata4* deletion (44).

In conclusion, detailed characterization of three new 46,XY DSD patients with and without CHD harboring heterozygous *GATA4* missense mutations in comparison to previously reported patients revealed possible explanations for phenotypical variability. Thereby incomplete penetrance, variable sensitivity, and oligogenic mechanisms may be considered. Interestingly, "double hits" combining a *GATA4* variant with *LHCGR* or *LRP4* variants have been found in two individuals with a 46,XY DSD phenotype only.

ETHICS STATEMENT

Approval for this study was obtained from ethical committees related to: (1) Pediatric Endocrinology, Hospital Infantil La Fe, Av. Campanar 21, 46009 Valencia, Spain. (2) Division of Pediatric Pulmology, Allergology, and Endocrinology, St. Anna Children's Hospital, Department of Pediatrics, Medical University of Vienna, Währinger Gürtel 18-20, 1090 Vienna, Austria. (3) Pediatric Endocrinology Section, Cruces University Hospital, BioCruces Health Research Institute, CIBERDEM, CIBERER, UPV/EHU, Plaza de Cruces 12, 48903, Barakaldo, Spain.

AUTHOR CONTRIBUTIONS

Study design: IM, CM, LA, and CF. Clinical work-up: CM, SR, AR, and LC. Genetic work-up: IM, MF-C, NC, AS, and LA. *In vitro* studies: IM, CF. Bioinformatic studies: AP. Data analysis: IM, AP, and CF. Data interpretation: IM, AP, LA, and CF. Manuscript preparation: IM, CM, LA, and CF. Manuscript approval: all.

ACKNOWLEDGMENTS

We thank all patients and their families for providing their clinical and genetic data for this study.

FUNDING

This work was supported in part by grants from the University of the Basque Country UPV/EHU (IT795-13) to LC, the Instituto de Salud Carlos III (www.isciii.es/; Madrid, *Spain*) Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER, http://www.ciberer.es/) (U-712) to MF-C,

the Agency for Management of University and Research Grants (AGAUR; *agaur.gencat.cat*), Barcelona, Spain (2009SGR31) to LA, and by the Beatriu de Pinós Fellowship 2014 (BP-B 00145) (AGAUR, Catalonia, Spain), the Instituto de Salud Carlos III (www.isciii.es/; Madrid, *Spain*) Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER; http://www.ciberer.es/) (U-712) and the private Foundation Bangerter-Rhyner (http://www.bangerter-stiftung.ch), Basel, Switzerland to NC. IL is supported by a personal research fellowship from the Spanish Pediatric Endocrine Society.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fendo.2018.00142/full#supplementary-material.

REFERENCES

1. Hughes IA, Houk C, Ahmed SF, Lee PA; Lawson Wilkins Pediatric Endocrine Society/European Society for Paediatric Endocrinology Consensus Group. Consensus statement on management of intersex disorders. *J Pediatr Urol* (2006) 2(3):148–62. doi:10.1016/j.jpurol.2006.03.004
2. Bashamboo A, McElreavey K. Mechanism of sex determination in humans: insights from disorders of sex development. *Sex Dev* (2016) 10(5–6):313–25. doi:10.1159/000452637
3. Simon MC. Gotta have GATA. *Nat Genet* (1995) 11(1):9–11. doi:10.1038/ng0995-9
4. Molkentin JD. The zinc finger-containing transcription factors GATA-4, -5, and -6. ubiquitously expressed regulators of tissue-specific gene expression. *J Biol Chem* (2000) 275(50):38949–52. doi:10.1074/jbc.R000029200
5. Morrisey EE, Ip HS, Tang Z, Parmacek MS. GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily. *J Biol Chem* (1997) 272(13):8515–24. doi:10.1074/jbc.272.13.8515
6. Fujiwara Y, Browne CP, Cunniff K, Goff SC, Orkin SH. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci U S A* (1996) 93(22):12355–8. doi:10.1073/pnas.93.22.12355
7. Evans T, Felsenfeld G. The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* (1989) 58(5):877–85. doi:10.1016/0092-8674(89)90940-9
8. Viger RS, Mertineit C, Trasler JM, Nemer M. Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Mullerian inhibiting substance promoter. *Development* (1998) 125(14):2665–75.
9. Kuo CT, Morrisey EE, Anandappa R, Sigrist K, Lu MM, Parmacek MS, et al. GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev* (1997) 11(8):1048–60. doi:10.1101/gad.11.8.1048
10. Molkentin JD, Lin Q, Duncan SA, Olson EN. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev* (1997) 11(8):1061–72. doi:10.1101/gad.11.8.1061
11. Eggers S, Sinclair A. Mammalian sex determination-insights from humans and mice. *Chromosome Res* (2012) 20(1):215–38. doi:10.1007/s10577-012-9274-3
12. Bouma GJ, Washburn LL, Albrecht KH, Eicher EM. Correct dosage of Fog2 and Gata4 transcription factors is critical for fetal testis development in mice. *Proc Natl Acad Sci U S A* (2007) 104(38):14994–9. doi:10.1073/pnas.0701677104
13. Crispino JD, Lodish MB, Thurberg BL, Litovsky SH, Collins T, Molkentin JD, et al. Proper coronary vascular development and heart morphogenesis depend on interaction of GATA-4 with FOG cofactors. *Genes Dev* (2001) 15(7):839–44. doi:10.1101/gad.875201
14. Tevosian SG, Albrecht KH, Crispino JD, Fujiwara Y, Eicher EM, Orkin SH. Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development* (2002) 129(19):4627–34.
15. Viger RS, Guittot SM, Anttonen M, Wilson DB, Heikinheimo M. Role of the GATA family of transcription factors in endocrine development, function, and disease. *Mol Endocrinol* (2008) 22(4):781–98. doi:10.1210/me.2007-0513
16. Nishida H, Miyagawa S, Vieux-Rochas M, Morini M, Ogino Y, Suzuki K, et al. Positive regulation of steroidogenic acute regulatory protein gene expression through the interaction between Dlx and GATA-4 for testicular steroidogenesis. *Endocrinology* (2008) 149(5):2090–7. doi:10.1210/en.2007-1265
17. Miyamoto Y, Taniguchi H, Hamel F, Silversides DW, Viger RS. A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. *BMC Mol Biol* (2008) 9:44. doi:10.1186/1471-2199-9-44
18. Fluck CE, Miller WL. GATA-4 and GATA-6 modulate tissue-specific transcription of the human gene for P450c17 by direct interaction with Sp1. *Mol Endocrinol* (2004) 18(5):1144–57. doi:10.1210/me.2003-0342
19. Eggers S, Sadedin S, van den Bergen JA, Robevska G, Ohnesorg T, Hewitt J, et al. Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. *Genome Biol* (2016) 17(1):243. doi:10.1186/s13059-016-1105-y
20. Harrison SM, Cronberg CF, Kenys M, Hill M, Grimsby GM, Baker LA. DNA copy number variations in patients with 46,XY disorders of sex development. *J Urol* (2014) 192(6):1801–6. doi:10.1016/j.juro.2014.06.040
21. Lourenço D, Brauner R, Rybczynska M, Nihoul-Fékété C, McElreavey K, Bashamboo A. Loss-of-function mutation in GATA4 causes anomalies of human testicular development. *Proc Natl Acad Sci U S A* (2011) 108(4):1597–602. doi:10.1073/pnas.1010257108
22. White S, Ohnesorg T, Notini A, Roeszler K, Hewitt J, Daggag H, et al. Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. *PLoS One* (2011) 6(3):e17793. doi:10.1371/journal.pone.0017793
23. Camats N, Pandey AV, Fernandez-Cancio M, Andaluz P, Janner M, Toran N, et al. Ten novel mutations in the NR5A1 gene cause disordered sex development in 46,XY and ovarian insufficiency in 46,XX individuals. *J Clin Endocrinol Metab* (2012) 97(7):E1294–306. doi:10.1210/jc.2011-3169
24. Krieger E, Vriend G. New ways to boost molecular dynamics simulations. *J Comput Chem* (2015) 36(13):996–1007. doi:10.1002/jcc.23899
25. Canutescu AA, Shelenkov AA, Dunbrack RL. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci* (2003) 12(9):2001–14. doi:10.1110/ps.03154503
26. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* (2003) 24(16):1999–2012. doi:10.1002/jcc.10349
27. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* (1983) 79(2):926–35. doi:10.1063/1.445869
28. Latronico AC, Anasti J, Arnhold IJ, Rapaport R, Mendonca BB, Bloise W, et al. Brief report: testicular and ovarian resistance to luteinizing hormone caused by inactivating mutations of the luteinizing hormone-receptor gene. *N Engl J Med* (1996) 334(8):507–12. doi:10.1056/NEJM199602223340805
29. Hirayama-Yamada K, Kamisago M, Akimoto K, Aotsuka H, Nakamura Y, Tomita H, et al. Phenotypes with GATA4 or NKX2.5 mutations in familial atrial septal defect. *Am J Med Genet A* (2005) 135(1):47–52. doi:10.1002/ajmg.a.30684
30. Garg V, Kathiriya IS, Barnes R, Schluterman MK, King IN, Butler CA, et al. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* (2003) 424(6947):443–7. doi:10.1038/nature01827
31. McCulley DJ, Black BL. Transcription factor pathways and congenital heart disease. *Curr Top Dev Biol* (2012) 100:253–77. doi:10.1016/B978-0-12-387786-4.00008-7
32. Nemer G, Fadlalah F, Usta J, Nemer M, Dbaibo G, Obeid M, et al. A novel mutation in the GATA4 gene in patients with Tetralogy of Fallot. *Hum Mutat* (2006) 27(3):293–4. doi:10.1002/humu.9410
33. Bose D, Valgundan D, Shetty M, Krishnappa J, Kutty AVM. Identification of intronic-splice site mutations in GATA4 gene in Indian patients with

congenital heart disease. *Mutat Res* (2017) 803–805:26–34. doi:10.1016/j.mrfmmm.2017.08.001
34. Zhang X, Wang J, Wang B, Chen S, Fu Q, Sun KA. Novel missense mutation of GATA4 in a Chinese family with congenital heart disease. *PLoS One* (2016) 11(7):e0158904. doi:10.1371/journal.pone.0158904
35. Yang YQ, Wang J, Liu XY, Chen XZ, Zhang W, Wang XZ, et al. Novel GATA4 mutations in patients with congenital ventricular septal defects. *Med Sci Monit* (2012) 18(6):CR344–50. doi:10.12659/MSM.882877
36. Pulignani S, Vecoli C, Sabina S, Foffa I, Ait-Ali L, Andreassi MG. 3'UTR SNPs and haplotypes in the GATA4 gene contribute to the genetic risk of congenital heart disease. *Rev Esp Cardiol (Engl Ed)* (2016) 69(8):760–5. doi:10.1016/j.rec.2016.03.004
37. Bashamboo A, Ledig S, Wieacker P, Achermann JC, McElreavey K. New technologies for the identification of novel genetic markers of disorders of sex development (DSD). *Sex Dev* (2010) 4(4–5):213–24. doi:10.1159/000314917
38. Bashamboo A, Brauner R, Bignon-Topalovic J, Lortat-Jacob S, Karageorgou V, Lourenco D, et al. Mutations in the FOG2/ZFPM2 gene are associated with anomalies of human testis determination. *Hum Mol Genet* (2014) 23(14):3657–65. doi:10.1093/hmg/ddu074
39. Bouchard MF, Taniguchi H, Viger RS. The effect of human GATA4 gene mutations on the activity of target gonadal promoters. *J Mol Endocrinol* (2009) 42(2):149–60. doi:10.1677/JME-08-0089
40. Vainio S, Heikkila M, Kispert A, Chin N, McMahon AP. Female development in mammals is regulated by Wnt-4 signalling. *Nature* (1999) 397(6718):405–9. doi:10.1038/17068
41. Martens JW, Verhoef-Post M, Abelin N, Ezabella M, Toledo SP, Brunner HG, et al. A homozygous mutation in the luteinizing hormone receptor causes partial leydig cell hypoplasia: correlation between receptor activity and phenotype. *Mol Endocrinol* (1998) 12(6):775–84. doi:10.1210/mend.12.6.0124
42. Latronico AC, Arnhold IJ. Gonadotropin resistance. *Endocr Dev* (2013) 24:25–32. doi:10.1159/000342496
43. Segaloff DL. Diseases associated with mutations of the human lutropin receptor. *Prog Mol Biol Transl Sci* (2009) 89:97–114. doi:10.1016/S1877-1173(09)89004-2
44. Jay PY, Bielinska M, Erlich JM, Mannisto S, Pu WT, Heikinheimo M, et al. Impaired mesenchymal cell function in Gata4 mutant mice leads to diaphragmatic hernias and primary lung defects. *Dev Biol* (2007) 301(2):602–14. doi:10.1016/j.ydbio.2006.09.050

Conflict of Interest Statement: The authors report no conflict of interest in this work.