GENETICS OF MÜLLERIAN APLASIA

Maria Sandbacka

Folkhälsan Institute of Genetics
Department of Medical Genetics, Haartman Institute and
Research Programs Unit, Genome-Scale Biology
University of Helsinki,
Helsinki, Finland

ACADEMIC DISSERTATION

To be publically discussed, with the permission of
the Medical Faculty of the University of Helsinki,
in Biomedicum Helsinki, Haartmaninkatu 8, lecture hall 3,
on November1st 2013, at 12 noon.

Helsinki 2013

Supervised by

Docent Kristiina Aittomäki, M.D, Ph.D.
Department of Genetics, HUSLAB
Helsinki University Central Hospital
Folkhälsan Institute of Genetics
Department of Medical Genetics, Haartman Institute and
Research Programs Unit, Genome-Scale Biology
University of Helsinki
Helsinki, Finland

Docent Hannele Laivuori, M.D, Ph.D.
Department of Medical Genetics, Haartman Institute
University of Helsinki and
Department of Obstetrics and Gynecology
Helsinki University Central Hospital
Helsinki, Finland

Reviewed by

Docent Miina Ollikainen, Ph.D.
Hjelt Institute, Department of Public Health
University of Helsinki
Helsinki, Finland

Docent Leila Unkila-Kallio, M.D., Ph.D.
Department of Obstetrics and Gynecology
Helsinki University Central Hospital
Helsinki, Finland

Official opponent

Professor Outi Hovatta, M.D., Ph.D.
Department of Clinical Sciences, Intervention and Technology (CLINTEC)
Karolinska Institutet, Stockholm and
Karolinska University Hospital Huddinge
Sweden

ISBN 978-952-10-9332-6 (paperback)
ISBN 978-952-10-9333-3 (PDF)
http://ethesis.helsinki.fi

Unigrafia Oy
Helsinki 2013

To Anders, Ellen and Axel

Contents

List of original publications

This thesis is based on the following four publications, which are referred to in the text by their Roman numerals. In addition, some unpublished results are presented.

I. **Sandbacka M,** Painter J, Puhakka M, Halttunen M, Laivuori H, Aittomäki K. Does the Y chromosome have a role in Müllerian aplasia? *Fertil Steril,* 94:120-5, 2010.

II. **Sandbacka M**, Bruce S, Halttunen M, Puhakka M, Lahermo P, Hannula-Jouppi K, Lipsanen-Nyman, M, Kere J, Aittomäki K, Laivuori H. Methylation of *H19* and its imprinted control region (*H19* ICR1) in Müllerian aplasia*. Fertil Steril*, 95:2703-6, 2011.

III. **Sandbacka M,** Halttunen M, Jokimaa V, Aittomäki K, Laivuori H. Evaluation of *SHOX* copy number variations in patients with Müllerian aplasia. *Orphanet J Rare Dis.* 6:53, 2011.

IV. **Sandbacka M**, Laivuori H, Freitas É, Halttunen M, Jokimaa V, Morin-Papunen L, Rosenberg C, Aittomäki K. *TBX6*, *LHX1* and copy number variations in the complex genetics of Müllerian aplasia. *Orphanet J Rare Dis.* 8:125, 2013.

Abbreviations

aCGH	array comparative genomic hybridization
Alk3	activin receptor-like kinase-3
AMH	anti-Müllerian hormone (also see MIF and MIS)
Amhr2	anti-Müllerian hormone receptor type 2 (see also Misr2)
BMP	bone morphogenetic protein
Bmpr1a	type 1 bone morphogenetic protein receptor
BWS	Beckwith-Wiedemann syndrome
CE	coelomic epithelium
CpG	cytosine-guanine dinucleotide
CTNNB1	catenin (cadherin-associated protein), beta 1
Dach1	dachshund 1
Dach2	dachshund 2
DES	diethylstilbestrol
DGV	Database of Genomic Variants
DLGH1	discs, large homolog 1
Dll1	delta-like 1
DNA	deoxyribonucleic acid
Emx2	empty spiracles homeobox 2
ESR1	estrogen receptor 1
ETENE	National Advisory Board on Social Welfare and Health Care Ethics
FDR	false discovery rate
FRD	female reproductive duct
FSH	follicle-stimulating hormone
GATA4	GATA binding protein 4
gDNA	genomic DNA
H19	H19, imprinted maternally expressed transcript (non-protein coding)
HGP	Human Genome Project
Hnf1β	HNF1B homeobox B, alias hepatic nuclear factor 1 (see also Tcf2)
Hoxa10	homeobox A10
Hoxa11	homeobox A11
Hoxa13	homeobox A13
ICR	imprinting control region
IM	intermediate mesoderm
IVF	in vitro fertilization
K-S	Kolmogorov-Smirnov statistical test
LAMC1	laminin, gamma 1 gene
LEF	lymphocyte enhancer factor
LH	luteinizing hormone
LHX1/lhx1	LIM homeobox 1 (human) / LIM homeobox protein 1 (mouse; see also Lim1)
Lim1	LIM homeobox protein (mouse)
MA	Müllerian aplasia
MAF	minor allele frequency
MD	Müllerian duct
MIF	Müllerian inhibiting factor

MIS	Müllerian inhibiting substance
Misr2	Müllerian inhibiting substance receptor type 2
MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2
MLPA	multiplex ligation-dependent probe amplification
Mmp2	matrix metallopeptidase 2
MRKH	Mayer-Rokitansky-Küster-Hauser
MURCS	Müllerian duct aplasia, Renal dysplasia and Cervical Somite
M-W U	Mann-Whitney U statistical test
OMIM	Online Mendelian Inheritance In Man
PAR1	pseudoautosomal region
PAX2	paired box 2
PBX1	pre-B-cell leukemia homeobox 1
PCR	polymerase chain reaction
qPCR	quantitative real-time PCR
RAR α,β,γ	retinoic acid receptors, subtypes α,β,γ
RARG	retinoic acid nuclear receptor gamma
RT-PCR	reverse transcriptase PCR
RXRA	retinoid X receptor, alpha
SHOX	short stature homeobox
SRS	Silver-Russel syndrome
SNP	single nucleotide polymorphism
Tcf2	T-cell factor 2
TAR	thrombocytopenia-absent radius syndrome
TF	transcription factor
TSPY1	testis-specific protein, Y-linked
WHO	World Health Organization
WD	Wolffian duct
Wnt4	wingless-related MMTV integration site 4
Wnt5a	wingless-related MMTV integration site 5a
Wnt7a	wingless-related MMTV integration site 7a
Wnt9b	wingless-related MMTV integration site 9b

Abstract

Müllerian aplasia (MA) is a congenital female reproductive disorder featured by loss of a functional uterus and vagina in otherwise healthy females. The prevalence world-wide is estimated to be at least 1:5000 female births. Only a few MA patients have been described with mutations or copy number variations (CNVs) and therefore the cause of the disorder is unknown for the majority of patients.

The aim of this study was to learn more about the genetics of MA. This was accomplished by studying candidate genes, CNVs and methylation defects, as well as by searching for new genes that could be involved in the development of the disorder. As a platform for the study, a large sample set of Finnish MA patients with well-characterized clinical data was used.

Initially the involvement of the male Y chromosome was studied in MA patients, because of the MA phenotype. During male sexual development, the Müllerian ducts (MDs) regress as a result of anti-Müllerian hormone (AMH) secreted by the developing testes, which in turn develops under the regulation of Y chromosomal genes, most importantly the testis-determining gene *SRY*. Testis-specific protein 1-Y-linked (*TSPY1*) was investigated as well as a large number of additional Y-chromosomal loci. The results were negative with no detectable presence of the studied fragments, indicating that Y chromosomal gene regulation is not a cause of the disorder in this patient cohort.

Epigenetic factors changing the expression of genes instead of the structure of the DNA itself have been implicated as a cause of MA. Therefore, DNA methylation studies of the imprinting control region (ICR1) of *H19*, (H19, imprinted maternally expressed transcript (non-coding)) and the gene itself were performed in order to evaluate its role in the development of MA. Aberrant methylation levels of *H19* ICR1 were not detected in MA patients, however aberrant methylation within *H19* was observed.

CNVs of the short stature homeobox gene (*SHOX*) were evaluated in the Finnish cohort, because of a finding suggesting *SHOX* duplications as MA causative. However, none of the Finnish patients showed presence of *SHOX* CNVs, possibly indicating population differences in the underlying cause of MA.

Finally, a genome-wide search for novel CNVs revealed nine rare CNVs in eight out of 50 (16%) studied patients. Of these nine CNVs, two had been previously reported in association with MA, namely deletions in 16p11.2 and 17q12. Further CNV screening in an enlarged patient sample set revealed four more 16p11.2 deletions, resulting in a total of 5/112 (4.5%)

MA patients with deletions in the region. The 16p11.2 region includes one gene with known function in embryonic development, the T-box gene *TBX6*, which therefore was chosen for further screening by Sanger sequencing. By this method, a novel splice site mutation 5´of exon 5 was found in two patients. In addition, two rare polymorphisms were significantly more common in patients than in controls, suggestive for a role in MA. LIM homeobox 1 (*LHX1)*, located within the 17q12 deletion region and recently reported as causative for MA in two patients, was also sequenced and three novel variants in five MA patients were identified, thereby strengthening its importance in the development of MA.

The major result of this study was the finding of a new gene, *TBX6*, linked to MA. Additionally, novel *LHX1* variants were found in MA patients, as well as rare CNVs. Furthermore, 4/112 (3.6%) patients were shown to carry both a *TBX6* or *LHX1* variant and a CNV, highlighting the complex and multifactorial background of MA.

So far, the genetics of MA has proved challenging to decipher. Here we were able to add a new gene, *TBX6*, to the few previously associated with MA and confirm the previously identified CNVs in 16p11.2 and 17q12 and *LHX1* mutations, in a small number of patients. This study strengthens the hypothesis of the complex inheritance of the condition. It also highlights the importance of validating new results in independent patient series as well as the importance of reporting both positive and negative results to the scientific community. The results of this study are inspiring for future research and will pave the way for new gene and mutation discoveries underlying MA.

1. Introduction

Throughout life, fertility and the ability to reproduce have been keystones for survival. In modern times, reproduction of each individual is not a prerequisite for the human population to survive, however, infertility and problems related to reproduction have become a growing concern in society.

Globally, infertility due to male and female factors affects one in every four couples in developing countries (World Health Organization, WHO). Female infertility can manifest as an inability to become pregnant, maintain a pregnancy or carry a pregnancy to a live birth. Müllerian anomalies are congenital disorders including malformations of the female reproductive duct (uterus, oviducts and vagina) associated with a wide range of fertility problems, ranging from increased risk of miscarriage to absolute infertility. Müllerian anomalies are identified in approximately 2-3% of females and the number might be even higher because these anomalies are frequently undiagnosed. Genetic factors have been suggested to contribute to the development of Müllerian anomalies (Hammoud *et al.* 2008).

Müllerian aplasia (MA) is a congenital disorder of the female reproductive duct characterized by lack of functional uterus and vagina in otherwise healthy females. It has an immense impact on a woman´s life due to infertility and inability to have a normal sex life without treatment. Psychosocial problems are also commonly associated with MA. The frequency of MA is estimated to be at least 1 in 5000 female births.

This study was conducted in order to gain further knowledge about the genetics behind MA and how it affects the normal development of the female reproductive organ. Although that may not help the treatment of this disorder, it would be of great importance to affected females to understand how MA develops and why they have been born with this abnormality.

2. Review of the literature

The development of a human being starts from fertilization and continues into adulthood. This literature review focuses on the development of the **female reproductive duct (FRD)** that takes place in the early embryo, and the disturbances in relation to that development.

2.1 Development of the female reproductive duct

The female **urogenital system** develops from the intermediate mesoderm (IM) of the embryo and includes the kidneys, ovaries, the urinary system and the FRD, which is composed of the oviducts (Fallopian tubes), uterus, cervix and vagina. The FRD derives from the Müllerian or paramesonephric ducts (MDs), which are dual ducts formed as invaginations of the coelomic epithelium (CE) of the developing urogenital ridge. The invagination of the MDs starts around embryonic day 11.5 (E11.5) in mice (Cunha 1975, Kobayashi *et al.* 2003). The formation of the MDs can be considered as a three-step event (Figure 1). In the first phase, cells in the CE are fated to become MD specific. This is evident as a thickening of CE cells and LIM homeobox protein 1 (*Lim1,* also called *Lhx1)* and paired box gene 2 (*Pax2*) expression at the site of MD initiation (Figure 1a) (Kobayashi *et al.* 2004, Orvis *et al.* 2007). In the second phase (Figure 1b), expression of wingless-related MMTV integration site 4 (*Wnt4*) from the mesonephros (primitive kidney) or the CE initiates the invagination of the specified CE cells and the MDs elongate caudally until they reach the Wolffian ducts (WDs) (Vainio *et al.* 1999, Orvis *et al.* 2007). The WDs (also called mesonephric ducts, because they connect the mesonephros to the cloaca) are the primordial anlage for the male reproductive system developing into the epididymis, vas deferens and seminal vesicle (Roberts *et al.* 2002).

The third and last phase of MD formation includes the elongation of the MDs (Figure 1c). This happens in close contact with the elongation of the WDs, and until recently it was thought that WD cells contribute to the formation of the MDs (Gruenwald 1941, Vainio *et al.* 1999, Kobayashi *et al.* 2005, Orvis *et al.* 2007). Experimental work performed on chicken and mouse embryos have more recently ruled out the necessity of WD derived cells for successful MD formation (Guioli *et al.* 2007, Orvis *et al.* 2007). However, expression of wingless-related MMTV integration site 9b (*Wnt9b*) in the WD cells seems to affect MD formation in a paracrine manner (Carroll *et al.* 2005). The elongation of the MDs proceeds posteriorly towards the urogenital sinus, which is derived from the endodermal germ layer. In the 8th week of human development (~E12.5 in mice), the MDs join and fuse with each other at the midline, giving rise to the characteristic morphological shape of the FRD with a lower/caudal/posterior one-luminar tube (upper vagina, cervix and uterus) and the

upper/cephalic/ anterior non-fused region (Fallopian tubes and infundibulum). The elongation of the MDs is completed by E13.5 in mice (Kobayashi *et al.* 2003, Kurita 2011, Fritsch *et al.* 2012). The lower part of the vagina is derived from the urogenital sinus with endodermal origin (Cunha 1975). Interactions between the urogenital sinus, MDs and WDs have been shown to be important for vaginal development (Fritsch *et al.* 2012).

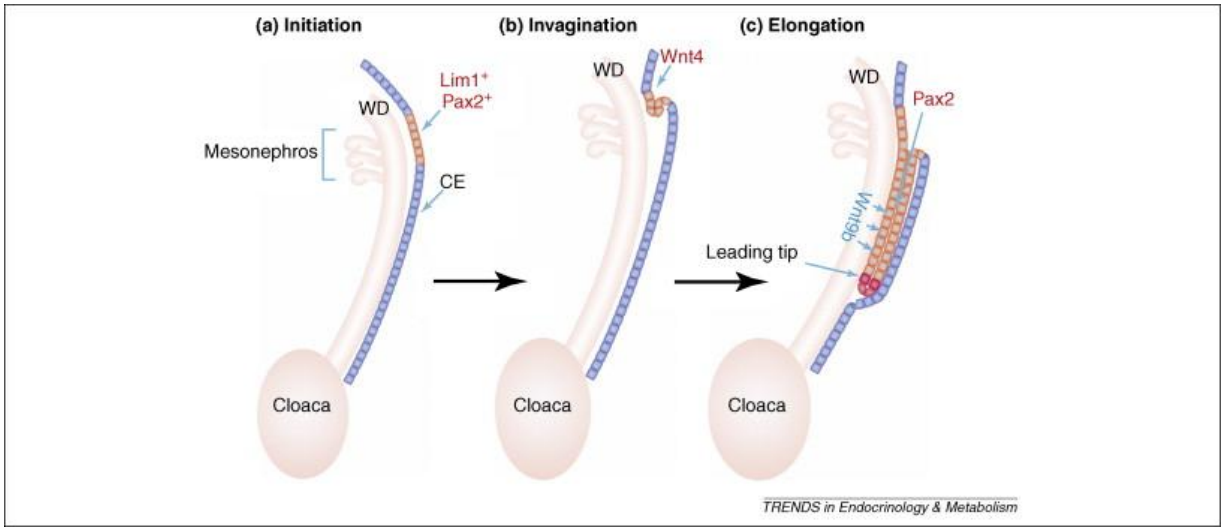

Figure 1: A schematic view of the development of the female reproductive duct (FRD). In the first phase (a) coelomic epithelium (CE) cells expressing *Lim1* and *Pax2* become specified for Müllerian duct (MD) formation (shown in orange). In the second phase (b) invagination occurs due to *Wnt4* expression and in the third phase (c) elongation of the MDs toward the urogenital sinus proceeds. *Pax2* is involved in the MD maintenance and elongation. *Wnt9b* is secreted from the Wolffian ducts (WDs) to promote MD elongation. Genes expressed in the WDs are shown in blue and in both WDs and MDs in red. Reprinted from Trends in Endocrinology & Metabolism, Volume 20, Ma L, Endocrine disruptors in female reproductive tract development and carcinogenesis, 357-363, Copyright (2009), with permission from Elsevier.

The MDs and the WDs are primordial ducts existing side by side in the developing embryo. It is not until the stage of sex differentiation that one of duct systems develops further, whereas the other one regresses.

2.1.1 Genetics of sex determination

Initially the embryo has bipotential gonads and possesses the possibility of developing into either sex. In humans, the sex is determined genetically with XX individuals becoming females and XY individuals becoming males. If the embryo expresses the Y-chromosomal testis-determining gene *SRY* the gonads differentiate into testes and start secreting testosterone, AMH (also called Müllerian inhibiting substance, MIS, or Müllerian inhibiting

factor, MIF) and insulin-growth factor (InsI3) (Nef *et al.* 2000). Testosterone promotes WD differentiation into the organs of the male reproductive system (epididymes, vas deferentia and seminal vesicles). AMH is a transforming growth factor-β (TGF-β) superfamily member secreted by the Sertoli cells in the testes (Josso *et al.* 1993). AMH eliminates the MDs (Behringer *et al.* 1994) resulting in absence of MD derivatives in the embryo, which in mice is completed at embryonic day 16.5 (Kobayashi *et al.* 2004). InsI3 is, together with testosterone and AMH, involved in the descent of the testis (Nef *et al.* 2000). During normal female development, with absence of the Y chromosomally located *SRY*, the male gonadal hormones are not produced and the anlage for the male reproductive system, the WDs, regress. The ovaries produce estrogen that enables the development of the MDs into uterus, Fallopian tubes, cervix and upper two thirds of the vagina (Figure 2) (reviewed in Kobayashi *et al.* 2003, Matzuk *et al.* 2008). The origin of the lower part of the vagina has been an issue of debate, but the general understanding is that it is a derivative of the urogenital sinus (reviewed in Kurita, 2011).

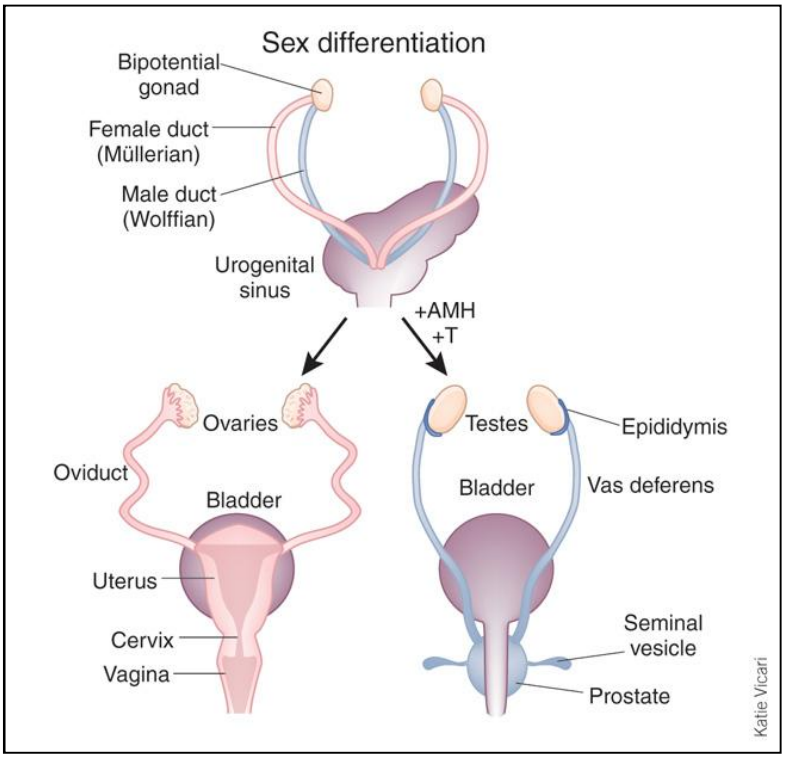

Figure 2: Development of the normal female and male reproductive ducts during sexual differentiation. AMH = Anti-Müllerian Hormone; T = Testosterone. Reprinted by permission from Macmillan Publisher Ltd: [Nature Medicine](Matzuk MM and Lamb DJ. The biology of infertility: research advances and clinical challenges. *Nat.Med.* 14:1197-1213), copyright (2008).

In puberty the adolescent takes a big leap into young adulthood, which involves many physical and psychosocial changes. In girls, the initial sign of puberty is normally breast and pubic hair development, followed by a growth spurt and menarche (onset of menses). The

normal age of pubertal onset is 8 years (Marshall *et al.* 1969, Nebesio *et al.* 2007) and puberty lasts on average about 3 years. However, individual differences vary greatly and the normal range is wide. Ethnicity along with genetic and environmental factors is known to impact pubertal timing (Euling *et al.* 2008). The Tanner stages describing five stages of breast and pubic hair development is widely used for assessing the various steps of sexual development during adolescence (Tanner 1962, Marshall *et al.* 1969). Female puberty is driven by a normally functioning hypothalamic-pituitary-ovarian axis with associated hormones, such as gonadotrophin releasing hormone (GnRH), follicle-stimulating hormone (FSH) and luteinizing hormone (LH), stimulating *i.e.* estrogen production, and follicular and endometrial growth (Nebesio *et al.* 2007, Bordini *et al.* 2011).

2.1.2 Genes and pathways in the development of the female reproductive duct

Although the development of the FRD is fairly well-known, the regulation of the genes involved in FRD development is still far from fully established. Most of the genes known to be involved in FRD regulation have emerged from mouse studies, where specifically mutant mouse strains with urogenital phenotypes have been informative. Genes required for successful development of the mouse FRD are briefly introduced in the following sections and their expression in the mouse reproductive duct is summarized in Table 1.

Lim1 encodes a transcription factor (TF) essential for head and kidney development (Shawlot *et al.* 1995). *Lim1* is also expressed in the IM differentiating into the WDs and the MDs (Barnes *et al.* 1994, Tsang *et al.* 2000). The female *Lim1* knockout mouse lacks a uterus and oviducts, but possesses ovaries (Kobayashi *et al.* 2004). ***Pax2***, also encoding a TF, is important for multiple steps of urogenital development. The *Pax2* homozygous (-/-) mutant mouse has absent kidneys and ureters. The WD and MD derivatives form initially but soon regress, resulting in complete lack of genital tract in both sexes (Torres *et al.* 1995). Empty spiracles homeobox 2 (***Emx2****)* mutant mice die soon after birth due to urogenital failure. Apparently, the initiated WDs degenerate and the MDs never start to form (Miyamoto *et al.* 1997). ***Wnt4*** is crucial for MD initiation, male sexual differentiation and female germ line maintenance. This was shown with a *Wnt4* knockout mouse where females are strikingly masculinized with absence of MDs and presence of WDs. The mutant also had activated testosterone production and diminished oocyte production (Vainio *et al.* 1999). Notable is that all four genes (*Lim1*, *Pax2*, *Emx2* and *Wnt4*) are involved both in MD formation and kidney development, suggesting shared underlying mechanisms in both pathways (Stark *et al.* 1994, Shawlot *et al.* 1995, Torres *et al.* 1995, Miyamoto *et al.* 1997, Tsang *et al.* 2000).

Two other genes from the *Wnt* family gene pathway, ***Wnt7a*** and ***Wnt9b***, homologues to *Drosophila* segment polarity genes, are crucial in MD formation. Mutant mouse studies have shown *Wnt7a* to be important for MD regression in normal male development. The female mutant mouse showed impaired development of the MDs, resulting in infertility due to absent oviducts and a less muscular and slender uterus. The male mutant showed persistence of the MDs (Parr *et al.* 1998). *Wnt9b* is expressed in both sexes in the WD epithelium during mouse E9.5-E14.5 and necessary for MD elongation. The mutant mouse model showed multiple urogenital defects, including lack of uterus and upper vagina in females (Carroll *et al.* 2005).

Hepatic nuclear factor 1β (***Hnf1β***, also called *Tcf2*), is a major player in epithelial cell development during organogenesis in several organs with tubular structures. Expression of *Hnf1β* is especially evident from an early phase in the development of the urogenital structures in mouse (Coffinier *et al.* 1999, Reber *et al.* 2001). In humans, *HNF1β* has been associated with maturity-onset diabetes of the young (MODY), with diabetes mellitus, renal cysts and other renal malformations (Kolatsi-Joannou *et al.* 2001, Bingham *et al.* 2002). Interestingly, also malformations of the FRD, such as bicornuate uterus, uterus didelphys (Bingham *et al.* 2002) and MA (Lindner *et al.* 1999) have been found in some of the patients with *HNF1β* mutations. *HNF1β* is also a risk gene for several forms of cancer, including endometrial and ovarian cancer (Kato *et al.* 2009). ***Dach1*** and ***Dach2***, both members of the *Dachshund* gene family of conserved transcriptional cofactors, have been shown as important players in MD development. Interestingly, the single *Dach1* or *Dach2* knockout mouse does not exhibit a FRD phenotype, but the combined mutant (*Dach1/Dach2*) has a severely disrupted FRD development. Also *Lim1* and *Wnt7a* expression is abnormal in these double knockouts (Davis *et al.* 2008).

In mouse mutants with deficient ***Wnt5a*** expression, the posterior part of the FRD (cervix and vagina) fails to form. The anterior parts of the FRD (Fallopian tubes and uterine horns) are present, however the uterine horns have a reduction in length and luminal changes compared to wild-type (Miller *et al.* 1998, Mericskay *et al.* 2004). Retinoic acid nuclear receptors (***RARs***, subtypes *α,β,γ*), which are transcriptional transducers of the retinoic signal, regulate development of several tissues and organs. Compound mutant mouse models of RARs have a wide range of developmental defects affecting neck, trunk and abdominal regions including defects of the urogenital system. Depending on the genotype of the double mutant, the defects vary from mild aplasia of the kidneys and apparently normal FRD to renal aplasia and lack of all MD derivatives (Mendelsohn *et al.* 1994). Renal and urogenital defects are also found in discs large homolog 1 (***Dlgh1***) mutants, such as hypoplastic kidneys and ureters, and absent vagina and seminal vesicles (Iizuka-Kogo *et al.*

2007). ***β-catenin*** knockouts are found with defective oviduct patterning and MD formation (Arango *et al.* 2005, Deutscher *et al.* 2007).

Table 1: Genes involved in mouse reproductive duct development, their expression in the Müllerian (MD) and Wolffian (WD) ducts and their female urogenital phenotype.

Gene	Expression	Knockout phenotype	Reference
Lim1	MD, WD	no uterus, no oviducts, normal ovaries	Kobayashi *et al.* 2004
Pax2	MD, WD	no kidneys, no ureters, no MD or WD derivatives	Torres *et al.* 1995
Emx2	MD, WD	no MD, no WD	Miyamoto *et al.* 1997
Wnt4	MD	no MDs but WDs (masculinized), ↑testosterone, ↓oocyte production	Vainio *et al.* 1999
Wnt7a	MD	(F) no oviducts, uterine aberrations (M) MD persistence	Parr & McMahon 1998
Wnt9b	WD	no upper vagina, no uterus	Carroll *et al.* 2005
Hnf1β/Tcf2	MD, WD	ND	Reber *et al.* 2001
Dach1	MD	no disruption in FRD	Davis *et al.*2008
Dach2	MD	no disruption in FRD	Davis *et al.*2008
Dach1/ Dach2	MD	the double knockout has hypoplastic oviducts and uterus, vaginal aplasia	Davis *et al.*2008
Wnt5a	MD	no cervix, no vagina, uterine horn malformation	Miller *et al.* 1998
RAR α,β,γ	ND	wide range from normal to lack of all MD derivatives	Mendelsohn *et al.* 1994
Dlgh1	ND	hypoplastic kidneys and ureters, no vagina	IIzuka-Kogo *et al.* 2007
β-catenin	MD	oviduct malformation, hypotrophic uterine horns	Deutscher *et al.*2007
Wt1	ND	no kidneys, no gonads	Kreidberg *et al.* 1993
Lamc1	MD, WD	no uterus, occasionally no oviducts	Willem *et al.* 2002
Hoxa10	MD (uterus)	no uterotubal junction, upper uterus with oviduct appearance	Benson *et al.* 1996
Hoxa11	MD (lower uterine segment and cervix)	reduced stroma and gland development in uterus	Gendron *et al.* 1997
Hoxa13	MD (ectocervix and upper vagina)	no caudal portion of MD	Warot *et al.* 1997
Pbx1	ND	no kidneys, no MDs	Schnabel *et al.* 2003b

ND = no data concerning expression in the reproductive duct or urogenital status
(F) =female, (M) = male

Wilm's tumour nephroblastoma, the most common intra-abdominal solid tumour in children, is associated with mutations in Wilms tumor 1 (***WT1)*** (Reddy *et al.* 1996). *Wt1* mutant mice fail to develop kidneys and gonads, suggesting a role for the gene in urogenital development (Kreidberg *et al.* 1993). Laminin, gamma 1 gene (***Lamc1)*** is a basal membrane component important in organ and tissue development. *Lamc1* mutant mice were found with blind-ending WDs and MDs and the females lack a uterus and occasionally also oviducts (Willem *et al.* 2002), suggesting that the gene has a role in urogenital development.

Last but not least, defects of the developmental homeobox (*Hox*) genes ***Hoxa10***, ***Hoxa11*** and ***Hoxa13*** have been reported as crucial for proper patterning of the FRD (Benson *et al.* 1996, Gendron *et al.* 1997, Warot *et al.* 1997, Zhao *et al.* 2001). The pre-B-cell leukemia homeobox 1 (*Pbx1*) gene is a coactivator of *Hoxa* genes and involved in skeletal development and patterning (Selleri *et al.* 2001), kidney formation (Schnabel *et al.* 2003a) and MD development (Schnabel *et al.* 2003b). Mouse studies have shown that *Pbx1* expression is essential for successful development of the urogenital ridge and of multiple organs evolving from that (Schnabel *et al.* 2003b).

From the aforementioned genes relevant in the development of the FRD three main pathways can be distinguished. These are the **AMH pathway**, the ***Wnt* pathway** and the ***Hoxa* pathway** (Table 1). All three are crucial for MD formation and all three have been in key focus when studying the underlying genetic defects of the FRD.

The AMH pathway is of interest for FRD development due to its key function in males, which is regression of the MDs. AMH signaling is mediated through a type 2 receptor (*Amhr2*, also called *Misr2*) expressed during fetal development in the mesenchyme cells surrounding the developing MDs and in the MD (Mishina *et al.* 1996, Nef *et al.* 2000, Kobayashi *et al.* 2011). Mutations in *AMH* or *AMHR2* cause persistent Müllerian duct syndrome (OMIM 261550) with presence of uterus and Fallopian tubes in males, who often have cryptorchidism (undescended testes) with otherwise normal male genitalia (Imbeaud *et al.* 1994). *Amh*-mutant male mice also have uterus and oviducts, but fully descendent testes (Behringer *et al.* 1994). AMH is probably the only ligand for *Amhr2*, because both *Amh* and *Amhr2* mutant male mice have the same phenotype (Mishina *et al.* 1996). The binding of AMH to *Amhr2* recruits a type 1 receptor mediating the AMH signal *in vivo*. Type 1 bone morphogenetic protein (BMP) receptor, *Bmpr1a* (also known as activin receptor-like kinase-3, *Alk3*), has been identified as such an AMH type 1 receptor, because conditional male mutants were found with uterus and oviducts (Jamin *et al.* 2002). The matrix metalloproteinase 2 (*Mmp2*) has been suggested as a downstream component of the AMH pathway involved in the apoptotic events causing MD degeneration, but the mutant *Mmp2* male mice showed no urogenital phenotype and thereby the mechanisms for the epithelial regression remains

unknown (Roberts *et al.* 2002). In women, AMH is mainly expressed by the granulosa cells in the ovaries and has been suggested as a marker of ovarian follicle reserve and female fertility (Bentzen *et al.* 2013).

The *Wnt* genes are a large family of secreted protein growth factors highly conserved between vertebrate species and acting in a wide variety of roles during development. *Wnt* signals are transduced through different intracellular pathways, of which the canonical "Wnt/β-catenin" pathway primarily regulates cell fate during development (Miller 2002). In the presence of a *Wnt* ligand and its binding to the Frizzled receptor, β-catenin is accumulated in the cytoplasm and imported to the nucleus where it serves as a transcriptional coactivator of TFs for the T-cell factor (TCF)/lymphocyte enhancer factor (LEF) family (reviewed in Rao *et al.* 2010).

Taken together, at least four *Wnt* family members (*Wnt4*, *Wnt5a*, *Wnt7a* and *Wnt9b*) are known to be important for successful FRD development (Miller *et al.* 1998, Parr *et al.* 1998, Vainio *et al.* 1999, Mericskay *et al.* 2004, Carroll *et al.* 2005). Interestingly, members of the *Wnt* gene family seem to be capable of substituting each other. This was shown in the work of Carroll *et al*, where the induction of the mesonephric and metanephric tubules and the elongation of the MD in *Wnt9b* mutants were rescued by *Wnt1* expression (Carroll *et al.* 2005). A link between the AMH and Wnt/β-catenin pathways was established when β-catenin was shown to mediate AMH signaling during MD regression in normal male development (Kobayashi *et al.* 2011). Thereby, β-catenin was shown to have dual roles in reproductive duct development: one for MD differentiation in females (Arango *et al.* 2005, Deutscher *et al.* 2007) and the other for MD regression in males (Kobayashi *et al.* 2011).

The *Hoxa* pathway is especially important in regulating the spatiotemporal interactions between TFs and signaling molecules in order to obtain the correct compartmentalization of the FRD. *Hoxa9* is expressed in the Fallopian tubes, *Hoxa10* and *Hoxa11* in the uterus, and *Hoxa13* in the upper vagina (Taylor *et al.* 1997).

Cooperation between the abovementioned genes and pathways as well as between factors acting downstream and upstream of them is fundamental for proper FRD development.

2.1.3 Epigenetics and genomic imprinting

The epigenetics phenomenon was suggested in 1942 by Conrad Waddington as a mechanism causing heritable alterations in gene function (expression) which do not involve changes in DNA sequence. There are three distinct forms of epigenetic regulation, namely

DNA methylation, histone modification and non-coding RNA (reviewed in Inbar-Feigenberg *et al.* 2013). This review discusses DNA methylation and imprinting.

DNA methylation is the most studied epigenetic phenomenon to date and involves cysteine methylation by DNA methylatransferase (DNMT) at cytosine-guanine dinuclotides (CpGs) in differentiated cells (Chen *et al.* 2011). Approximately 30 million CpG sites are found in the human genome, *i.e.* they cover about 1% of the genome (Fouse *et al.* 2010). The CpG sites are primarily located in GC rich promoter regions of genes, where, in general, a low level of DNA methylation means transcriptional activity of the gene and a high level of DNA methylation means gene silencing (reviewed in Inbar *et al.* 2013).

Genomic imprinting occurs during early development and is an epigenetic phenomenon, where only one parental allele of a gene is expressed and the other is silenced. The regulation of genomic imprinting is through epigenetic mechanisms, which involves differential DNA methylation at so called differentially methylated regions (DMRs). DMRs are defined as either germline DMRs, where methylation is inherited from the male or female gamete and is maintained throughout development, or tissue-specific DMRs, at which the parental-specific methylation mark is set after fertilization. DMRs act as imprinting control regions (ICRs) regulating the parental-of-origin manner of gene expression of one or several imprinted genes (John *et al.* 1996, Inbar-Feigenberg *et al.* 2013).

Genetic factors are considered the major player in human disorders, but epigenetic factors (epimutations) are known to also contribute. Epimutations may alter gene expression, leaving the DNA sequence intact. Moreover, they can be heritable through cell divisions, and are thereby good candidates for explaining the etiology of some disorders or aberrant phenotypes. Lynch syndrome, a form of hereditary nonpolyposis colon cancer (HNPCC) was one of the first examples where epimutations, namely in the DNA mismatch repair gene *MLH1* (mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)), were associated with cancer susceptibility (Gazzoli *et al.* 2002, Peltomaki 2012). Another example is the congenital Silver-Russell syndrome (SRS), a growth restriction syndrome, where hypomethylation of the ICR of the *H19* gene is the major cause of the disorder (Penaherrera *et al.* 2010). Interestingly, SRS patients with skeletal and urogenital malformations as well as MA have been reported, suggesting a role for epimutations in MA (Bliek *et al.* 2006, Bruce *et al.* 2009).

2.1.4 Abnormalities of the female reproductive duct

Congenital malformations of the FRD are thought to arise during early embryogenesis and often occur in conjunction with specific syndromes, where FRD abnormalities are one symptom among many (Table 2).

Table 2. Syndromes associated with malformations of the FRD.

Syndrome	Phenotype	FRD status	Cause	OMIM[1] Reference
Al-Awadi/Raas-Rothschild	severe malformations of upper limbs, hypoplastic pelvis, abnormal genitalia	MA	Autosomal recessive WNT7A mutation	276820 (Al-Awadi *et al.* 1985)
Androgen insensitivity (testicular feminization)	females with 46,XY, female external genitalia, breast development, abdominal testes and in some cases no pubic and axillary hair	MA	X-linked recessive, androgen receptor (AR) mutation	300068, 312300
Bardet-Biedl	Ciliopathy, renal abnormalities	occasionally vaginal atresia, MA	genetically heterogeneous	209900
CATCH22 (22q11 deletion)	parathyroid hypoplasia, cardiac malformations, cleft palate	occasionally MA	22q11.2del (TBX1 haplo-insufficiency)	188400 (Sundaram *et al.* 2007)
Cat-eye	Coloboma of iris, heart and renal defects, other malformations	occasionally MA	22q11	115470 (Schinzel *et al.* 1981)
Hand-Foot-Genital	small feet with short great toes, abnormal thumbs	genital tract duplication	*HOXA13* mutation	140000
Silver-Russel (SRS)	growth retardation, craniofacial features, variable malformations	occasionally MA	H19 ICR1 hypomethylation (20-60%); matUPD7[2] (10%)	180860 (Bliek *et al.* 2006, Bruce *et al.* 2009)
Tetraamelia	tetraphocomelia, craniophacial abnormalities	urogenital malformations	Autosomal recessive, one family with *WNT3* mutation	273395 (Niemann *et al.* 2004)
Thrombocyto-penia absent radius (TAR)	low platelet count, absence of radius	occasionally MA	1q21.1	274000 (Klopocki *et al.* 2007)
Urogenitaladysplasia (hereditary renal dysplasia)	Lethal renal abnormalities	occasionally MA	22q13.31 UPK3A, PAX2, RET mutations	191830

MA = Müllerian aplasia

[1]OMIM (Online Mendelian Inheritance in Man) phenotype accession number.

[2]matUPD7 = maternal uniparental disomy of chromosome 7

2.2 Müllerian aplasia

2.2.1 Definition and diagnosis

Müllerian aplasia (MA) is defined as congenital aplasia of a functional uterus and vagina with normal female karyotype (46,XX) and secondary sexual characteristics, and usually normal functioning ovaries (Griffin *et al.* 1976, Simpson 1999). MA is mostly diagnosed in puberty when adolescent females are referred to a gynecologist due to primary amenorrea (no menstruation). Pubic and axillary hair growth is normal, as well as the external genitals, but instead of a normal-length vagina only a vaginal pouch extending approximately 1-3 cm is present. Ultrasonography is used to detect the status of the uterus, the Fallopian tubes and the ovaries, and the finding is often confirmed by magnetic resonance imaging (MRI), less often by laparoscopy. Karyotyping is performed to exclude chromosomal abnormalities as a cause of the disorder. Hormonal levels (FSH, LH, estradiol and progesterone) are measured, if clinical signs of acne or hirsutism appear.

According to a population-based study in Finland, the incidence of MA is at least 1:5000 female births (Aittomäki *et al.* 2001) and approximately the same worldwide (Griffin *et al.* 1976, Folch *et al.* 2000, Morcel *et al.* 2007). MA is also referred to as **MURCS association** (Müllerian duct aplasia, Renal dysplasia and Cervical Somite anomalies, OMIM 601076), because renal and skeletal malformations occur in 20-40% of the patients. The malformations are mostly minor including renal aplasia, horseshoe kidney, scoliosis and milder vertebral defects (Griffin *et al.* 1976, Carson *et al.* 1983, Simpson 1999, Pittock *et al.* 2005). Hearing defects (Letterie *et al.* 1991, Strubbe *et al.* 1994), cardiac malformations (Gilliam *et al.* 2002, Kula *et al.* 2004) and digital anomalies (Strubbe *et al.* 1987, Massafra *et al.* 1988) have been reported in some MA cases. Reports of MA patients with associated anorectal malformations (Gilliam *et al.* 2002, Wester *et al.* 2012) or with a uterus (Doyle *et al.* 2009) exist, but these characteristics are not compatible with the clinical definition of MA (Jones 2006).

The most common form of MA is the **Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome**, with its name originating from the four clinicians who originally described the characteristics of MA; namely August Franz Joseph Mayer, Carl Freiherr von Rokitansky, Hermann Küster, and G. A. Hauser. In MRKH, remnants of the uterine cornu are present and connected by a thin streak of connective tissue (Figure 3). The Fallopian tubes are also present on one (unilateral) or both sides (bilateral). MRKH is a heterogeneous group of phenotypes, which can be further divided into subtypes depending on the status and symmetry of Fallopian tubes, uterine remnants and ovaries. Some studies define patients as either **typical** (type A, type I or isolated) including patients with only genital malformations, or **atypical** (type B or

II) including patients with genital malformations and associated characteristics of MURCS (Strubbe *et al.* 1993, Oppelt *et al.* 2006). Depending on the diagnostic method used (laparoscopy, ultrasound, X-ray or magnetic resonance examination) and the amount of clinical data available, the patients cannot always be classified into A or B and are so-called unspecified MRKH. A new classification system called VCUAM (Vagina, Cervix, Uterus, Adnexa and associated Malformations), where malformations can be assigned to the precise organ subgroup, has been proposed in order to provide a clinical classification that more precisely reflects the entire genital malformation (Oppelt *et al.* 2005a). Within this study, all three types of MRKH (type A, B and unspecified) are classified as MA. The rarest form of MA is total MA, where all Müllerian derivatives (uterus, Fallopian tubes and upper vagina) are missing (Figure 3) (Griffin *et al.* 1976).

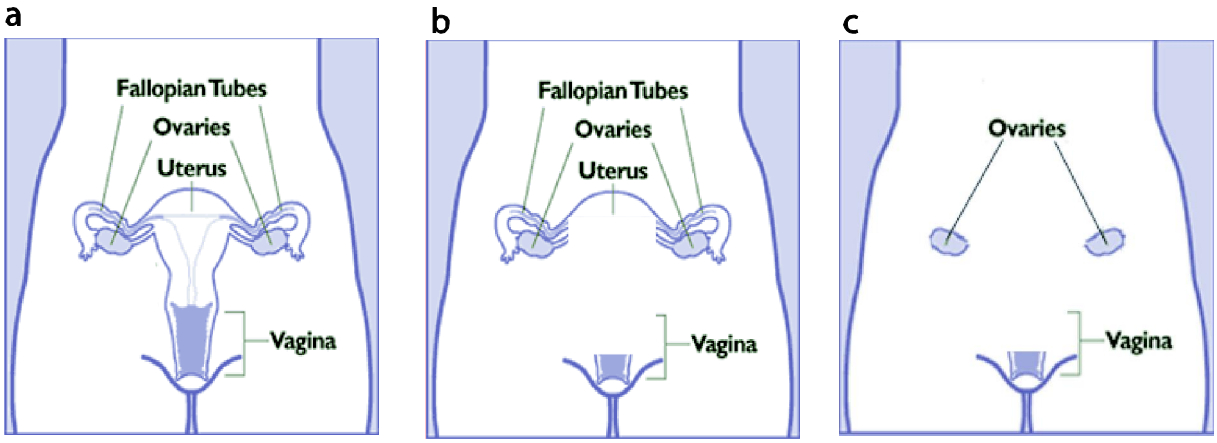

Figure 3. Illustration of the female reproductive duct in a a) normal female structure, b) MRKH and c) total Müllerian aplasia (MA) patient. Figure modified from Wikimedia (http://commons.wikimedia.org/wiki/File:Scheme_female_reproductive_system-en.svg).

MA as the main phenotype in an otherwise healthy woman can be referred to as **non-syndromic**. MA can also occur in conjunction with other syndromes and is then often referred to as **syndromic** MA. Examples of syndromes associated with MA are presented in Table 2. It is not clear if the same genetic variations could underlie both syndromic and non-syndromic forms of MA. This thesis, however, is focused on the non-syndromic form of MA.

In conjunction with the structural abnormalities of MA, it is important to note the psychosocial elements of the disorder. MA is mostly diagnosed in adolescence, at the age of 14-16 years, in the period of transition from childhood to adulthood. It is a sensitive period in development, when the body as well as the mind is going through many changes and self-esteem is forming. The diagnosis of MA, to be found without uterus and normal vagina with associated infertility is profoundly traumatizing. This diagnosis, at such a vulnerable age, must have a large impact on a girl´s life. Questions, such as how is normal sex life possible, will the structural abnormality hinder finding of a spouse and establishing a family, are all

relevant issues to a woman with MA. All these aspects make MA one of the most difficult congenital female reproductive disorders to deal with.

2.2.2 Treatment

The vaginal aplasia in MA patients can be treated by nonsurgical or surgical methods. Timing for treatment is best planned when the patient is emotionally mature and desires correction. The first-line method in Finland and several other countries is the nonsurgical so call Frank´s method (Frank 1938, Committee on Adolescent Health Care 2013). It is based on dilation of the vagina from its original length (1-3 cm) until about 10 cm, which is considered normal for sexual intercourse. The dilation is done by the patient herself, using a series of dilators with growing size. The method requires personal engagement of the patient, with daily self-dilation for 30 minutes to 2 hours for several months to years. However, complications associated with surgery can be avoided when using this method (Laufer 2002, Committee on Adolescent Health Care 2013). The success rate of achieving an anatomically functioning vagina using vaginal dilation in highly motivated patients is reported to be 90-95% (Roberts *et al.* 2001, Edmonds *et al.* 2012). Surgical methods are available for patients who are unsuccessful with dilators or for other reasons prefer this alternative. All methods are based on creation of a neovagina either by using skin or peritoneum implantation (e.g. William´s and Davydov´s method), small bowel or by gradual mechanical stretching of vaginal pouch using a traction device (Vecchietti´s method), followed by use of vaginal dilators post-operatively. The surgical methods can be painful with risk of scarring and incontinence among other complications (Committee on Adolescent Health Care 2013, Pizzo *et al.* 2013).

For the infertility, the only existing treatment is surrogacy. Utilizing in vitro fertilization (IVF) surrogacy enables MA patients to have their own biological children with their spouse as they usually have normal functioning ovaries (Beski *et al.* 2000). At present, the legislation does not allow this treatment in many countries, Finland being among those. An initiative from the National Advisory Board on Social Welfare and Health Care Ethics (ETENE, http://www.etene.fi/en) concerning surrogacy treatment in Finland has been put forward to the Ministry of Justice in order to allow women infertile due to absence of uterus to have surrogacy. Uterus transplantation would be another possibility for infertility treatment. To date these procedures have been unsuccessful, and would require an immense amount of research and ethical discussion in order to become a realistic options for the patients (Brannstrom 2013, Del Priore *et al.* 2013). Even more far-reaching is the possibility of artificial wombs, by which the embryo would grow and develop outside of the human body in an artificial uterus (Bulletti *et al.* 2011).

2.2.3 Genetic background

Pattern of inheritance

Most MA patients are sporadic cases, but familial occurrence has also been documented and affected sib-pairs have been reported in several publications (Jones *et al.* 1972, Griffin *et al.* 1976, Shokeir 1978, van Lingen *et al.* 1998, Morcel *et al.* 2007, Gervasini *et al.* 2010). Based on a study in Saskatchewan with 16 families, where several female family members were affected with different degrees of MA whereas the males showed no deleterious effects, a sex-limited autosomal dominant inheritance pattern was suggested (Shokeir 1978). However, other studies were not able to support this mode of inheritance (Carson *et al.* 1983, van Lingen *et al.* 1998).

Additionally, in the study by Petrozza *et al*. including 34 surrogate pregnancies, with oocytes from women with MA , 17 baby girls and 17 baby boys were born, of which all were healthy, except one male child with a middle ear defect. None of the girls had MA, making it rather unlikely that the disorder would be caused by a single dominant mutation (Petrozza *et al.* 1997). The most logical explanation for MA is a multifactorial or polygenic inheritance, which in rare traits such as MA would have a low recurrence risk (1-2%) for first-degree relatives. However, in some instances or certain populations, a dominant or recessive mode of inheritance is possible (Simpson 1999).

Candidate genes

Several candidate genes have been suggested for MA. Most of these genes are members of the *AMH*, *HOXA* or *WNT* pathways, or other known genes with crucial roles during early development of the female urogenital structures. *AMH*, which is of interest due to its role in MD regression in males (Behringer *et al.* 1994), and its receptor *AMHR2* were primarily investigated in MA patients, but no mutations were found (Resendes *et al.* 2001, Zenteno *et al.* 2004, Oppelt *et al.* 2005b). Candidate gene studies including *HOXA7*, *HOXA9*, *HOXA10*, *HOXA11* and *HOXA13*, which are all required for successful MD formation, were also negative (Karnis *et al.* 2000, Timmreck *et al.* 2003, Burel *et al.* 2006, Lalwani *et al.* 2008, Dang *et al.* 2012, Ekici *et al.* 2013). Mutation screening efforts in MA patients involving *WT1* (van Lingen *et al.* 1998), the retinoic acid nuclear receptors *RARG* and *RXRA* (Cheroki *et al.* 2006), β-catenin (*CTNNB1*) (Drummond *et al.* 2008), *PBX1* (Ma *et al.* 2011), *PAX2* (Burel *et al.* 2006, Wang *et al.* 2012) *DLGH1* and *LAMC1* (Ravel *et al.* 2012) were also unable to find pathogenic gene variations.

Contradictory results regarding a possible association between decreased levels of the galactose-1-phosphate uridyl transferase (GALT) enzyme and MRKH have been reported (Cramer *et al.* 1996, Klipstein *et al.* 2003, Zenteno *et al.* 2004). Cramer *et al.* reported a *GALT*

N314D mutation in 6/13 (46%) MRKH patients compared to 16/113 (14%) controls. The authors argued that decreased maternal or fetal *GALT* expression due to the N314D mutation could result in vaginal agenesis (Cramer *et al.* 1996). However, two other studies in 15 and 32 MA patients, respectively, failed to replicate these results, thereby suggesting that decreased *GALT* expression does not associate with MA (Klipstein *et al.* 2003, Zenteno *et al.* 2004). The carrier frequency of the N314D variation has been reported as high in controls in other studies (Morland *et al.* 1998, Stefansson *et al.* 2001) and the association Cramer *et al.* found might therefore be due to their low patient number (Zenteno *et al.* 2004).

The first positive finding of a mutation in a patient with MA was reported in 2004. Biason-Lauber and coworkers screened *WNT4* in a girl with MRKH and unilateral renal agenesis (Biason-Lauber *et al.* 2004). They found a heterozygous loss-of-function mutation (E226G) situated in exon 5 of the gene (Figure 4). The mutation was shown to prevent correct lipid modification. The mutant molecule was less hydrophobic than the normal one, and thereby trapped in the cell. The decreased *WNT4* expression resulted in impaired MD formation as well as elevated androgen production. The phenotype of the patient was strikingly similar to the mouse mutant model *Wnt4-/-* (Vainio *et al.* 1999, Biason-Lauber *et al.* 2004). Shortly after this report, three more heterozygous missense *WNT4* mutations were reported in three patients with MRKH and androgen excess (Figure 4) (Biason-Lauber *et al.* 2007, Philibert *et al.* 2008, Philibert *et al.* 2011). However, *WNT4* mutations were not found in other cohorts of patients with the classical form of MRKH without androgen excess (Clement-Ziza *et al.* 2005, Ravel *et al.* 2009). Therefore, MRKH in conjunction with hyperandrogenism can be regarded as an entity of its own or as a rare subtype of MA (OMIM 158330) (Clement-Ziza *et al.* 2005, Biason-Lauber *et al.* 2007). Genetic studies of other *WNT* genes (*WTN5A*, *WNT7A* and *WNT9B)* in MA patients have been negative (Timmreck *et al.* 2003, Ravel *et al.* 2009, Dang *et al.* 2012).

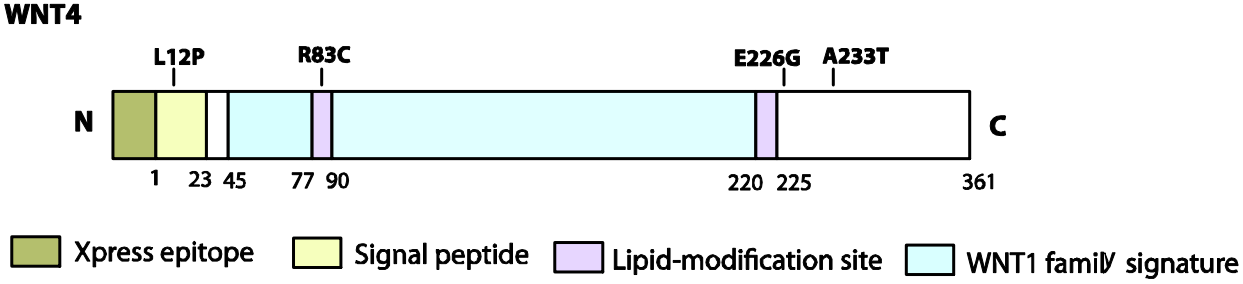

Figure 4: *WNT4* mutations found in patients with MRKH and hyperandrogenism. The mutations are all heterozygous, nonsynonumous and located in exons 1 (L12P; Philibert *et al.* 2008), 2 (R83C; Biason-Lauber *et al.* 2007) and 5 (E226G; Biason-Lauber *et al.* 2004 and A233T; Philibert *et al.* 2011) of the gene. Figure modified from Biason-Lauber *et al.* (2004).

Recently, mutations in another gene, *LHX1,* were observed in two patients with MA (Ledig *et al.* 2011, Ledig *et al.* 2012). The mutations were a missense mutation in exon 4 (p.R264G) and a heterozygous adenine duplication at position 25 in exon 1, leading to a frame shift and premature stop codon at amino acid 33 (c.25dup; p.Arg9LysfsX25). The patient with the exon 4 mutation was reported as MRKH type I, whereas the patient with the exon 1 mutation was reported as MRKH type II with unilateral renal agenesis, both largely resembling the *Lhx1* mutant mouse phenotype with renal agenesis and absent uterus and oviducts (Kobayashi *et al.* 2004). *LHX1* is a member of the LIM homeodomain TF family containing two LIM domains, a central homeodomain possessing DNA-binding activity and a transactivation domain at the C-terminus (Bozzi *et al.* 1996). Downstream targets of *LHX1* have not been reported to date (Ledig *et al.* 2012). In mouse, *Lhx1* expression starts in the MD epithelium at E11.5 in both sexes. In females, the expression persists at least until E16.5, whereas in males it starts to weaken at E14.5 corresponding to the ongoing MD regression (Kobayashi *et al.* 2004).

Presence of *TSPY1* in patients with MA has also been suggested to play a part in the development of MA. This is primarily based on a study by Plevraki and coworkers, where four out of six MA patients were reported to carry fragments of *TSPY1* (Plevraki *et al.* 2004). *TSPY1* has been suggested as a candidate gene for gonadoblastoma and to have a proliferative role in spermatogenesis (Lau *et al.* 2009).

Copy number variations

Several CNVs (microdeletions or microduplications) have been reported in association with MA (Table 3). Most of these are rare, observed only in one or two patients and have not been reported in the Database of Genomic Variants (DGV). Therefore, their importance in the etiology of MA is inconclusive. However, four chromosomal regions, 1q21.1, 16p11.2, 17q12 and 22q11.2, stand out from the rest because they have been found in three, four, nine and six patients, respectively (Cheroki *et al.* 2006, Klopocki *et al.* 2007, Sundaram *et al.* 2007, Cheroki *et al.* 2008, Bernardini *et al.* 2009, Ledig *et al.* 2011, Nik-Zainal *et al.* 2011).

1q21.1 is a known locus for TAR (OMIM 274000) characterized by reduction of platelet number and absent radius (a bone in the forearm) and two TAR patients were reported with MA (Klopocki *et al.* 2007, Ledig *et al.* 2011). One patient with complete uterine and vaginal aplasia in conjunction with fused external labia and undetected ovaries by ultrasound was reported with CNVs in the TAR locus and 22q11.22 (Cheroki *et al.* 2008). The 22q11.2 locus is known for DiGeorge (OMIM 188400), velocardiofacial (OMIM 192430) and 22q11.2 distal deletion syndrome (Ben-Shachar *et al.* 2008) characterized by a variable phenotype including facial dysmorphic features, congenital heart defects and behavioral difficulties. Four patients were reported with 22q11.2 deletion and MA (Cheroki *et al.* 2006, Sundaram

et al. 2007, Cheroki *et al.* 2008, Nik-Zainal *et al.* 2011), whereas one MA patient was reported with duplications in 22q11.21 and 12q23.1 (Ledig *et al.* 2011).

Table 3. Summary of copy number variations (CNVs) detected in patients with MA.

Locus	CNV	Size	Patients	Phenotype	Reference
1q21.1	del	~0.5 Mb	2	TAR, MA	Klopocki *et al.* 2007, Ledig *et al.* 2011
1q21.1 and 22q11.22	dup dup	2.7 Mb 0.6 Mb	1	syndromic MA	Cheroki *et al.* 2008
2q11.2	dup	1.30 Mb	1	MA	Nik-Zainal *et al.* 2011
2q13 and 17q12	del del	0.1 Mb 1.4 Mb		MA	Ledig *et al.* 2011
3p21.31	dup	0.1 Mb	1	MA	Ledig *et al.* 2011
4q34→qter	del	ND	1	MA[a]	Bendavid *et al.* 2007
4q32.2	del	0.34 Mb	1	MA	Ledig *et al.* 2011
4q28.3	del	0.1 Mb	1	MA	Ledig *et al.* 2011
6p21.2, 6q25.1 and 6q25.2	dup dup dup	0.2 Mb 0.4 Mb 0.4 Mb	1	MA	Ledig *et al.* 2011
10q24.33	dup	56 kb	1	MA	Ledig *et al.* 2011
12q23.1 and 22q11.21	dup dup	0.1 Mb 0.4 Mb	1	MA	Ledig *et al.* 2011
12q24.12 and Xp11.3	dup dup	0.1 Mb 0.2 Mb	1	MA	Ledig *et al.* 2011
16p11.2	del	~0.55 Mb	4	MA	Nik-Zainal *et al.* 2011
17q12	del	1.4 Mb	4	MA	Nik-Zainal *et al.* 2011
17q12	del	1.5 Mb	2	MA, one with mild dysmorphism, one with renal cysts	Bernardini *et al.* 2009
17q12	del	1.8 Mb	1	MA	Ledig *et al.* 2011
17q12	del	1.2 Mb	1	syndromic MA (mental imparement)	Cheroki *et al.* 2008
22q11.2	del	0.39 Mb	1	22q11.2 distal syndrome, MA	Nik-Zainal *et al.* 2011
22q11.2	del	ND	2	22q11.2 syndrome, MA	Sundaram *et al.* 2007
22q11.21	del	2.6 Mb	1	22q11.2 syndrome, MA	Cheroki *et al.* 2006, Cheroki *et al.* 2008
Xp11.1	del	0.1 Mb	1	MA	Ledig *et al.* 2011
Xp22.2	dup	0.4 Mb	1	MA	Ledig *et al.* 2011
Xq21.31	del	1 Mb	1	syndromic MA	Cheroki *et al.* 2008

[a] mother had the same deletion but not MA, instead cardiac defect and Fallopian tube cancer
ND=no data; TAR= thrombocytopenia, absent radius

Four MA patients without features of other syndromes were recently reported with deletions of 16p11.2 (Nik-Zainal *et al.* 2011), a CNV previously documented in association with autism (Kumar *et al.* 2008, Weiss *et al.* 2008), schizophrenia (McCarthy *et al.* 2009), developmental delay (Rosenfeld *et al.* 2010) and obesity (Walters *et al.* 2010). 17q12 deletions have been reported in six MA patients without additional syndromes (Ledig *et al.*

2011, Nik-Zainal *et al.* 2011) and in three MA patients with either renal cysts, mild dysmorphic features (Bernardini *et al.* 2009) or mental impairment (Cheroki *et al.* 2008). CNVs in 17q12 have previously been associated with several phenotypes including renal malformations and cysts, growth restriction, speech problems (Mefford *et al.* 2007, Nagamani *et al.* 2010) and autism (Loirat *et al.* 2010). The 17q12 region includes *LHX1* and *HNF1β*, for which mutations were described in 4/63 MA patients (Ledig *et al.* 2011, Ledig *et al.* 2012) and in 2/4 patients with mild diabetes and MA (Lindner *et al.* 1999), respectively.

Partial duplications of the short stature homeobox (*SHOX*) gene have been reported in five MA patients (two sporadic, three familial). Two sisters carried the same duplication inherited from their unaffected father, whereas two healthy sisters and their healthy mother did not have the duplication (Gervasini *et al.* 2010). *SHOX* is a homeobox gene located on the pseudoautosomal region (PAR1) of the X (Xp22) and Y (Yp11.3) chromosomes. Mutations and CNVs in *SHOX* have previously been associated with idiopathic short stature (ISS, OMIM 300582), Turner syndrome, dyschondrosteosis (Leri-Weill syndrome, LWD, OMIM 127300) and Langer mesomelic dysplasia (OMIM 249700) (Ellison *et al.* 1997, Belin *et al.* 1998, Shears *et al.* 1998, Rao *et al.* 2010, Benito-Sanz *et al.* 2011).

Epigenetic factors and expression studies

DNA hypomethylation of the imprinting control region ICR1 of *H19* has been associated with genital and skeletal malformations in Silver Russel syndrome (SRS) patients (Bliek *et al.* 2006). Two SRS patients with extreme *H19* ICR1 hypomethylation and MA have been described (Bruce *et al.* 2009), thereby suggesting a role for epigenetic factors in the etiology of MA. Reports of monozygotic twins where one twin is healthy and the other has MA also suggest the possibility of epigenetic involvement (Lischke *et al.* 1973, Steinkampf *et al.* 2003, Duru *et al.* 2009). The surrogate pregnancies with oocytes from MA women resulting in only healthy baby girls (Petrozza *et al.* 1997) also argue for possible epigenetic regulation or epimutation in the development of the disorder.

One study including genome-wide methylation and expression data from patients with MA has been published. This combined methylome and transcriptome study was based on uterine remnant tissue from seven and eight MA patients, respectively. The study revealed nine genes relevant for FRD development that were both differently methylated and differently expressed in patient compared to control samples. Based on these results, the authors suggest GATA binding protein 4 (*GATA4*) and *WT1* as well as the estrogen receptor 1 (*ESR1)* as good candidate genes for further studies (Rall *et al.* 2011). From mouse studies *Gata4* and *Wt1* were shown to form a complex that synergistically binds to the sex

determining gene *Sry* as well as to the *Amh* promoter, thereby affecting the expression of *AMH* responsible for MD regression (Miyamoto *et al.* 2008). Estrogens, on the other hand, are important for normal uterine development. Mutant mouse models have hypoplastic uterus and vagina, illuminating a role for the gene in normal FRD development (Couse *et al.* 1999). The HOXA gene pathway was also highlighted as relevant for FRD development, because both *HOXA5* and *HOXA9* showed aberrant expression in MA patients compared to control samples (Rall *et al.* 2011).

Teratogens

Teratogens are reagents or factors that can cause developmental defects when the fetus is exposed to these substances. Diethylstilbestrol (DES) is a known teratogen and an endocrine disrupting chemical that was frequently prescribed to pregnant women during the 1940s to the 1970s to prevent miscarriage. However, DES was found to have both teratogenic and oncogenic effects on many organs, including the male and female reproductive systems, when exposure occurred *in utero*. DES might not directly affect the formation of the MDs, but from mouse models it is known that DES affects several genes in the *Wnt* and *Hoxa* gene pathways, thereby affecting FRD development (Ma 2009).

2.3 Genetic methods for studying Müllerian aplasia

The human genome comprises approximately 3 billion base pairs (bp) of deoxyribonucleic acid (DNA) tightly packed into 23 pairs of chromosomes (22 pairs of autosomes, one pair of sex chromosomes) in each cell of an individual excluding germ cells. An immense international research effort, the Human Genome Project (HGP) aiming at determining the entire DNA sequence of the human genome and identifying all human genes, was completed in 2003 with the first draft of the genetic structure of the human genome published in 2001 (Lander *et al.* 2001) and a more detailed version in 2007 (Levy *et al.* 2007). The number of genes transcribed from the human genome and translated into functional proteins is estimated to be 20-23 000, much lower than anticipated. However, the noncoding parts of the genome are thought to harbor many regulatory and biochemical functions still to be elucidated. The genes are divided into coding units, exons, and noncoding units, introns. By differential splicing of the introns, several protein products with different lengths and functions and differing time windows of expression can be created. One of the basic features of the human genome is sequence variability between every individual, which is estimated to occur in one base pair per every 1000 nucleotides (The International HapMap Consortium 2005, Ng *et al.* 2008).

2.3.1 Genome-wide approach

Genome-wide approaches are often the starting point for finding genomic variations associated with a given disorder. Genome-wide approaches include array-based platforms utilizing DNA or RNA as the starting material for studying genetic differences.

In MA, array comparative genomic hybridization (aCGH) has been utilized for genome-wide identification of CNVs (microdeletions or microduplications) that might associate with the disorder. This method can identify much smaller changes (within an individual gene) than the classical karyotyping using differently labeled patient and control samples hybridized against up to millions of probes located throughout the genome. The higher the frequency of the probes, the better the sensitivity, coverage and specificity in defining the start and end points of CNVs.

CNVs can also be studied using single nucleotide polymorphism (SNP) genotyping and multiplex ligation-dependent probe amplification (MLPA). SNPs are single-base changes in the DNA and the most common form of variation in the genome. Successive SNP deletions or duplications are indicative of a larger genomic aberration (CNV). SNP genotyping is commonly used for comparing genotype frequencies between patients and controls. Since MA patients are mostly sporadic cases, family studies are rarely possible. MLPA is a multiplex method for detecting intragenic CNVs, *i.e.* gain or loss of exons or entire genes, especially suitable for small-scale screening. MLPA has been used successfully also in MA studies (Gervasini *et al.* 2010).

Gene expression studies based on tissue-extracted RNA are troublesome in MA, because we are investigating a "missing" tissue. However, in one study the remnants of the uterus were collected in laparoscopy-assisted surgery for neovagina construction. DNA extracted from the same specimens was used for genome-wide DNA methylation studies (Rall *et al.* 2011).

Over the past years, next-generation sequencing (NGS) methods have emerged and taken a leader position in finding disease-causing genes. Exome sequencing, *i.e.* sequencing of all protein-coding genes in the genome, is widely used and has been a successful approach in finding the underlying cause for both monogenic and complex disorders. Whole-genome sequencing is more challenging with an enormous amount of data produced. The statistical analyses are laborious, especially when studying multifactorial disorders (Bamshad *et al.* 2012). Studies in MA patients utilizing NGS methods have not yet been published.

2.3.2 Candidate gene approach

Candidate genes are usually chosen based on previous knowledge of metabolic pathways or animal models. In MA, the first candidate studies were based on genes known to affect FRD development from mouse knockout studies. Thus far, all have been negative, except for *WNT4* (reviewed in section 2.2.3).

Another candidate gene approach is to study genes located within CNV regions known to associate with MA. Within these CNV regions, genes previously known to be important in mesodermal or early embryonic development have been chosen for further screening. With this approach, *LHX1* mutations within the 17q12 deletion region were found in women with MA (reviewed in section 2.2.3).

2.3.3 Mutation screening

When a candidate gene of interest has been selected, the next step is to screen for nucleotide changes *i.e.* mutations primarily in the protein-coding DNA sequence segregating with the disease-phenotype, and not found in a selected set of healthy control samples. Polymorphisms, which are nucleotide changes common in the population and not disease-associated, are important to rule out.

By Sanger sequencing different types of mutations can be found. A **point mutation** is a change of one nucleotide in the DNA sequence, through a single base substitution, deletion or insertion. A **silent** or **synonymous mutation** is a point mutation, which does not alter the encoded amino acid and therefore has no effect on the protein function. However, a silent mutation can have a pathogenic effect if located in a functionally important part of the gene. A **nonsynonymous** or **missense mutation** is a point mutation where the nucleotide substitution causes an amino acid change. Missense mutations can be rare polymorphisms, but are more likely to be pathogenic if they occur in a highly conserved base or amino acid, or in a functionally important gene region. A **nonsense mutation** is a point mutation that introduces a premature stop codon in the transcript. Depending on the site in which it occurs, a nonsense mutation can result in a shortened non-functional protein product (Strachan *et al.* 2004). **Frameshift mutations** are caused by deletions, duplications or insertions of one or more nucleotides in the protein-coding DNA sequence. These can result in a disrupted reading-frame, also with a shortened non-functional protein product as the end result. A **splice-site mutation** is a nucleotide change occurring intronically in the highly conserved splice donor site, branch site or splice acceptor site at either 5´ or 3´end of an exon. The mutation causes incorrect splicing of the following exon, often resulting in a premature stop codon and a shortened non-functional protein product (Strachan *et al.* 2004).

3. Aims of the study

The aim of this doctoral thesis was to investigate the genetic background of MA.

The following specific aims were set for the study:

I. To determine if Y chromosomal genetic material and in particular the earlier reported *TSPY1* fragment is found in Finnish MA patients.

II. To investigate MA patient DNA methylation status of the *H19* ICR1 region on 11p15 known to be involved in growth and related developmental disorders.

III. To evaluate *SHOX* CNVs as a cause of MA in Finnish patients.

IV. To find new candidate genes for MA.

4. Patients and methods

4.1 Patients and controls

The Finnish MA project was initiated in the Helsinki University Central Hospital (HUCH)/Department of Obstetrics and Gynecology in the mid-1970s when clinical data collection from MA patients began. Subsequently, detailed clinical data and blood samples were collected through all five university hospitals in Finland from 1978 to 1993. Since the end of the 1990s, the enrolment of MA patients has been ongoing at HUCH. In 2011, a new enrolment of MA patients was started through Oulu University Central Hospital and Turku University Central Hospital.

To date, clinical data has been collected for 250 Finnish patients and two foreign patients originating from Russia and the Middle East. DNA samples are available from 140 patients and 36 relatives (29 mothers, 4 fathers and 3 healthy sisters from 31 families in total). Peripheral blood RNA samples have been collected from only a few patients. The majority of Finnish patients have the MRKH phenotype, while total aplasia comprises less than 10% of MA (personal communication, Minna Puhakka, M.D.) The majority of the MA patients are sporadic cases. However, one sibling pair with one sister having MRKH and the other having total MA was included in the study. Urinary tract malformations including renal defects occur in 24% of the patients and skeletal malformations in 48% of the patients (personal communication, Minna Puhakka M.D.). Considering that MA is a rare syndrome with an incidence of 1:5000 new born girls, the yearly number of girls born in Finland with this syndrome is around six. When considering the sensitivity of the syndrome, the collected sample set and its size in relation to the size of the Finnish population is remarkably large. To our knowledge this series is also one of the largest MA sample sets in the world. A German research group recently published a clinical overview/retrospective study of a MA cohort, comprising 284 MRKH women, the largest cohort thus far published (Oppelt *et al.* 2012). Molecular genetic studies of MA have mostly been performed in smaller cohort sizes including 1-63 patients (Biason-Lauber *et al.* 2004, Plevraki *et al.* 2004, Plevraki *et al.* 2004, Gervasini *et al.* 2010, Ledig *et al.* 2011, Nik-Zainal *et al.* 2011, Philibert *et al.* 2011).

The control group consisted of 200 healthy Finnish females with at least one normal pregnancy. The blood samples were collected at the HUCH. Five healthy male controls were obtained from the Finnish Red Cross for validation of Study I. Detailed data on the number of patients and controls included in each study are given in Table 4.

Table 4: Number of MA patients, healthy relatives, and controls included in each study. Study II and IV include methods for which the investigated number of patients and controls differs from the total number (shown in brackets).

Study	Patients	Healthy relatives	Healthy controls
I	110	20[a]	105[b]
II	83 (80 in qPCR; 38 in HM27)	-	20 (6 in HM27)
III	101	-	115
IV	112 (50 in aCGH; 12 in SNP genotyping; 6 in qPCR; 3 in RT-PCR)	-	200 (100 in MLPA; 150/180 for *LHX1* studies; 6 in qPCR; 3 in RT-PCR)

[a] 13 mothers, 4 fathers, and 3 sisters from different families.
[b] five of the controls are healthy males from the Finnish Red Cross
HM27=Genome-wide Infinium HumanMethylation27 Bead Arrays

4.2 Ethical issues

Informed consent was obtained from all participants prior to enrolment in the study. For patients <18 years, signed consent was also obtained from the guardian. The study protocol has been approved by the Finnish Ministry of Social Affairs and Health and reviewed by the Ethics Committee of the Department of Obstetrics and Gynecology of the HUCH.

4.3 Methods

A summary of all methods used in this thesis project is included in Table 5. The most relevant methods are separately summarized in the following sections.

Table 5. Summary of used genetic methods.

Method	Used in Study
DNA extraction from peripheral blood	I, II, III, IV
RNA extraction from peripheral blood	IV
Polymerase chain reaction (PCR)	I, II, III, IV
Real-time quantitative polymerase chain reaction (q-PCR)	II
Reverse transcriptase-PCR (RT-PCR)	IV
Gel electrophoresis	I, II, III, IV
Sanger sequencing	I, II, III, IV
Multiplex ligation-dependent probe amplification (MLPA)	III, IV
Genome-wide DNA methylation microarrays	II
Comparative genomic hybridization (aCGH)	IV
Genome-wide single-nucleotide polymorphism (SNP) genotyping	IV

4.3.1 Extraction of DNA and RNA from peripheral blood

Genomic DNA was extracted from EDTA or heparin blood collection tubes with the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN) including protein precipitation according to the manufacturer's recommendation or by the phenol-chloroform method. DNA quality and quantity was analyzed by NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Genomic RNA was extracted from PAXgene Blood RNA Tubes (PreAnalytiX Gmbh, Hombrechtikon, Switzerland) by the PAXGene RNA Kit (Qiagen, Hilden, Germany) according to the recommended protocol. The extracted RNA was further DNase-treated using the DNA-*free*™ kit (Applied Biosystems, Foster City, CA, USA) to remove genomic DNA (gDNA).

4.3.2 Polymerase chain reaction

Polymerase chain reaction (PCR) is a well-established method for amplifying a large amount of a specific stretch of DNA, optimally 200-800 bp in length. Ph.D. Kary Mullis was acknowledged with the Nobel Prize in 1993 for the development of the method into an important biochemical technique. The DNA fragment of interest is enzymatically replicated by the use of short oligonucleotides (primers) complementary to the flanking region of the fragment and DNA polymerase through repeated series (thermal cycling) of heating, annealing, extension and cooling. Primers for the DNA fragments of interest were designed using different publically available software programs, such as Primer3 (http://frodo.wi.mit.edu/primer3/input.htm) and ExonPrimer (http://ihg.gsf.de/ihg/ExonPrimer.html). The specificity of the primers was tested and sequence comparisons were performed using Ensembl (http://www.ensembl.org/index.html), UCSC (http://www.genome.ucsc.edu/cgi-bin/hgGateway) or NCBI (http://www.ncbi.nlm.nih.gov/) Genome Browsers. Databases including known SNPs, e.g. dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/), the 1000 Genomes database (http://browser.1000genomes.org/index.html) and Exome Variant Server (EVS, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA; http://evs.gs.washington.edu/EVS/), were used when designing optimal primers with high specificity.

The PCR reactions in this study were run as single (Study I-IV), multiplex (Study I) or nested reactions (Study I). Real-time quantitative PCR (qPCR) assessing the specific amount of an amplified PCR product was used in Study II. Reverse transcriptase PCR (RT-PCR) was utilized in Study IV. RT-PCR enables PCR amplification of RNA through conversion of RNA into complementary DNA (cDNA) by a reverse transcriptase enzyme. Thereafter a standard PCR

reaction can be performed. RT-PCR is especially useful for investing the effects of gDNA mutations on the transcriptional (cDNA) level.

The sizes of the amplified PCR products were determined by the use of gel electrophoresis and ethidium bromide or MidoriGreen DNA stain, which are intercalating agents used as fluorescent tags for nucleic acid binding. The PCR products were then further analyzed by traditional Sanger sequencing using an automated ABI3730XL DNA Analyzer (Applied Biosystems).

4.3.3 Multiplex ligation-dependent probe amplification (MLPA)

MLPA is a multiplex PCR method for determining the relative copy number of several DNA or RNA sequences at a time. In Study III we used the commercially available SALSA MLPA kit for *SHOX* (MRC-Holland, Amsterdam, Netherlands) CNV detection, whereas in Study IV we designed synthetic MLPA probes for *TBX6* to detect intragenic deletions or duplications. The probe design has been described in detail in the Patients and methods' section of publication IV. Each MLPA probe has a unique length, which can be detected using capillary electrophoresis and visualized by e.g. GeneMapper software version 4.0 (Applied Biosystems). Differences in relative copy numbers between the studied probes can be detected using MRC Coffylazer MLPA-Data Software (MRC-Holland) or using a calculation based method (summarized in the Patients and method-section of publication III and IV).

4.3.4 Genome-wide microarray-based methods

Several DNA-based microarray methods were used in this project to screen for genomic aberrations on a genome-wide level. Infinium HumanMethylation27 Bead Arrays (HM27), Illumina, San Diego, CA, USA) were used in Study II for studying DNA methylation. The technology allows single CpG resolution of >27,000 CpG sites using 500 ng of bisulphite-converted DNA as starting material. Treatment of DNA by bisulfite (ZymoResearch, Orange, CA, USA) results in conversion of unmethylated cytosin (C) to uracil, leaving methylated 5-methylcytosin unaffected, thus enabling detection of methylation status in patient versus control samples. HumanOmni2.5-8 BeadChip v1.0 (Illumina) was used for single-nucleotide polymorphism (SNP) genotyping in Study IV. This technique enables high-resolution SNP detection of almost 2.4 million markers, both common and rare minor allele frequencies (MAF, >2.5%) according to 1000 Genomes Project data including CNVs, from 200 ng of DNA. Both array projects were performed by the Technology Centre, Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland and the results were analysed by Genome Studio software (Illumina).

Comparative genomic hybridization (aCGH) arrays (180 K platform, Oxford Gene Technology, Yarnton, Oxford, UK) are widely used for high-resolution detection of CNVs (deletions or duplications) in the genome. The 180 K platform with genome-wide coverage and an overall median probe spacing of 13 Kb was used for CNV screening in Study IV, which was performed in collaboration with Érika Freitas, Ph.D., and Assistant Professor Carla Rosenberg, Ph.D., at the Institute of Biosciences, University of São Paulo, Brazil. Genomic Workbench software (Agilent Technologies, Santa Clara, CA, USA) was used for CNV calling.

4.3.5 Statistical methods

Statistical t-tests (Kolmogorov-Smirnov (K-S) and Mann-Whitney U (M-W U), PASW Statistics 18 software (SPSS, Chicago, IL) were used to test for the statistical significance of the findings. K-S can be used when assessing data sets showing normal distribution, whereas M-W U is optimized for unequal data distribution. Therefore, K-S was utilized in Study II to assess the difference in methylation status in patients compared to controls for most of the *H19* CpG sites, whereas M-W U was utilized in Study II to assess CpG sites showing unequal distribution, as well as to assess the methylation difference in the *H19-HpaII* site 25 between patients and controls. In Study IV differences in SNP allele frequency between patients and controls were calculated using M-W U. Correction for multiple testing adjusts p-values derived from the numerous statistical tests performed. This was performed using the conservative Bonferroni method in Study IV, and in Study II using the false discovery rate (FDR; www.r-project.org), which shows the expected proportion of false positives in the data.

4.3.6 Freely available databases

Several different freely available online databases were utilized throughout this thesis. In addition to those databases already mentioned, UniProtKB (http://www.uniprot.org/) was used for protein sequences and functional information and MutationTaster (Schwarz *et al.* 2010), PON-P (http://bioinf.uta.fi/PON-P/) including PhD-SNP 2.0.6, PolyPhen 2.0.22, SIFT 4.0.3, SNAP 1.0.8 and I-Mutant 3.0.6 programs used for analysis, SpliceMan (http://fairbrother.biomed.brown.edu/spliceman/index.cgi), SplicePort (http://spliceport.cbcb.umd.edu/SplicingAnalyser.html) and Human Splicing Finder (http://www.umd.be/HSF/) software programs were used for *in silico* prediction of genetic variants. CpG Plot (http://www.ebi.ac.uk/Tools/emboss/cpgplot/) and CpG Island Searcher (http://www.uscnorris.com/cpgislands2/cpg.aspx) were used for identifying promoter regions. Database of Genomic Variants (DGV, http://projects.tcag.ca/variation) was used for information on structural variants of the human genome. Online Mendelian Inheritance In Man (http://www.omim.org/) is a freely available online catalogue including human genes and associated genomic disorders.

5. Results and Discussion

5.1 The Y chromosome and Müllerian aplasia (Study I)

In this study the focus was set to investigate if the Y chromosomal DNA, or more precisely, if presence of any Y chromosomal fragments in the genetic make-up of MA patients, could have an effect on the development of the disorder.

We studied DNA from 110 Finnish MA patients and 20 healthy relatives for *TSPY1* and 38 additional loci covering the Y chromosome (Figure 5) by PCR amplification (single, nested and multiplex reactions). The primer information and amplification protocols for 33 of the 38 loci (the so called Y panel) were acquired from the HUCH, Laboratory of Prenatal Diagnostics, where they were routinely used for male infertility screening. Five additional markers were designed within our study in order to improve the coverage of especially the short arm of the chromosome.

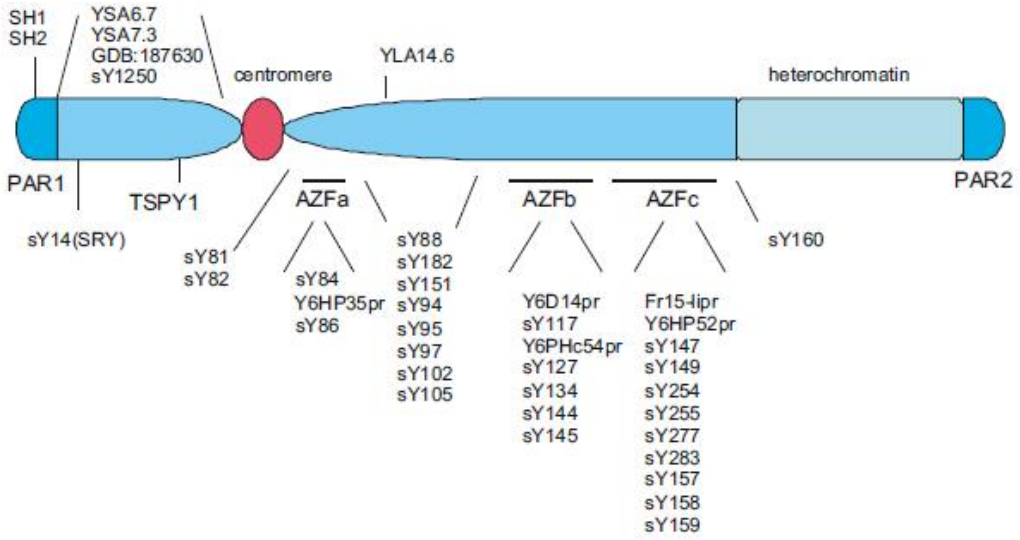

Figure 5: Schematic view of the investigated loci and their locations on the Y chromosome (Sandbacka *et al.* 2009). PAR = pseudoautosomal region, AZF = Azoospermia Factor.

All studied Y chromosomal fragments were absent in the Finnish female patients, indicating that the corresponding Y chromosomal regions are not the cause of MA in the investigated sample set.

Presence of *TSPY1* was previously suggested as the primary cause of MA in the study of Plevraki and colleagues (Plevraki *et al.* 2004). Their result was based on positive amplification by nested PCR in 4/6 MA patients. Nested PCR involves a second round of PCR

amplification using a previous PCR product as starting material. This can increase the yield of an amplicon that is only vaguely, if at all, detected in the first PCR round. In our study, using the same PCR primers as Plevraki and colleagues, the nested PCR reactions for *TSPY1* amplification were unreliable resulting in false positive results due to unspecific binding of the primers. Similar observations were reported by Álvarez-Nava and coworkers, where positive amplification results alternated with negative ones, and two healthy female controls showed positive amplification (Alvarez-Nava *et al.* 2004). Therefore, we designed new primers specific only to *TSPY1*, which replaced the need for nested PCR, and performed single PCRs in all Finnish patients, 100 healthy females who all had biological children, 20 healthy relatives (13 mothers, 4 fathers and 3 sisters) and five healthy males from the Finnish Red Cross as controls. All males (fathers and male controls) were positive for *TSPY1*, whereas all females (patients, female relatives and controls) were negative.

In addition, to further investigate if any other fragments from the Y chromosome might be involved in the development of MA, we studied 38 additional Y loci. None of the female patients showed presence of any of the 38 studied fragments, whereas the male samples showed strong amplification of each of them. However, for a few markers some female samples (both patients and controls) generated PCR products, but these were not the same size as in male samples. Repetitive PCR reactions showed inconsistent results indicating unspecific binding of the PCR primers, confirmed by Sanger sequencing. The reason for the unspecific binding is probably the highly repetitive DNA sequence structure of the Y chromosome, as well as a high frequency of shared genomic sequences between the X and Y chromosomes.

During normal male development the Y-chromosomally expressed *SRY* activates *AMH* expression from the testis resulting in MD regression. Therefore, if fragments with Y-chromosomal origin would be present in MA patients (46,XX), they could cause MD regression and the MA phenotype. In the study by Plevraki and coworkers (Plevraki *et al.* 2004) four out of six MA patients were reported to carry fragments of the Y-chromosomal *TSPY1* supporting above-mentioned hypothesis. However, in our study, which included almost 20 times more patients, none of the patients were positive for *TSPY1* amplification by PCR. This raises the question of the reliability of the results from the Plevraki group. *TSPY1* has previously been suggested as a putative gene for gonadoblastoma (Lau *et al.* 2009) and *TSPY1* CNVs have been suggested to associate with spermatogenesis (Shen *et al.* 2013). Further studies of *TSPY1* or other Y-originating fragments in MA patients have not been reported and no further associations of the gene with MD regression has been shown to date.

Taken together, we investigated if fragments of the Y chromosomal DNA are present in MA patient samples indicating a role for Y chromosomal genes in the etiology of the disorder. Our results clearly show that the studied Y chromosomal regions are not involved in MA, at least not in this Finnish patient cohort. Our results also show the importance of confirming specificity of nested PCR products by sequencing and not by gel electrophoresis alone, as in the study by Plevraki and coworkers (Plevraki *et al.* 2004).

5.2 DNA methylation studies of *H19* in Müllerian aplasia (Study II)

Epigenetic studies, such as DNA methylation measurements, are of importance when studying imprinted genes involved in genetic disorders. *H19* is an example of such a gene, which has been suggested to associate with MA. Namely, in a study by Bliek and coworkers, hypomethylation (low methylation compared to the normal state) of the *H19* imprinting control region ICR1 was shown to associate with Silver Russel syndrome (SRS) patients with genital and skeletal malformations (Bliek *et al.* 2006). Hypomethylation of *H19* (Gicquel *et al.* 2005, Eggermann *et al.* 2006) and maternal uniparental disomy of chromosome 7 (matUPD7) are previously known important causes of SRS (Moore *et al.* 1999), but also genital malformations were reported in these patients. Furthermore, Bruce *et al*. described two Finnish SRS patients with extreme *H19* ICR1 hypomethylation (Bruce *et al.* 2009), and interestingly, both patients also had MA. These findings inspired us to study if 1) hypomethylation of the *H19* ICR1 site 25 is associated with the MA phenotype and if 2) altered methylation of the *H19* gene *per se* is associated with MA.

Eighty Finnish MA patients were studied for DNA methylation status at the *H19* ICR1-*HpaII* site 25 by qPCR, according to Bruce *et al.* (Bruce *et al.* 2008). In practice, this was performed by comparing the relative amount of qPCR products amplified from DNA samples digested with the methylation-sensitive *HpaII* restriction enzyme (unable to cut methylated cytosine) versus undigested samples. The methylation level of the *H19* ICR1 site 25 did not differ significantly between MA patients and controls with mean methylation 47.5% (SD, 7.0%) for patients and 50.2% (SD, 10.8%) for controls (Table 6). The results indicate that hypomethylation of the investigated site is not associated with the MA phenotype.

Table 6. Statistical comparison of methylation values (%) from the qPCR studies of the *H19* ICR1.

Sample	Number (N)	Mean (std)	Median	p value
All MA patients	80	47.5 (7.0)	46.1	0.215[a]
Controls	20	50.2 (10.8)	49.6	

[a]p value derived from Mann-Whitney U test.

H19 belongs to one of the most studied complexes of imprinted genes, the *IGF2-H19* locus on 11p15.5. Insulin-like growth factor 2 (*IGF2*) and *H19* are neighboring genes regulated by the same ICR1 complex (also called the *H19* DMR*)*, which is a paternally methylated germline DMR (Weksberg *et al.* 2002). *IGF2* is expressed from the paternally derived allele, whereas *H19* is expressed from the maternally derived allele. The *IGF2-H19* locus is associated with two clinically opposite disorders, SRS and Beckwith-Wiedemann syndrome (BWS), displaying almost mirror phenotypes. SRS patients suffer from intrauterine and postnatal growth retardation defects (Eggermann 2010), whereas BWS patients suffer from fetal and postnatal overgrowth (Choufani *et al.* 2010). BWS can be caused by hypermethylation of ICR1 on the maternal allele promoting *IGF2* expression from both alleles instead of only one. On the other hand, hypomethylation of ICR1 on the paternal allele promotes expression of *H19* from both alleles, resulting in downregulation of *IGF2* and a SRS phenotype (Eggermann 2010). *CTCF* (or CCCTC-binding factor) is a major mediator of imprinted gene regulation, and hypomethylation of the *HpaII* site 25 situated nearby the CTCF site 3 within the *H19* ICR1 was reported to associate with genital malformations in SRS patients (Gicquel *et al.* 2005, Bruce *et al.* 2009)

To further investigate the general methylation level of *H19*, 38 MA patients and six controls, of which 35 patients and all controls were included in the qPCR study, were screened using HumanMethylation27 methylation arrays (HM27; Illumina). HM27 is a genome-wide methylation bead chip covering >27000 CpG sites, of which 16 sites are situated in *H19*. A total of 3/16 CpG sites investigated showed statistically significant methylation differences (p-value <0.05) between patients and controls (Table 7). The methylation differences between patients and controls were, however, small in all studied sites. Therefore, after FDR correction for multiple testing, the methylation levels were not significant and the effect of this on a transcriptional level is uncertain. The observed methylation level was 61-98% in both patients and controls in all studied sites within *H19* (CpG sites 1-14), whereas the CpG sites 15 and 16 situated 5´upstream of the transcribed region had a lower methylation level (37-75%).

Table 7. Location of studied CpG sites *H19* and their methylation data based on HumanMethylation27 microarray (Illumina).

	Probe name	Location within *H19*[a]	CpG island	MA patients (n=38) mean	(std)	p[b]	q[c]	Controls (n=6) mean	(std)
1	cg11492040	exon 5	1	0.96	(0.01)	0.785	0.897	0.96	(0.01)
2	cg06197492	exon 5	1	0.84	(0.04)	0.150	0.400	0.86	(0.03)
3	cg10602543	exon 5	1	0.82	(0.03)	0.270	0.540	0.83	(0.03)
4	cg23977670	exon 5	1	0.89	(0.02)	0.555	0.807	0.90	(0.01)
5	cg11716026	exon 5	1	0.83	(0.04)	0.416	0.666	0.79	(0.10)
6	**cg22172494**	**exon 3**	**2**	0.81	(0.05)	**0.012**	**0.107**	0.86	(0.04)
7	cg26808784	intron 2-3	2	0.91	(0.02)	0.971	0.971	0.91	(0.01)
8	**cg25852472**	**intron 1-2**	**2**	0.81	(0.03)	**0.014**	**0.107**	0.84	(0.04)
9	cg15269875	exon 1	2	0.93	(0.01)	0.111	0.355	0.93	(0.01)
10	cg21167159	exon 1	2	0.89	(0.03)	0.908	0.969	0.89	(0.03)
11	cg13145013	exon 1	2	0.96	(0.01)	0.782	0.897	0.96	(0.01)
12	cg15317267	exon 1	2	0.92	(0.01)	0.183	0.418	0.93	(0.01)
13	**cg04715462**	**exon 1**	**2**	0.98	(0.01)	**0.020**	**0.107**	0.98	(0.00)
14	cg07342901	exon 1	[d]	0.93	(0.02)	0.606	0.808	0.92	(0.03)
15	cg17769238	5´upstream	3	0.59	(0.05)	0.369	0.656	0.57	(0.03)
16	cg02657360	5´upstream	4	0.48	(0.05)	0.063	0.252	0.44	(0.04)

[a] according to Ensembl GRCh37(hg19).
[b] p value derived from independent samples t-test performed between MA patients and controls.
[c] p value after false discovery rate (FDR) correction.
[d] not located on a known CpG island.

H19 is a non-coding RNA abundantly expressed in mammalian development with largely unknown biological function. *H19* belongs to a network of imprinted genes associated with embryonic growth (Gabory *et al.* 2009, Gabory *et al.* 2010) and it is therefore possible that the observed methylation reduction has an impact on expression of other genes acting within the same network and with a function in MD development. It has also been shown that *H19* acts as a tumour suppressor (Yoshimizu *et al.* 2008), and that the first exon of the gene encodes a microRNA that regulates placental growth, which might be relevant for normal growth of the fetus (Keniry *et al.* 2012). The three CpG sites showing significant differences in p-values between patients and controls in the HM27 study are all situated within the same CpG island and one of the sites is located in exon 1 (Figure 6). One can therefore speculate that reduced methylation levels within the *H19* gene itself could contribute to the MA phenotype. The results should, however, be validated in a larger patient cohort including more controls in order to obtain robust statistical significance for the findings.

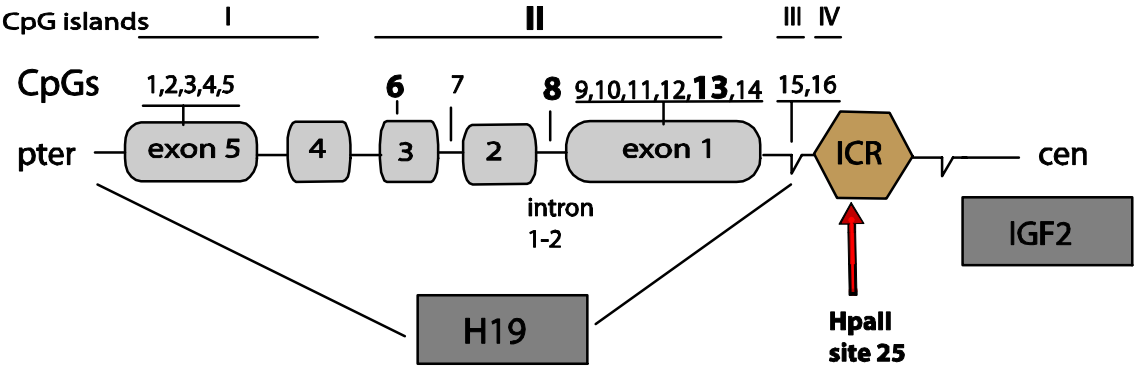

Figure 6. Schematic view of the *H19* genomic structure and locations of the HumanMethylation27 Bead Array CpG sites with corresponding CpG islands. Three CpG sites (6, 8 and 13) show statistically significant differences in methylation levels between MRKH patients and controls. The *H19* ICR1 studied by qRT-PCR is indicated by an arrow. IGF2 = Insulin-like growth factor 2 gene.

In summary, hypomethylation of *H19* ICR1 is not a cause of MA in the investigated sample set including 80 Finnish patients with MA. The previously reported association between SRS patients with reproductive duct malformations and *H19* ICR1 hypomethylation were observed only in two patients per study (Bliek *et al.* 2006, Bruce *et al.* 2009), and could therefore be correlated to a phenotype other than MA, or be coincidental. Growth restriction, a main feature of SRS, is also not included in the MA phenotype spectrum, and none of the MA patients have been reported to have other features suggestive of SRS.

Our results on the patients´ methylation state of several CpGs within *H19* compared to controls were not statistically significant after adjusting for FDR but are, however, indicative and interesting for further studies in larger patient cohorts or in other populations. In particular, studies involving other CTCF-sites within *H19* ICR1 and their methylation levels, or other imprinted genes, could be relevant to the disorder. This, in addition to previous findings of monozygotic twins discordant for MA (Lischke *et al.* 1973, Steinkampf *et al.* 2003, Duru *et al.* 2009), as well as healthy female children born to MA patients (Petrozza *et al.* 1997), encourage further epigenetic studies in MA for finding the underlying cause for at least some of the patients.

5.3 *SHOX* copy number variations in patients with Müllerian aplasia (Study III)

In this study our aim was to evaluate the role of *SHOX* CNVs in the Finnish cohort of MA patients. This was based on a previous report by Gervasini *et al.* suggesting that partial duplications of the *SHOX* gene are associated with MA (Gervasini *et al.* 2010).

DNA samples from 111 Finnish MA patients and 115 healthy controls were studied for CNVs in *SHOX* using the commercial MLPA assay (SALSA MLPA kit P018-E1 SHOX, MRC-Holland, Amsterdam, the Netherlands). *SHOX* is alternatively spliced into two main transcripts, *SHOXa* and *SHOXb*, with a non-coding first exon followed by six protein-coding exons, of which the last exon differs between the two transcripts (6a or 6b)(Leka *et al.* 2006). We used an MLPA assay that included 43 probes covering the first six exons of *SHOX*, its promoter region, a unique intron sequence before exon 6b and a region downstream of the gene. The S*HOX* probes for exons 1-3 and 5-6a were the same as in the MLPA assay used by Gervasini and coworkers (PO18-B), whereas the probes for exon 4 and 6b had been replaced by the manufacturer in the assay used in our study (P018-E1). All aberrant amplification results were confirmed in a second independent MLPA analysis.

All investigated MA samples within our study had normal amplification of *SHOX*, in other words two copies of the gene were detected in all patient and control samples. In five MA patient samples and seven control samples, CNVs downstream of *SHOX* were found (Figure 7; Table 8).

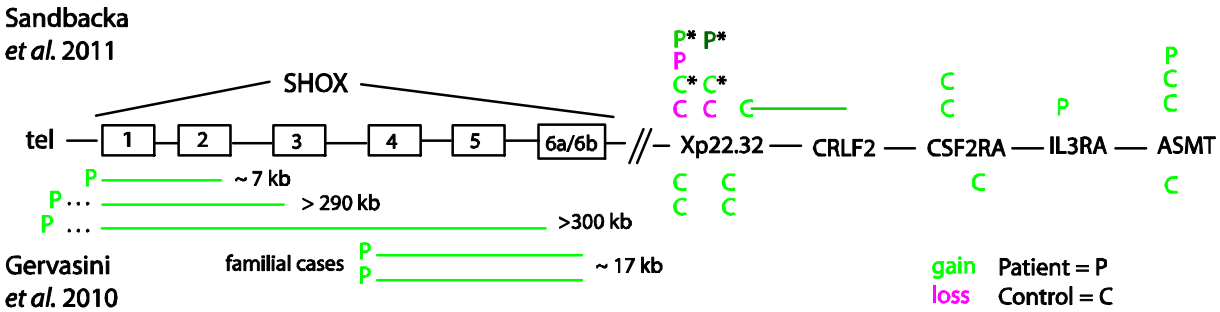

Figure 7. Comparison of CNV results within and downstream of *SHOX* detected in MA patients (P)and controls (C) using MLPA. CNVs reported by Sandbacka *et al.* (2011) are illustrated above the schematic view of the studied genomic region, whereas CNVs reported by Gervasini *et al.* (2010) are illustrated below. CNVs spanning more than two probes are indicated by lines, whereas CNVs detected by one probe are indicated with a letter. Gains are in green and losses in purple. Samples containing two CNVs are indicated by *. Tel = telomere.

All observed CNVs downstream of *SHOX* were previously reported in the Database of Genomic Variants (DGV; http://projects.tcag.ca/variation/) and are therefore most likely common polymorphisms. Furthermore, the 12 samples with CNVs detected downstream of *SHOX* contained one CNV each, except one patient sample and one control sample containing two aberrations. Another control sample had a larger duplication consisting of 11 probes (Table 8).

Table 8. CNVs (gains and losses) downstream of *SHOX* detected in Finnish MA patients and controls using MLPA assays (SALSA MLPA kit P018-E1 SHOX, MRC-Holland).

Region/gene	Aberration	MLPA probe	Cases	Controls	Reference
Xp22.32-PAR1	Gain	09335-L15508[a]	1*	1*	DGV[b], Gervasini *et al*. 2010
Xp22.32-PAR1	Loss	09335-L15508	1	1	DGV
Xp22.32-PAR1	Gain	14697-L16348	1*	1*	DGV
Xp22.32-PAR1	Loss	14697-L16348		1	DGV
Xp22.32-PAR1 to CRLF2	Gain	13296-L15336 - 13911-L16505		1	all 11 probes reported in DGV
CSF2RA	Gain	10251-L15502		2	DGV
IL3RA	Gain	13597-L15055	1		DGV
ASMT	Gain	01153-L00712[c]	2	1	DGV, Gervasini *et al*. 2010

[a] the probe sequence partly overlaps with MLPA probe 5650-L5104 from version PO18-B SHOX used by Gervasini *et al*. 2010.
[b] DGV (Database of Genomic Variants, http://projects.tcag.ca/variation)
*one patient and one control sample contain two aberrations each
[c] the probe corresponds to MLPA probe 1153-L0712 from version PO18-B SHOX used by Gervasini *et al*. 2010.

The *SHOX* gene is situated in the pseudoautosomal region of the X (Xp22) and Y (Yp11.3) chromosomes, and is therefore normally present in two copies in both sexes. Mutations and CNVs (duplications and deletions) in *SHOX* have previously been reported to contribute to several short stature disorders, such as idiopathic short stature (ISS, OMIM 300582), Turner syndrome (Ellison *et al.* 1997, Rao *et al.* 1997), dyschondrosteosis (Leri-Weill syndrome, LWD, OMIM 127300) and Langer mesomelic dysplasia (OMIM 249700)(Belin *et al.* 1998, Shears *et al.* 1998, Benito-Sanz *et al.* 2011). Previous reports of patients with duplications in *SHOX* are rare and include a total of only five patients, of whom none were reported with MA (Grigelioniene *et al.* 2001, Ogata *et al.* 2001, Roos *et al.* 2009, Thomas *et al.* 2009, D'haene *et al.* 2010, Benito-Sanz *et al.* 2011). For one of the patients the duplication was also found in her healthy sister, thereby questioning the functional effect of the CNV (Grigelioniene *et al.* 2001).

In the study by Gervasini *et al*. (2010) including 30 MA patients, the familial MA patients (3/3) and 7% of the sporadic patients (2/27) were reported to have partial duplications of *SHOX* corresponding to two or more exons of the gene and ranging in size from 7 to >300 kb (Figure 7). The entire *SHOX* gene was not duplicated in any of the samples (Gervasini *et al.* 2010). In the familial cases, the research group showed that both sisters with MA shared the same duplication as well as the same haplotype for the aberrant region as their healthy father, which was lacking from the two healthy sisters and their mother. This would correspond to a sex-limited dominant inheritance transmitted through a healthy father, suggested already in 1978 by Shokeir (Shokeir 1978).

In our study including a more than three times larger patient cohort (101 Finnish MA patients), of whom two were familial (sisters), the expected number of patients with *SHOX* aberrations if the frequencies were the same would be nine including both sisters and seven sporadic patients. However, none of the studied patients in our study were found to carry intragenic CNVs in *SHOX*. Our results do not confirm their findings, but on the contrary suggest that *SHOX* CNVs are not the cause of MA, at least not in the Finnish patient cohort, and are unlikely to be a major cause of MA worldwide. We chose to use the same experiment for detecting deletions and duplications in *SHOX*, the MLPA method, as the Gervasini group, in order to achieve comparable results. The intragenic *SHOX* probes for exon 4 and 6b had been replaced in the MLPA assay when compared to that in the Italian report, but since the intragenic *SHOX* duplications reported by Gervasini and coworkers all spanned more than one probe, corresponding CNVs would also have been detected in the Finnish cohort if they existed. However, in a complex trait population differences may exist and underlie the discordant results. Additionally, the positive findings by the Italian group might be coincidental without actual association to the MA phenotype.

5.4 *TBX6*, *LHX1* and copy number variations in Müllerian aplasia (Study IV)

The focus of this study was to find novel genes and CNVs associated with the development of MA. At the start of the study, mutations in only two genes, *WNT4* and *LHX1*, were known to cause MA and the number of patients reported with mutations (a total of six) was small (Biason-Lauber *et al.* 2004, Biason-Lauber *et al.* 2007, Philibert *et al.* 2008, Philibert *et al.* 2011, Ledig *et al.* 2011, Ledig *et al.* 2012).

Fifty MA patients, all diagnosed by laparoscopy, were investigated with aCGH (180 K platform, Oxford Gene Technology, Yarnton, Oxford, UK) revealing rare CNVs in 8/50 (16%) (Table 9). Seven of the CNVs were deletions found in one patient each and two were

duplications, found in the same patient. Five of the deletions (5p14.3, 9q21.13, 11q13.4, 15q26.1 and 16p13.3) and both duplications (19q13.11 and 19q13.12) were novel and not reported in DGV, whereas two deletions had been previously reported in association with MA. These were the 17q12 (Cheroki *et al.* 2008, Bernardini *et al.* 2009, Ledig *et al.* 2011, Nik-Zainal *et al.* 2011) and 16p11.2 deletions (Nik-Zainal *et al.* 2011) sized 1.7 Mb and 0.53 Mb, respectively.

Table 9. CNVs in a set of MA patients detected using aCGH.

Locus	Patient ID	Size	CNV	Genes within CNV	Confirmation method
5p14.3	28	1.6 Mb	Del	CDH18	SNP array, qPCR
9q21.13	3	95 Kb	Del	TMEM2	SNP array
11q13.4	2	54 Kb	Del	CHCHD8, PAAF1	SNP array, qPCR
15q26.1	4	96 Kb	Del	ZNF774, IQGAP1	SNP array, qPCR
16p11.2	69	0.53 Mb	Del	SPN, QPRT, C16orf54, MAZ, PRRT2, C16orf53, MVP, CDIPT, LOC440356, SEZ6L2, ASPHD1, KCTD13, TMEM219, TAOK2, HIRIP3, INO80E, DOC2A, C16orf92, FAM57B, ALDOA, PPP4C, TBX6, YPEL3, GDPD3, MAPK3, LOC100271831	SNP array
16p13.3	42	143 Kb	Del	A2BP1	SNP array, qPCR
17q12	24	1.7 Mb	Del	TBC1D3C, CCL3L1, CCL3L3, CCL4L2, CCL4L1, TBC1D3H, TBC1D3C, TBC1D3G, ZNHIT3, MYO19, PIGW, GGNBP2, DHRS11, MRM1, LHX1, AATF, ACACA, C17orf78, TADA2L, DUSP14, AP1GBP1, DDX52, HNF1B, LOC284100	SNP array
19q13.11[a]	49	194 Kb	Dupl	RHPN2, GPATCH1, WDR88, LRP3, SLC7A10	SNP array, qPCR
19q13.12[a]	49	0.6 Mb	Dupl	LSR, USF2, HAMP, MAG, CD22, FFAR1, FFAR3, FFAR2, KRTDAP, DMKN, SBSN, GAPDHS, TMEM147, ATP4A, HAUS5, RBM42, ETV2, COX6B1, UPK1A, ZBTB32, MLL4, TMEM149, U2AF1L4, PSENEN, LIN37, HSPB6, C19orf55, SNX26, PRODH2	SNP array, qPCR

[a] likely to include non-duplicated regions, possibly due to other chromosomal rearrangements e.g. inversions within the region or non-functional aCGH probes.
CNV = copy number variations, Del = deletion, Dupl = duplication.

The 17q12 region embraces 24 genes including *LHX1*, for which mutations in MA were reported (Ledig *et al.* 2011, Ledig *et al.* 2012), and *HNF1B*, which has been studied in MA but thus far no mutations have been found (Bernardini *et al.* 2009, Ledig *et al.* 2011). 17q12 deletions of the same size have previously been associated with renal malformations and cysts, growth restriction, speech problems (Mefford *et al.* 2007, Nagamani *et al.* 2010) and autism (Loirat *et al.* 2010), whereas duplications have been associated with cognitive impairment, epilepsy, and renal and urinary malformations (Mefford *et al.* 2007, Nagamani *et al.* 2010). Deletions in 16p11.2 have previously been documented in association with autism (Kumar *et al.* 2008, Weiss *et al.* 2008), developmental delay (Rosenfeld *et al.* 2010) and obesity (Walters *et al.* 2010), whereas duplications in 16p11.2 have been reported in association with autism (Weiss *et al.* 2008), developmental delay (Rosenfeld *et al.* 2010), and schizophrenia (McCarthy *et al.* 2009). 16p11.2 includes at least 26 genes, of which *TBX6* is the only gene with known function in embryonic development (White *et al.* 2003). We therefore continued to investigate *TBX6* and *LHX1* as candidate genes for MA.

MLPA and Sanger sequencing of *TBX6* was performed in the patient cohort that at this point was extended to embrace 112 MA patients (110 Finnish and two foreign). MLPA revealed monoallelic deletion of *TBX6* in 5/112 patients, of which one was the patient with the 0.53 Mb deletion on 16p11.2 originally detected by aCGH. Validation studies by SNP genotyping (HumanOmni2.5-8 BeadChip v1.0, Illumina, San Diego, CA, USA) confirmed that also the other four patients had the same 16p11.2 deletion including *TBX6*.

Sanger sequencing of *TBX6* revealed a novel splice site variant (c.622-2A>T; g.30100162T>A) in 2/112 MA patients (1.8%) (Table 10). This heterozygous A>C change is located in the highly conserved splice acceptor site (AG) of exon 5. The variant was not observed when sequencing 200 healthy female control samples, not reported in dbSNP, the 1000 Genomes database, in 1532 alleles of Finnish ancestry (A-P Sarin and A. Palotie, personal communication, 2013) or in the Exome Variant Server (EVS), where *TBX6* sequence data is available for >12000 alleles of African American and European American ancestry. The splice site variant is located within the evolutionally highly conserved DNA-binding T-box element of the gene (Figure 8). *In silico* prediction programs indicated that the mutation would disrupt the correct splicing of the transcript. RT-PCR revealed at least three different transcripts present in both patients and healthy controls.

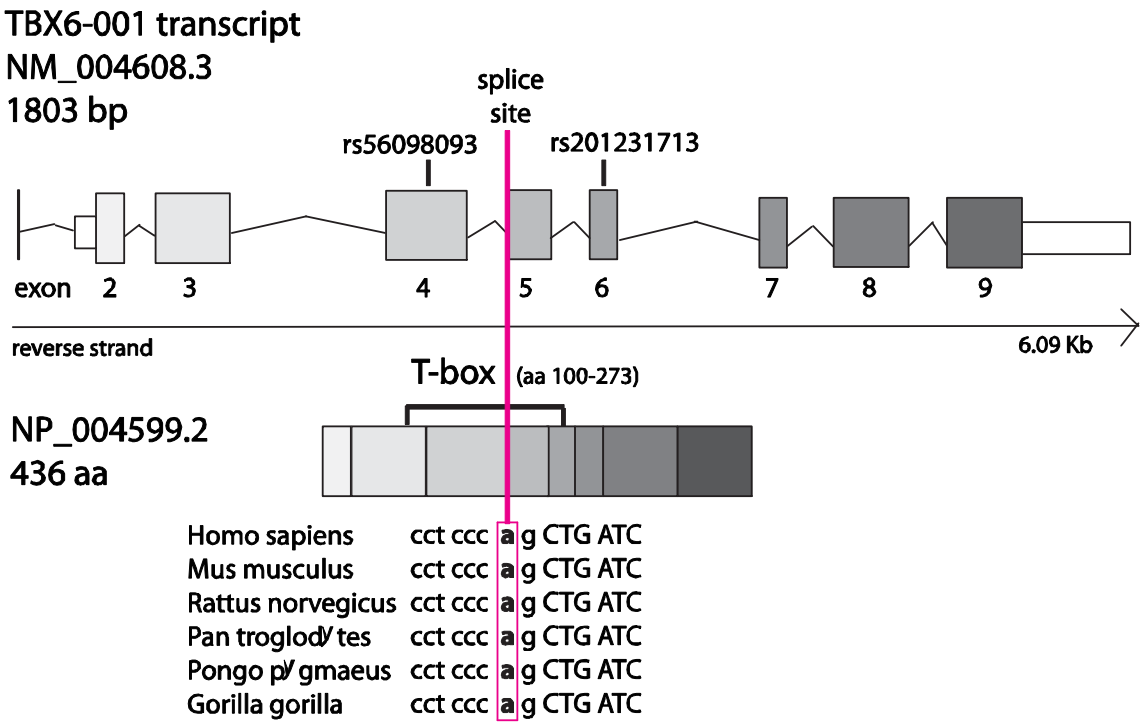

Figure 8. A schematic view of *TBX6* with arrows indicating the c.622-2A>T mutation, and the rare missense variants in exon 4 (g.30100401C>T; p.Gly162Ser; rs56098093) and exon 6 (g.30099890C>T; p.Arg272Gln; rs201231713) of the gene (Sandbacka *et al.* 2013).

By sequencing we also found six *TBX6* variants, all previously reported in dbSNP (Table 10). Interestingly, two of the variants were rare (exon 4 rs56098093 and 6 rs201231713) and statistically significantly more common in patients than in controls. The MAF for rs56098093 was 8.0% for patients compared to 2.0% for controls and for rs201231713 5.8% for patients compared to 2.0% for controls. The differences were statistically significant (*P*-value 0.0021 (rs56098093) and 0.0002 (rs201231713), also after correction for multiple testing according to Bonferroni (*P*-value/2 < 0.05), indicating that the variants might increase the risk for the disorder. Noteworthy is that two patients were homozygous for both of the variants, whereas none of the 200 sequenced controls were homozygous for either of them.

Table 10. A summary of *TBX6* and *LHX1* variants found by Sanger sequencing 112 MA patients (110 Finnish and 2 foreign). The novel and rare variants are shown in bold.

Gene and location within gene	**Variant[a] Predicted change**	**rs number[b]**	**Genotype of patients[d]** N=112	**Genotype of controls[d]** N=200 (*TBX6*) N=180 (*LHX1*)
TBX6, intron 2	g.30102391G>A c.118+6C>T	rs112565029	CC: 107 (95.5%) CT: 5 (4.5%) TT: 0 (0%)	CC: 187 (93.5%) CT: 13 (6.5%) TT: 0 (0%)
TBX6, exon 4	g.30100402G>A p.Ser161=	rs147485102	CC: 111 (99.1%) CT: 1 (0.9%) TT: 0 (0%)	CC: 198 (99%) CT: 2 (1%) TT: 0 (0%)
***TBX6,* exon 4**	**g.30100401C>T p.Gly162Ser**	**rs56098093**	**GG: 97 (86.6%) GA: 12 (10.7%) AA: 3 (2.7%)**	**GG: 192 (96%) GA: 8 (4%) AA: 0 (0%)**
***TBX6,* intron 4**	**g.30100162T>A c.622-2A>T**	**-**	**AA: 11 (98.2%) AT: 2 (1.8%) TT: 0 (0%)**	**AA: 200 (100%) AT: 0 (0%) TT: 0 (0%)**
***TBX6,* exon 6**	**g.30099890C>T p.Arg272Gln**	**rs201231713**	**GG: 101 (90.2%) GA: 9 (8.0%) AA: 2 (1.8 %)**	**GG: 198 (99%) GA: 2 (1%) AA: 0 (0 %)**
TBX6, intron 7	g.30098022G>A c.914-6C>T	rs200310768	CC: 109 (97.3%) CT: 3 (2.7%) TT: 0 (0%)	CC: 194 (97%) CT: 6 (3%) TT: 0 (0%)
TBX6, exon 9	g.30097630C>T p.Pro409=	rs2289292	GG: 32 (28.6%) GA: 55 (49.1%) AA: 25 (22.3%)	GG: 67 (33.5%) GA: 95 (47.5%) AA: 38 (19%)
***LHX1,* exon 1**	**g.35295505G>C p.Cys4Ser**	**-**	**GG: 111 (99.1%) GC: 1 (0.9%) CC: 0 (0%)**	**GG: 180 (100%) GC: 0 (0%) CC: 0 (0%)**
***LHX1,* exon 5**	**g.35300142C>A p.Pro312His**	**-**	**CC: 109 (97.3%) CA: 3 (2.7%) AA: 0 (0%)**	**CC: 180 (100%) CA: 0 (0%) AA: 0 (0%)**
***LHX1,* exon 5**	**g.35300202C>G p.Pro332Arg**	**TMP_ESP_17_35300202[c]**	**CC: 111 (99.1%) CG: 1 (0.9%) GG: 0 (0%)**	**CC: 180 (100%) CG: 0 (0%) GG: 0 (0%)**

[a]variations presented according to genomic reference sequence for *TBX6*: (g.) NC_000016.9, coding DNA reference sequence (c.) NM_004608.3 and protein reference sequence (p.) NP_004599.2, (Genome Build 37.3, dbSNP), and for *LHX1*: (g.) NC_000017.10, (c.) NM_005568.3 and (p.) NP_005559.2 (Genome Build 37.3, dbSNP).

[b]rs numbers for previously known SNPs according to dbSNP or Exome Variant Server (EVS).

[c]1/12783 alleles reported in EVS.

[d]number of genotypes reported as: reference allele/reference allele ; reference allele/alternative allele ; alternative allele/alternative allele.

Mouse knockout studies have shown *Tbx6* to be important for somite segmentation in the paraxial mesoderm during embryogenesis (Chapman *et al.* 1998). A link between the Wnt signaling pathway and *Tbx6* is suggested through interactions with Notch-ligand delta-like 1 (*Dll1*) (White *et al.* 2003). Moreover, a spontaneous mouse model for *Tbx6*, the homozygous Tbx6^{rv} (rib-vertebrae) has in addition to segmentation problems skeletal and urogenital

malformations (Theiler *et al.* 1985, Watabe-Rudolph *et al.* 2002), thereby resembling the MA phenotype. In humans, a Macedonian family has been reported with a missense mutation in the last codon of the *TBX6* transcript resulting in segmentation defects of the vertebrae (Sparrow *et al.* 2013). *TBX6* has also been associated with congenital scoliosis in the Chinese Han population (Fei *et al.* 2010).

Sequencing of *LHX1* in the patient cohort revealed three novel missense variants in 5/112 (4.4%) MA patients (Table 10; Figure 9). None of the variants are reported in dbSNP, and only the p.Pro332Arg was previously reported in EVS, in 1/12783 alleles. *In silico* predictions indicated the exon 1 variant as deleterious, while the data for the other two were discrepant. The patient with the exon 1 mutation was from the Middle East and the four other patients Finnish.

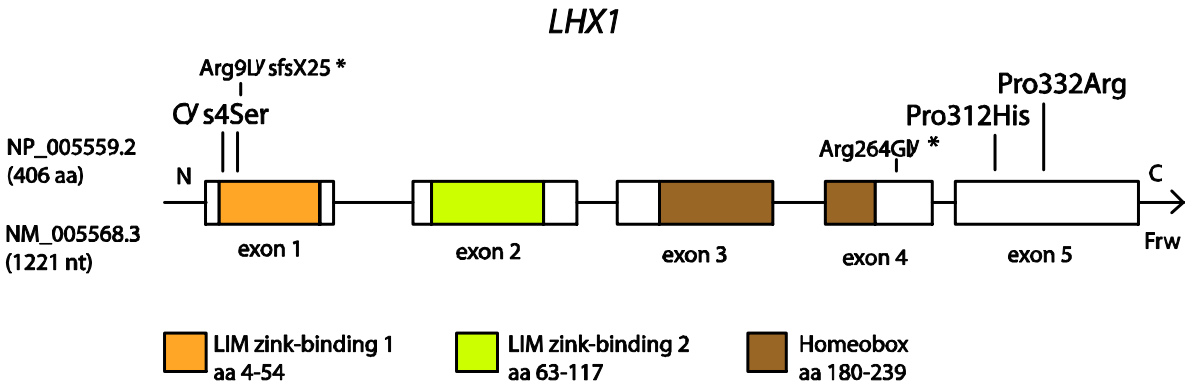

Figure 9: A schematic view of *LHX1* with the novel DNA variants (resulting in Cys4Ser, Pro312His and Pro332Arg) shown at corresponding locations at nucleotide and protein levels. The two previously known *LHX1* variants (Arg9LysfsX25 and Arg264Gly; Ledig *et al.* 2011 and 2012) are indicated by *. Aa corresponds to amino acid.

Taken together, we were able to identify *TBX6* as a new gene associated with the development of MA. Our findings strengthening this statement included a deletion of 16p11.2, embracing *TBX6*, found in 5/112 MA patients, a novel *TBX6* splice site mutation found in 2/112 patients, and two rare *TBX6* variants in exon 4 and exon 6, both significantly more common in patients (8% and 5%, respectively) than in controls (2%). We also found three novel *LHX1* variants in 5/112 patients, strengthening the relevance of the gene in MA development. Additionally, we found seven rare CNVs not previously associated with MA, denoting the genes located within the aberrant regions as important candidate genes in further MA studies. Altogether, we identified either CNVs or *TBX6* or *LHX1* variants in 30/112 (26.8%) MA patients. CNVs were found in 12/112 (10.7%), patients, novel *TBX6* or *LHX1* variants in 7/112 (6.3%) patients, and rare *TBX6* variants in 15/112 (13.4%) patients. Importantly, four patients (4/112, 3.6%) were carrying variants in both *TBX6* and *LHX1* or a *TBX6* variant in combination with a CNV, supporting the multifactorial etiology of MA.

6. General discussion, concluding remarks and future prospects

The aim of this study was to investigate the genetic background and identify novel genetic causes of MA. This syndrome is rare, however affecting at least 1 in 5000 female births. In the Finnish population with a size of 5.4 million and approximately 60 000 births per year, the estimated number of new cases is at least six (0.0002 x 30 000) per year. With an average life expectancy of 83 years for women in Finland, this means that at least 500 females have MA in our country today. The MA frequency is estimated to be approximately the same world-wide.

MA is regarded as one of the difficult disorders of female health. It is featured by congenital loss of functional uterus and vagina and is commonly diagnosed at the age of 14-16 due to primary amenorrhea (absent menses). The vagina is undeveloped and exists only as a small pouch and therefore treatment is needed for commencing a normal sex life. The lack of uterus leads to infertility, even if MA patients mostly have normal oocyte-producing ovaries. Legal and ethical acceptance for surrogacy could provide MA patients with biological children. Uterine transplantation could offer another possibility for MA women to become pregnant in the future, but is still in an initial phase and necessitates much research and ethical discussion in order to become a real-life alternative. Therefore, adoption is presently the only possibility for parenthood among MA women in Finland. The diagnosis of MA often leads to psychosocial problems. Many patients will not share the diagnosis even with their family and closest friends due to the sensitivity of the syndrome.

Previously, only mutations in two genes, namely *WNT4* and *LHX1*, were linked to MA and only in a handful of patients world-wide. In addition to these mutations, CNVs were suggested as causative for the disorder, but again in a small number of patients. Therefore, the cause of the disorder is still unknown for the majority of patients and the genetics of MA is considered multifactorial with mutations in several genes or genomic rearrangements with incomplete penetrance and variable expressivity playing part in its etiology. However, for a subgroup of patients or for specific populations, a monogenic form of MA might exist.

This study investigated the role of Y chromosomal fragments, CNVs as well as DNA methylation defects to underlie MA in a large Finnish cohort of MA patients, the second largest MA cohort world-wide published to date. We have also performed collaborative studies with researchers both in Sweden and Brazil in order to improve our methodological knowledge and to test our study hypothesis. The results show that Y chromosomal factors and *SHOX* CNVs do not have a role in the etiology of the disorder in Finnish patients. However, as MA is a complex multifactorial disorder, population-differences may occur. Our

studies on the DNA methylation status of the imprinted gene *H19* and its control region ICR1 indicate that DNA methylation and epigenetic factors in general can have a role for a subgroup of MA patients. However, further studies in this field are needed.

This study confirmed the previous findings that CNVs in 16p11.2 and 17q12 are associated with MA. Furthermore, we identified *TBX6* as a new gene with mutations in MA patients as we found two Finnish patients to carry splice mutations. In addition to the mutation, we also found two rare variants to be more frequent in patients than in healthy controls, and we therefore suggest that they are important in MA etiology. We were also able to reveal three novel mutations in five patients in the *LHX1* gene, recently found to be linked to MA, as well as rare CNVs, a basis for future candidate gene studies. To summarize, within the frameworks of this study, 6% (7/112) patients were found with mutations most likely causative of the disorder and 24% (27/112) with suggestive causes (rare variants and CNVs) adding up to 30% (34/112) of the patients with a genetic cause that could underlie MA. We were also able to show that four MA patients (3.4%) carried variations in both *TBX6* and *LHX1* or in *TBX6* in combination with CNVs exemplifying the complex genetics of MA.

Limitations in this study include the problem with investigating a tissue that is missing - expression studies based on tissue samples are not possible. Additionally, the time point of MD development (or regression), which occurs in the early weeks of human embryonic development, restrict the research possibilities. Functional studies based on our *TBX6* and *LHX1* mutation findings to prove their role in the development of MA would be important to perform. The private nature of MA also challenges the possibilities of collecting samples from families, which would be valuable for haplotype and inheritance screening. Genome-wide association studies in order to reveal new candidate loci have been performed in several complex disorders such as diabetes and cancer. However, these studies require thousands or tens of thousands of samples in order to give reliable results, which is not possible in rare and sensitive disorders like MA. Increased knowledge of MA and other reproductive disorders in the general population would be helpful for both the patients and their families as well as for the research within the field.

In this study several negative findings were encountered. Our study designs and methods of choice were selected based on high efficiency, sensitivity and reliability and therefore we believe that these are true findings. Our negative findings can be due to previous studies being performed in small patient cohorts with different population backgrounds or that the original findings are not relevant for MA. Establishment of an international MA consortium with collaboration and data sharing between research groups world-wide could be a good platform for future discoveries in the field.

Future goals within the MA project are to further explore the candidate genes within the newly described CNVs and also add them to the previously associated CNVs in order to find gene families or pathways that might be involved in the development of MA. Further, combining genome-wide SNP genotype with DNA methylation data could give clues to new genes or genetic regions with shared haplotypes in conjunction with aberrant DNA methylation patterns for future candidate gene studies. Mutation screening by exome sequencing is presently being pursued in the project. Samples from healthy family members would be valuable to gather and include, in order to exclude variants without biological relevance to the disorder. Finding out more about the underlying causes of MA, could in the future also enable genetic counseling for the patients.

In this study, we are able to give an answer for the probable genetic cause of MA to almost one third of the Finnish MA patients, which is a significant number compared to what was previously known. Every new discovery is a step forward in MA research and even if there is no treatment today for the infertility caused by MA, gradually deciphering the genetic causes of MA might help the patients and families in their everyday life.

7. Acknowledgements

This study was carried out at the Folkhälsan Institute of Genetics, the Department of Medical Genetics, and the Research Programs Unit, Genome-Scale Biology, University of Helsinki during 2006-2013. I wish to thank the former and current heads of these institutes, Professors Per-Henrik Groop, Anna-Elina Lehesjoki, Päivi Peltomäki, Irma Järvelä, Olli Jänne and Sampsa Hautaniemi for providing excellent working facilities in Biomedicum Helsinki.

I am grateful to the following organizations, foundations and funds for financially supporting my work: Victoriastiftelsen, Samfundet Folkhälsan, Päivikki and Sakari Sohlberg Foundation, Medicinska Understödsföreningen Liv och Hälsa, Finska Läkaresällskapet, Paulo Foundation, Waldemar von Frenckell Foundation, Svenska Kulturfonden, Njurstiftelsen, University of Helsinki Funds and the Chancellor travel fund of the University of Helsinki.

I cordially thank the patients and the family members that participated in the study and thereby made it possible.

I wish to thank the pre-examiners, docents Leila Unkila-Kallio and Miina Ollikainen, for their thorough examination of my thesis and valuable comments and support. Docent Leila Unkila-Kallio is also thanked for being my Ph.D. follow-up group member together with docent Sirpa Kivirikko. Your enthusiasm and positive attitude towards my work has been encouraging, and our yearly meetings have been a source of inspiration.

I wish to express my gratitude to my supervisors, docents Kristiina Aittomäki and Hannele Laivuori for introducing me to the interesting field of reproductive genetics. I am grateful to Kristiina for sharing her vast expertise in genetics and medicine and to Hannele for her knowledge, everlasting positive attitude and encouragement over the years.

I warmly acknowledge the following clinical collaborators for building up the unique sample series I had the opportunity to work with: Pauli Kajanoja for initiating the Finnish Müllerian aplasia project in the mid-1970s, Hille Eroila for data and sample collection in the early days of the project, Mervi Halttunen for active enrollment of patients at HUCH since the late 1990s and her encouraging support and quick response to any of my queries, Varpu Jokimaa and Laure Morin-Papunen for joining the project in 2011. Minna Puhakka is warmly thanked for putting the clinical data into a comprehensive format for me to work with and for her genealogy efforts in the project.

I am indebted to Ph.D. Jodie Painter, who was working in the group when I started. Your help and encouragement during the first study of my thesis made me believe in my

capabilities of continuing with the project even without a senior person by my side in the lab. Jodie is also warmly thanked for her friendship and for revising the language of my thesis. My fellow Ph.D. student Milja Kaare is warmly thanked for sharing the ups and downs with me in the project and for helping with sample collection when I was on maternity leave. I have missed both of you ever since you left the group. Anne Polvi is warmly thanked for her work with the exome data and for sharing her vast knowledge in the field. Hanna Nurmi, Susanna Mehtälä, Katariina Hirvonen and Viivi Nevanlinna are thanked for helping me out at times when extra hands have been needed in the lab.

I owe a special thanks to docent Carina Wallgren-Pettersson for welcoming me with open arms into her group and the field of neuromuscular disorders when I graduated as a fresh M.Sc. from Åbo Akademi University in 2001. I am grateful for your continuous enthusiasm towards my work, even if the focus changed during the years from muscles into reproductive organs.

Several collaborators have contributed to the study and I am deeply grateful to Harriet von Koskull for sharing information on Y-chromosome screening, Professor Juha Kere, Sara Bruce and Ingegerd Fransson for inviting me to Karolinska Institutet, Sweden, to learn and perform qPCR, Carla Rosenberg and Érika Freitas at the São Paulo University, Brazil, for their aCGH expertise, Professor Sampsa Hautaniemi, Markku Niemi and Sirkku Karinen for bioinformatics expertise, Professor Aarno Palotie and Antti-Pekka Sarin for sharing genotype data, Professor Päivi Peltomäki, Annette Gylling, Sippy Kaur, Saila Saarinen, Anni Niskakoski and Satu Valo for sharing their MLPA expertise and Sippy Kaur together with Suvi Savola for teaching me aCGH many years ago in Professor Sakari Knuutila´s lab. The staff of the Finnish Institute of Molecular Medicine (FIMM), in particular Päivi Lahermo, Mari Kaunisto, Pekka Ellonen and Kati Donner are warmly thanked for assistance in several genome-wide studies.

I am thankful to the research nurses Sinikka Lindh, Päivi Ihamuotila and Tuula Soppela for helping with blood sample collections whenever needed. Sinikka is also warmly thanked for her kind help in genealogy and logistical matters. Katja Kuosa is acknowledged for helping with sample logistics and Aila Riikonen, Jaana Welin-Haapamäki, Stephan Keskinen, Nina Forss, Marjatta Valkama and Madeleine Avellan for helping with practical and financial matters during the years.

I cordially thank each and every one of my present and former colleagues at the Folkhälsan Research Center for your kind advice, methodological expertise, warm friendship and joyful company both at work and during numerous get-togethers and conference trips throughout the years. Without your help I would have been extremely lonely and lost, both in a scientific and a mental manner. My special thanks go to Kati Donner and Ann-Liz Träskelin

for your unconditional friendship, support and help in ANY kind of matters; my former and present roommates Nina Aula, Anne Juvonen, Eija Siintola, Anni Laari, Inken Körber and Otto Manninen for the relaxed atmosphere and everyday-life discussions; Kirsi Kiiski for MLPA trouble-shooting; Vilma-Lotta Lehtokari for language check-ups; Anna Naukkarinen, Minttu Marttila, Mervi Kuronen, Kaisa Kettunen and Päivi Söderholm for sharing happiness and sorrow; pH Jonson for your Norwegian humor and superior kindness, and to my Ph.D. "soul mates" Saara Tegelberg and Jaakko Sarparanta for sharing ups and downs and practical hints along the thorny doctoral path – best of luck to both of you in your forthcoming defences! Outside of Folkhälsan, Annukka Lahtinen is thanked for her nice travel company and Sequencher-backup.

Jag vill också tacka alla vänner när och fjärran för det stöd och den glädje ni gett mig. Speciellt vill jag tacka min barndomsvän Karin för våra pratstunder om allt mellan himmel och jord, Johanna och Ela för våra årliga sammanstrålningar, Nancy & Oje, Michaela & Roger, Eva & Roger, Lotta & Kjell och Patte för alla minnesvärda stunder vi delat. Er vänskap betyder mycket för mig och har varit en härlig motvikt till forskningen.

Biomaffian Britt, Carina, Sofia, Soffi och Totta tackas för den oförglömliga Biolog-tiden. Trots att avstånden mellan oss nuförtiden är betydligt längre än de var i Åbo, har vår vänskap bestått. Jag är tacksam för att ni agerat som bollplank när frustrationerna varit stora, och samtidigt glatts med mig i stunder då solen glimtat fram.

Sist men inte minst vill jag varmt tacka mina föräldrar Ea och Gösta för att ni alltid stöttat mig i de val jag gjort i livet och mina systrar Carina och Birgit med familjer för att ni alltid finns där i alla situationer. Mina svärföräldrar Anneli och Björn tackas varmt för sitt engagemang och för att ni alltid är redo att ställa upp för barn och barnbarn. Mina svägerskor Anna-Karin och Maria med familjer tackas för er stora hjälpsamhet och varma omtanke. Kära Anders, det STÖRSTA tacket sparar jag till dig, eftersom du har stöttat mig och stått ut med mitt tidvis varierande humör, mina tänjbara arbetstimmar och pendlandets frustrationer under hela detta avhandlingsarbete. Ellen och Axel, ni är de ljuvligaste varelser på jorden och tack vare er påminns jag varje dag om vad som är det viktigaste i livet.

Helsingfors, 1.10.2013

Maria Sandholm

8. References

Aittomäki K, Eroila H, Kajanoja P. (2001). A population-based study of the incidence of Müllerian aplasia in Finland. *Fertil Steril.* 76:624-625.

Al-Awadi SA, Teebi AS, Farag TI, Naguib KM, el-Khalifa MY. (1985). Profound limb deficiency, thoracic dystrophy, unusual facies, and normal intelligence: a new syndrome. *J.Med.Genet.* 22:36-38.

Alvarez-Nava F, Soto M, Fernandez E, Zenteno JC. (2004). Nested polymerase chain reaction--not ready for gold standard status yet! *Fertil.Steril.* 82:1471; author reply 1471-2.

Arango NA, Szotek PP, Manganaro TF, Oliva E, Donahoe PK, Teixeira J. (2005). Conditional deletion of beta-catenin in the mesenchyme of the developing mouse uterus results in a switch to adipogenesis in the myometrium. *Dev.Biol.* 288:276-283.

Bamshad MJ, Shendure JA, Valle D, Hamosh A, Lupski JR, Gibbs RA, Boerwinkle E, Lifton RP, Gerstein M, Gunel M, Mane S, Nickerson DA, Centers for Mendelian Genomics. (2012). The Centers for Mendelian Genomics: a new large-scale initiative to identify the genes underlying rare Mendelian conditions. *Am.J.Med.Genet.A.* 158A:1523-1525.

Barnes JD, Crosby JL, Jones CM, Wright CV, Hogan BL. (1994). Embryonic expression of Lim-1, the mouse homolog of Xenopus Xlim-1, suggests a role in lateral mesoderm differentiation and neurogenesis. *Dev.Biol.* 161:168-178.

Behringer RR, Finegold MJ, Cate RL. (1994). Mullerian-inhibiting substance function during mammalian sexual development. *Cell.* 79:415-425.

Belin V, Cusin V, Viot G, Girlich D, Toutain A, Moncla A, Vekemans M, Le Merrer M, Munnich A, Cormier-Daire V. (1998). SHOX mutations in dyschondrosteosis (Leri-Weill syndrome). *Nat.Genet.* 19:67-69.

Bendavid C, Pasquier L, Watrin T, Morcel K, Lucas J, Gicquel I, Dubourg C, Henry C, David V, Odent S, Leveque J, Pellerin I, Guerrier D. (2007). Phenotypic variability of a 4q34-->qter inherited deletion: MRKH syndrome in the daughter, cardiac defect and Fallopian tube cancer in the mother. *Eur.J.Med.Genet.* 50:66-72.

Benito-Sanz S, Barroso E, Heine-Suñer D, Hisado-Oliva A, Romanelli V, Rosell J, Aragones A, Caimari M, Argente J, Ross JL, Zinn AR, Gracia R, Lapunzina P, Campos-Barros A, Heath KE. (2011). Clinical and molecular evaluation of SHOX/PAR1 duplications in Leri-Weill dyschondrosteosis (LWD) and idiopathic short stature (ISS). *J.Clin.Endocrinol.Metab.* 96:E404-12.

Ben-Shachar S, Ou Z, Shaw CA, Belmont JW, Patel MS, Hummel M, Amato S, Tartaglia N, Berg J, Sutton VR, Lalani SR, Chinault AC, Cheung SW, Lupski JR, Patel A. (2008). 22q11.2 distal deletion: a recurrent genomic disorder distinct from DiGeorge syndrome and velocardiofacial syndrome. *Am.J.Hum.Genet.* 82:214-221.

Benson GV, Lim H, Paria BC, Satokata I, Dey SK, Maas RL. (1996). Mechanisms of reduced fertility in Hoxa-10 mutant mice: uterine homeosis and loss of maternal Hoxa-10 expression. *Development.* 122:2687-2696.

Bentzen JG, Forman JL, Johannsen TH, Pinborg A, Larsen EC, Andersen AN. (2013). Ovarian antral follicle subclasses and anti-mullerian hormone during normal reproductive aging. *J.Clin.Endocrinol.Metab.* 98:1602-1611.

Bernardini L, Gimelli S, Gervasini C, Carella M, Baban A, Frontino G, Barbano G, Divizia MT, Fedele L, Novelli A, Bena F, Lalatta F, Miozzo M, Dallapiccola B. (2009). Recurrent microdeletion at 17q12 as a cause of Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome: two case reports. *Orphanet J.Rare Dis.* 4:25.

Beski S, Gorgy A, Venkat G, Craft IL, Edmonds K. (2000). Gestational surrogacy: a feasible option for patients with Rokitansky syndrome. *Hum.Reprod.* 15:2326-2328.

Biason-Lauber A, Konrad D, Navratil F, Schoenle EJ. (2004). A WNT4 mutation associated with Müllerian-duct regression and virilization in a 46,XX woman. *N.Engl.J.Med.* 351:792-798.

Biason-Lauber A, De Filippo G, Konrad D, Scarano G, Nazzaro A, Schoenle EJ. (2007). WNT4 deficiency--a clinical phenotype distinct from the classic Mayer-Rokitansky-Kuster-Hauser syndrome: a case report. *Hum.Reprod.* 22:224-229.

Bingham C, Ellard S, Cole TR, Jones KE, Allen LI, Goodship JA, Goodship TH, Bakalinova-Pugh D, Russell GI, Woolf AS, Nicholls AJ, Hattersley AT. (2002). Solitary functioning kidney and diverse genital tract malformations associated with hepatocyte nuclear factor-1beta mutations. *Kidney Int.* 61:1243-1251.

Bliek J, Terhal P, van den Bogaard MJ, Maas S, Hamel B, Salieb-Beugelaar G, Simon M, Letteboer T, van der Smagt J, Kroes H, Mannens M. (2006). Hypomethylation of the H19 gene causes not only Silver-Russell syndrome (SRS) but also isolated asymmetry or an SRS-like phenotype. *Am.J.Hum.Genet.* 78:604-614.

Bordini B, Rosenfield RL. (2011). Normal pubertal development: Part I: The endocrine basis of puberty. *Pediatr.Rev.* 32:223-229.

Bozzi F, Bertuzzi S, Strina D, Giannetto C, Vezzoni P, Villa A. (1996). The exon-intron structure of human LHX1 gene. *Biochem.Biophys.Res.Commun.* 229:494-497.

Brannstrom M (2013). Uterus transplantation. *Fertil.Steril.* 99:348-349.

Bruce S, Hannula-Jouppi K, Lindgren CM, Lipsanen-Nyman M , Kere J. (2008). Restriction site-specific methylation studies of imprinted genes with quantitative real-time PCR. *Clin.Chem.* 54:491-499.

Bruce S, Hannula-Jouppi K, Peltonen J, Kere J, Lipsanen-Nyman M. (2009). Clinically distinct epigenetic subgroups in Silver-Russell syndrome: the degree of H19 hypomethylation associates with phenotype severity and genital and skeletal anomalies. *J.Clin.Endocrinol.Metab.* 94:579-587.

Bulletti C, Palagiano A, Pace C, Cerni A, Borini A, de Ziegler D. (2011). The artificial womb. *Ann.N.Y.Acad.Sci.* 1221:124-128.

Burel A, Mouchel T, Odent S, Tiker F, Knebelmann B, Pellerin I, Guerrier D. (2006). Role of HOXA7 to HOXA13 and PBX1 genes in various forms of MRKH syndrome (congenital absence of uterus and vagina). *J.Negat Results Biomed.* 5:4.

Carroll TJ, Park JS, Hayashi S, Majumdar A, McMahon AP. (2005). Wnt9b plays a central role in the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system. *Dev.Cell.* 9:283-292.

Carson SA, Simpson JL, Malinak LR, Elias S, Gerbie AB, Buttram VC Jr, Sarto GE. (1983). Heritable aspects of uterine anomalies. II. Genetic analysis of Mullerian aplasia. *Fertil.Steril.* 40:86-90.

Chapman DL, Papaioannou VE. (1998). Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature.* 391:695-697.

Chen ZX, Riggs AD. (2011). DNA methylation and demethylation in mammals. *J.Biol.Chem.* 286:18347-18353.

Cheroki C, Krepischi-Santos AC, Rosenberg C, Jehee FS, Mingroni-Netto RC, Pavanello Filho I, Zanforlin Filho S, Kim CA, Bagnoli VR, Mendonca BB, Szuhai K, Otto PA. (2006). Report of a del22q11 in a patient with Mayer-Rokitansky-Küster-Hauser (MRKH) anomaly and exclusion of WNT-4, RAR-gamma, and RXR-alpha as major genes determining MRKH anomaly in a study of 25 affected women. *Am.J.Med.Genet.A.* 140:1339-1342.

Cheroki C, Krepischi-Santos AC, Szuhai K, Brenner V, Kim CA, Otto PA, Rosenberg C. (2008). Genomic imbalances associated with mullerian aplasia. *J.Med.Genet.* 45:228-232.

Choufani S, Shuman C, Weksberg R. (2010). Beckwith-Wiedemann syndrome. *Am.J.Med.Genet.C.Semin.Med.Genet.* 154C:343-354.

Clement-Ziza M, Khen N, Gonzales J, Cretolle-Vastel C, Picard JY, Tullio-Pelet A, Besmond C, Munnich A, Lyonnet S, Nihoul-Fekete C. (2005). Exclusion of WNT4 as a major gene in Rokitansky-Kuster-Hauser anomaly. *Am.J.Med.Genet.A.* 137:98-99.

Coffinier C, Barra J, Babinet C, Yaniv M. (1999). Expression of the vHNF1/HNF1beta homeoprotein gene during mouse organogenesis. *Mech.Dev.* 89:211-213.

Committee on Adolescent Health Care (2013). Committee opinion: no. 562: mullerian agenesis: diagnosis, management, and treatment. *Obstet.Gynecol.* 121:1134-1137.

Couse JF, Korach KS. (1999). Estrogen receptor null mice: what have we learned and where will they lead us? *Endocr.Rev.* 20:358-417.

Cramer DW, Goldstein DP, Fraer C, Reichardt JK. (1996). Vaginal agenesis (Mayer-Rokitansky-Kuster-Hauser syndrome) associated with the N314D mutation of galactose-1-phosphate uridyl transferase (GALT). *Mol.Hum.Reprod.* 2:145-148.

Cunha GR (1975). The dual origin of vaginal epithelium. *Am.J.Anat.* 143:387-392.

Dang Y, Qin Y, Tang R, Mu Y, Li G, Xia M, Chen ZJ. (2012). Variants of the WNT7A gene in Chinese patients with mullerian duct abnormalities. *Fertil.Steril.* 97:391-394.

Davis RJ, Harding M, Moayedi Y, Mardon G. (2008). Mouse Dach1 and Dach2 are redundantly required for Mullerian duct development. *Genesis.* 46:205-213.

Del Priore G, Saso S, Meslin EM, Tzakis A, Brannstrom M, Clarke A, Vianna R, Sawyer R, Smith JR. (2013). Uterine transplantation--a real possibility? The Indianapolis consensus. *Hum.Reprod.* 28:288-291.

Deutscher E, Hung-Chang Yao H. (2007). Essential roles of mesenchyme-derived beta-catenin in mouse Mullerian duct morphogenesis. *Dev.Biol.* 307:227-236.

D'haene B, Hellemans J, Craen M, De Schepper J, Devriendt K, Fryns JP, Keymolen K, Debals E, de Klein A, de Jong EM, Segers K, De Paepe A, Mortier G, Vandesompele J, De Baere E. (2010). Improved molecular diagnostics of idiopathic short stature and allied disorders: quantitative polymerase chain reaction-based copy number profiling of SHOX and pseudoautosomal region 1. *J.Clin.Endocrinol.Metab.* 95:3010-3018.

Doyle JO, Laufer MR. (2009). Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome with a single septate uterus: a novel anomaly and description of treatment options. *Fertil.Steril.* 92:391.e17-391.e19.

Drummond JB, Rezende CF, Peixoto FC, Carvalho JS, Reis FM, De Marco L. (2008). Molecular analysis of the beta-catenin gene in patients with the Mayer-Rokitansky-Kuster-Hauser syndrome. *J.Assist.Reprod.Genet.* 25:511-514.

Duru UA, Laufer MR. (2009). Discordance in Mayer-von Rokitansky-Küster-Hauser Syndrome noted in monozygotic twins. *J.Pediatr.Adolesc.Gynecol.* 22:e73-5.

Edmonds DK, Rose GL, Lipton MG, Quek J. (2012). Mayer-Rokitansky-Kuster-Hauser syndrome: a review of 245 consecutive cases managed by a multidisciplinary approach with vaginal dilators. *Fertil.Steril.* 97:686-690.

Eggermann T, Schönherr N, Meyer E, Obermann C, Mavany M, Eggermann K, Ranke MB, Wollmann HA. (2006). Epigenetic mutations in 11p15 in Silver-Russell syndrome are restricted to the telomeric imprinting domain. *J.Med.Genet.* 43:615-616.

Eggermann T (2010). Russell-Silver syndrome. *Am.J.Med.Genet.C.Semin.Med.Genet.* 154C:355-364.

Ekici AB, Strissel PL, Oppelt PG, Renner SP, Brucker S, Beckmann MW, Strick R. (2013). HOXA10 and HOXA13 sequence variations in human female genital malformations including congenital absence of the uterus and vagina. *Gene.* 518:267-272.

Ellison JW, Wardak Z, Young MF, Gehron Robey P, Laig-Webster M, Chiong W. (1997). PHOG, a candidate gene for involvement in the short stature of Turner syndrome. *Hum.Mol.Genet.* 6:1341-1347.

Euling SY, Selevan SG, Pescovitz OH, Skakkebaek NE. (2008). Role of environmental factors in the timing of puberty. *Pediatrics.* 121 Suppl 3:S167-71.

Fei Q, Wu Z, Wang H, Zhou X, Wang N, Ding Y, Wang Y, Qiu G. (2010). The association analysis of TBX6 polymorphism with susceptibility to congenital scoliosis in a Chinese Han population. *Spine (Phila Pa.1976).* 35:983-988.

Folch M, Pigem I, Konje JC. (2000). Mullerian agenesis: etiology, diagnosis, and management. *Obstet.Gynecol.Surv.* 55:644-649.

Fouse SD, Nagarajan RO, Costello JF. (2010). Genome-scale DNA methylation analysis. *Epigenomics.* 2:105-117.

Frank R (1938). The formation of an artificial vagina without operation. *Am.J.Obstet.Gynecol.* 35:1053-1057.

Fritsch H, Richter E, Adam N. (2012). Molecular characteristics and alterations during early development of the human vagina. *J.Anat.* 220:363-371.

Gabory A, Ripoche MA, Le Digarcher A, Watrin F, Ziyyat A, Forné T, Jammes H, Ainscough JF, Surani MA, Journot L, Dandolo L. (2009). H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development.* 136:3413-3421.

Gabory A, Jammes H, Dandolo L. (2010). The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays.* 32:473-480.

Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. (2002). A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res.* 62:3925-3928.

Gendron RL, Paradis H, Hsieh-Li HM, Lee DW, Potter SS, Markoff E. (1997). Abnormal uterine stromal and glandular function associated with maternal reproductive defects in Hoxa-11 null mice. *Biol.Reprod.* 56:1097-1105.

Gervasini C, Grati FR, Lalatta F, Tabano S, Gentilin B, Colapietro P, De Toffol S, Frontino G, Motta F, Maitz S, Bernardini L, Dallapiccola B, Fedele L, Larizza L, Miozzo M. (2010). SHOX duplications found in some cases with type I Mayer-Rokitansky-Kuster-Hauser syndrome. *Genet.Med.* 12:634-640.

Gicquel C, Rossignol S, Cabrol S, Houang M, Steunou V, Barbu V, Danton F, Thibaud N, Le Merrer M, Burglen L, Bertrand AM, Netchine I, Le Bouc Y. (2005). Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat.Genet.* 37:1003-1007.

Gilliam L, Schulman LP. (2002). Tetralogy of Fallot, imperforate anus, and Müllerian, renal, and cervical spine (MURCS) anomalies in a 15-year-old girl. *J Pediatr Adolesc Gynecol.* 15:231-233.

Griffin JE, Edwards C, Madden JD, Harrod MJ, Wilson JD. (1976). Congenital absence of the vagina. The Mayer-Rokitansky-Kuster-Hauser syndrome. *Ann.Intern.Med.* 85:224-236.

Grigelioniene G, Schoumans J, Neumeyer L, Ivarsson A, Eklof O, Enkvist O, Tordai P, Fosdal I, Myhre AG, Westphal O, Nilsson NO, Elfving M, Ellis I, Anderlid BM, Fransson I, Tapia-Paez I, Nordenskjold M, Hagenas L, Dumanski JP. (2001). Analysis of short stature homeobox-containing gene (SHOX) and auxological phenotype in dyschondrosteosis and isolated Madelung deformity. *Hum.Genet.* 109:551-558.

Gruenwald P (1941). The relation of the growing Müllerian duct to the Wolffian duct and its importance for the genesis of malformations. *Anat. Rec.* 81:1-19.

Guioli S, Sekido R, Lovell-Badge R. (2007). The origin of the Mullerian duct in chick and mouse. *Dev.Biol.* 302:389-398.

Hammoud AO, Gibson M, Peterson CM, Kerber RA, Mineau GP, Hatasaka H. (2008). Quantification of the familial contribution to mullerian anomalies. *Obstet.Gynecol.* 111:378-384.

Iizuka-Kogo A, Ishidao T, Akiyama T, Senda T. (2007). Abnormal development of urogenital organs in Dlgh1-deficient mice. *Development.* 134:1799-1807.

Imbeaud S, Carre-Eusebe D, Rey R, Belville C, Josso N, Picard JY. (1994). Molecular genetics of the persistent mullerian duct syndrome: a study of 19 families. *Hum.Mol.Genet.* 3:125-131.

Inbar-Feigenberg M, Choufani S, Butcher DT, Roifman M, Weksberg R. (2013). Basic concepts of epigenetics. *Fertil.Steril.*

Jamin SP, Arango NA, Mishina Y, Hanks MC, Behringer RR. (2002). Requirement of Bmpr1a for Mullerian duct regression during male sexual development. *Nat.Genet.* 32:408-410.

John RM, Surani MA. (1996). Imprinted genes and regulation of gene expression by epigenetic inheritance. *Curr.Opin.Cell Biol.* 8:348-353.

Jones K (2006). Smith´s Recognizable Patterns of Malformation.

Jones HW Jr, Mermut S. (1972). Familial occurrence of congenital absence of the vagina. *Am.J.Obstet.Gynecol.* 114:1100-1101.

Josso N, Lamarre I, Picard JY, Berta P, Davies N, Morichon N, Peschanski M, Jeny R. (1993). Anti-mullerian hormone in early human development. *Early Hum.Dev.* 33:91-99.

Karnis MF, Stelling JR, Lalwani SI, Bhagavath B, Pan HA, Davis AJ, Reindollar RH, Gray MR. (2000). Mutation analysis of the HOXA13 gene in patients with congenital absence of the uterus and vagina. *J. Soc. Gynecol. Invest.* 7:172A.

Kato N, Motoyama T. (2009). Hepatocyte nuclear factor-1beta(HNF-1beta) in human urogenital organs: its expression and role in embryogenesis and tumorigenesis. *Histol.Histopathol.* 24:1479-1486.

Keniry A, Oxley D, Monnier P, Kyba M, Dandolo L, Smits G, Reik W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat.Cell Biol.* 14:659-665.

Klipstein S, Bhagavath B, Topipat C, Sasur L, Reindollar RH, Gray MR. (2003). The N314D polymorphism of the GALT gene is not associated with congenital absence of the uterus and vagina. *Mol.Hum.Reprod.* 9:171-174.

Klopocki E, Schulze H, Strauss G, Ott CE, Hall J, Trotier F, Fleischhauer S, Greenhalgh L, Newbury-Ecob RA, Neumann LM, Habenicht R, Konig R, Seemanova E, Megarbane A, Ropers HH, Ullmann R, Horn D, Mundlos S. (2007). Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am.J.Hum.Genet.* 80:232-240.

Kobayashi A, Behringer RR. (2003). Developmental genetics of the female reproductive tract in mammals. *Nat.Rev.Genet.* 4:969-980.

Kobayashi A, Shawlot W, Kania A, Behringer RR. (2004). Requirement of Lim1 for female reproductive tract development. *Development.* 131:539-549.

Kobayashi A, Kwan KM, Carroll TJ, McMahon AP, Mendelsohn CL, Behringer RR. (2005). Distinct and sequential tissue-specific activities of the LIM-class homeobox gene Lim1 for tubular morphogenesis during kidney development. *Development.* 132:2809-2823.

Kobayashi A, Stewart CA, Wang Y, Fujioka K, Thomas NC, Jamin SP, Behringer RR. (2011). beta-Catenin is essential for Mullerian duct regression during male sexual differentiation. *Development.* 138:1967-1975.

Kolatsi-Joannou M, Bingham C, Ellard S, Bulman MP, Allen LI, Hattersley AT, Woolf AS. (2001). Hepatocyte nuclear factor-1beta: a new kindred with renal cysts and diabetes and gene expression in normal human development. *J.Am.Soc.Nephrol.* 12:2175-2180.

Kreidberg JA, Sariola H, Loring JM, Maeda M, Pelletier J, Housman D, Jaenisch R. (1993). WT-1 is required for early kidney development. *Cell.* 74:679-691.

Kula S, Saygili A, Tunaoglu FS, Olgunturk R. (2004). Mayer-Rokitansky-Kuster-Hauser syndrome associated with pulmonary stenosis. *Acta Paediatr.* 93:570-572.

Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH,Jr, Dobyns WB, Christian SL. (2008). Recurrent 16p11.2 microdeletions in autism. *Hum.Mol.Genet.* 17:628-638.

Kurita T (2011). Normal and abnormal epithelial differentiation in the female reproductive tract. *Differentiation.* 82:117-126.

Lalwani S, Wu HH, Reindollar RH, Gray MR. (2008). HOXA10 mutations in congenital absence of uterus and vagina. *Fertil.Steril.* 89:325-330.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A,

Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature.* 409:860-921.

Lau YF, Li Y, Kido T. (2009). Gonadoblastoma locus and the TSPY gene on the human Y chromosome. *Birth Defects Res.C.Embryo.Today.* 87:114-122.

Laufer MR (2002). Congenital absence of the vagina: in search of the perfect solution. When, and by what technique, should a vagina be created? *Curr.Opin.Obstet.Gynecol.* 14:441-444.

Ledig S, Schippert C, Strick R, Beckmann MW, Oppelt PG, Wieacker P. (2011). Recurrent aberrations identified by array-CGH in patients with Mayer-Rokitansky-Kuster-Hauser syndrome. *Fertil.Steril.* 95:1589-1594.

Ledig S, Brucker S, Barresi G, Schomburg J, Rall K, Wieacker P. (2012). Frame shift mutation of LHX1 is associated with Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome. *Hum.Reprod.* 27:2872-2875.

Leka SK, Kitsiou-Tzeli S, Kalpini-Mavrou A, Kanavakis E. (2006). Short stature and dysmorphology associated with defects in the SHOX gene. *Hormones (Athens).* 5:107-118.

Letterie GS, Vauss N. (1991). Mullerian tract abnormalities and associated auditory defects. *J.Reprod.Med.* 36:765-768.

Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC. (2007). The diploid genome sequence of an individual human. *PLoS Biol.* 5:e254.

Lindner TH, Njolstad PR, Horikawa Y, Bostad L, Bell GI, Sovik O. (1999). A novel syndrome of diabetes mellitus, renal dysfunction and genital malformation associated with a partial deletion of the pseudo-POU domain of hepatocyte nuclear factor-1beta. *Hum.Mol.Genet.* 8:2001-2008.

Lischke JH, Curtis CH, Lamb EJ. (1973). Discordance of vaginal agenesis in monozygotic twins. *Obstet.Gynecol.* 41:920-924.

Loirat C, Bellanne-Chantelot C, Husson I, Deschenes G, Guigonis V, Chabane N. (2010). Autism in three patients with cystic or hyperechogenic kidneys and chromosome 17q12 deletion. *Nephrol.Dial.Transplant.* 25:3430-3433.

Ma J, Qin Y, Liu W, Duan H, Xia M, Chen ZJ. (2011). Analysis of PBX1 mutations in 192 Chinese women with Mullerian duct abnormalities. *Fertil.Steril.* 95:2615-2617.

Ma L (2009). Endocrine disruptors in female reproductive tract development and carcinogenesis. *Trends Endocrinol.Metab.* 20:357-363.

Marshall WA, Tanner JM. (1969). Variations in pattern of pubertal changes in girls. *Arch.Dis.Child.* 44:291-303.

Massafra C, Bartolozzi M, Bartolozzi P, Scillone L. (1988). Rokitansky-Kuster-Hauser syndrome with ectrodactyly. *Acta Obstet.Gynecol.Scand.* 67:557-560.

Matzuk MM, Lamb DJ. (2008). The biology of infertility: research advances and clinical challenges. *Nat.Med.* 14:1197-1213.

McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, Perkins DO, Dickel DE, Kusenda M, Krastoshevsky O, Krause V, Kumar RA, Grozeva D, Malhotra D, Walsh T, Zackai EH, Kaplan P, Ganesh J, Krantz ID, Spinner NB, Roccanova P, Bhandari A, Pavon K, Lakshmi B, Leotta A, Kendall J, Lee YH, Vacic V, Gary S, Iakoucheva LM, Crow TJ, Christian SL, Lieberman JA, Stroup TS, Lehtimaki T, Puura K, Haldeman-Englert C, Pearl J, Goodell M, Willour VL, Derosse P, Steele J, Kassem L, Wolff J, Chitkara N, McMahon FJ, Malhotra AK, Potash JB, Schulze TG, Nothen MM, Cichon S, Rietschel M, Leibenluft E, Kustanovich V, Lajonchere CM, Sutcliffe JS, Skuse D, Gill M, Gallagher L, Mendell NR, Wellcome Trust Case Control Consortium, Craddock N, Owen MJ, O'Donovan MC, Shaikh TH, Susser E, Delisi LE, Sullivan PF, Deutsch CK, Rapoport J, Levy DL, King MC, Sebat J. (2009). Microduplications of 16p11.2 are associated with schizophrenia. *Nat.Genet.* 41:1223-1227.

Mefford HC, Clauin S, Sharp AJ, Moller RS, Ullmann R, Kapur R, Pinkel D, Cooper GM, Ventura M, Ropers HH, Tommerup N, Eichler EE, Bellanne-Chantelot C. (2007). Recurrent reciprocal genomic rearrangements of 17q12 are associated with renal disease, diabetes, and epilepsy. *Am.J.Hum.Genet.* 81:1057-1069.

Mendelsohn C, Lohnes D, Decimo D, Lufkin T, LeMeur M, Chambon P, Mark M. (1994). Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants. *Development.* 120:2749-2771.

Mericskay M, Kitajewski J, Sassoon D. (2004). Wnt5a is required for proper epithelial-mesenchymal interactions in the uterus. *Development.* 131:2061-2072.

Miller C, Pavlova A, Sassoon DA. (1998). Differential expression patterns of Wnt genes in the murine female reproductive tract during development and the estrous cycle. *Mech.Dev.* 76:91-99.

Miller JR (2002). The Wnts. *Genome Biol.* 3:REVIEWS3001.

Mishina Y, Rey R, Finegold MJ, Matzuk MM, Josso N, Cate RL, Behringer RR. (1996). Genetic analysis of the Mullerian-inhibiting substance signal transduction pathway in mammalian sexual differentiation. *Genes Dev.* 10:2577-2587.

Miyamoto N, Yoshida M, Kuratani S, Matsuo I, Aizawa S. (1997). Defects of urogenital development in mice lacking Emx2. *Development.* 124:1653-1664.

Miyamoto Y, Taniguchi H, Hamel F, Silversides DW, Viger RS. (2008). A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. *BMC Mol.Biol.* 9:44-2199-9-44.

Moore GE, Abu-Amero S, Wakeling E, Hitchins M, Monk D, Stanier P, Preece M. (1999). The search for the gene for Silver-Russell syndrome. *Acta Paediatr.Suppl.* 88:42-48.

Morcel K, Camborieux L, Programme de Recherches sur les Aplasies Mulleriennes, Guerrier D. (2007). Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome. *Orphanet J.Rare Dis.* 2:13.

Morland SJ, Jiang X, Hitchcock A, Thomas EJ, Campbell IG. (1998). Mutation of galactose-1-phosphate uridyl transferase and its association with ovarian cancer and endometriosis. *Int.J.Cancer.* 77:825-827.

Nagamani SC, Erez A, Shen J, Li C, Roeder E, Cox S, Karaviti L, Pearson M, Kang SH, Sahoo T, Lalani SR, Stankiewicz P, Sutton VR, Cheung SW. (2010). Clinical spectrum associated with recurrent genomic rearrangements in chromosome 17q12. *Eur.J.Hum.Genet.* 18:278-284.

Nebesio TD, Eugster EA. (2007). Current concepts in normal and abnormal puberty. *Curr.Probl.Pediatr.Adolesc.Health.Care.* 37:50-72.

Nef S, Parada LF. (2000). Hormones in male sexual development. *Genes Dev.* 14:3075-3086.

Ng PC, Levy S, Huang J, Stockwell TB, Walenz BP, Li K, Axelrod N, Busam DA, Strausberg RL, Venter JC. (2008). Genetic variation in an individual human exome. *PLoS Genet.* 4:e1000160.

Niemann S, Zhao C, Pascu F, Stahl U, Aulepp U, Niswander L, Weber JL, Muller U. (2004). Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family. *Am.J.Hum.Genet.* 74:558-563.

Nik-Zainal S, Strick R, Storer M, Huang N, Rad R, Willatt L, Fitzgerald T, Martin V, Sandford R, Carter NP, Janecke AR, Renner SP, Oppelt PG, Oppelt P, Schulze C, Brucker S, Hurles M, Beckmann MW, Strissel PL, Shaw-Smith C. (2011). High incidence of recurrent copy number variants in patients with isolated and syndromic Mullerian aplasia. *J.Med.Genet.* 48:197-204.

Ogata T, Matsuo N, Fukushima Y, Saito M, Nose O, Miharu N, Uehara S, Ishizuka B. (2001). FISH analysis for apparently simple terminal deletions of the X chromosome: identification of hidden structural abnormalities. *Am.J.Med.Genet.* 104:307-311.

Oppelt P, Renner SP, Brucker S, Strissel PL, Strick R, Oppelt PG, Doerr HG, Schott GE, Hucke J, Wallwiener D, Beckmann MW. (2005a). The VCUAM (Vagina Cervix Uterus Adnex-associated Malformation) classification: a new classification for genital malformations. *Fertil.Steril.* 84:1493-1497.

Oppelt P, Strissel PL, Kellermann A, Seeber S, Humeny A, Beckmann MW, Strick R. (2005b). DNA sequence variations of the entire anti-Mullerian hormone (AMH) gene promoter and AMH protein expression in patients with the Mayer-Rokitanski-Kuster-Hauser syndrome. *Hum.Reprod.* 20:149-157.

Oppelt P, Renner SP, Kellermann A, Brucker S, Hauser GA, Ludwig KS, Strissel PL, Strick R, Wallwiener D, Beckmann MW. (2006). Clinical aspects of Mayer-Rokitansky-Kuester-Hauser syndrome: recommendations for clinical diagnosis and staging. *Hum.Reprod.* 21:792-797.

Oppelt PG, Lermann J, Strick R, Dittrich R, Strissel P, Rettig I, Schulze C, Renner SP, Beckmann MW, Brucker S, Rall K, Mueller A. (2012). Malformations in a cohort of 284 women with Mayer-Rokitansky-Kuster-Hauser syndrome (MRKH). *Reprod.Biol.Endocrinol.* 10:57-7827-10-57.

Orvis GD, Behringer RR. (2007). Cellular mechanisms of Mullerian duct formation in the mouse. *Dev.Biol.* 306:493-504.

Parr BA, McMahon AP. (1998). Sexually dimorphic development of the mammalian reproductive tract requires Wnt-7a. *Nature.* 395:707-710.

Peltomaki P (2012). Mutations and epimutations in the origin of cancer. *Exp.Cell Res.* 318:299-310.

Penaherrera MS, Weindler S, Van Allen MI, Yong SL, Metzger DL, McGillivray B, Boerkoel C, Langlois S, Robinson WP. (2010). Methylation profiling in individuals with Russell-Silver syndrome. *Am.J.Med.Genet.A.* 152A:347-355.

Petrozza JC, Gray MR, Davis AJ, Reindollar RH. (1997). Congenital absence of the uterus and vagina is not commonly transmitted as a dominant genetic trait: outcomes of surrogate pregnancies. *Fertil.Steril.* 67:387-389.

Philibert P, Biason-Lauber A, Rouzier R, Pienkowski C, Paris F, Konrad D, Schoenle E, Sultan C. (2008). Identification and functional analysis of a new WNT4 gene mutation among 28 adolescent girls with primary amenorrhea and mullerian duct abnormalities: a French collaborative study. *J.Clin.Endocrinol.Metab.* 93:895-900.

Philibert P, Biason-Lauber A, Gueorguieva I, Stuckens C, Pienkowski C, Lebon-Labich B, Paris F, Sultan C. (2011). Molecular analysis of WNT4 gene in four adolescent girls with mullerian duct abnormality and hyperandrogenism (atypical Mayer-Rokitansky-Kuster-Hauser syndrome). *Fertil.Steril.* 95:2683-2686.

Pittock ST, Babovic-Vuksanovic D, Lteif A. (2005). Mayer-Rokitansky-Kuster-Hauser anomaly and its associated malformations. *Am.J.Med.Genet.A.* 135:314-316.

Pizzo A, Lagana AS, Sturlese E, Retto G, Retto A, De Dominici R, Puzzolo D. (2013). Mayer-rokitansky-kuster-hauser syndrome: embryology, genetics and clinical and surgical treatment. *ISRN Obstet.Gynecol.* 2013:628717.

Plevraki E, Kita M, Goulis DG, Hatzisevastou-Loukidou H, Lambropoulos AF, Avramides A. (2004). Bilateral ovarian agenesis and the presence of the testis-specific protein 1-Y-linked gene: two new features of Mayer-Rokitansky-Kuster-Hauser syndrome. *Fertil.Steril.* 81:689-692.

Rall K, Barresi G, Walter M, Poths S, Haebig K, Schaeferhoff K, Schoenfisch B, Riess O, Wallwiener D, Bonin M, Brucker S. (2011). A combination of transcriptome and methylation analyses reveals embryologically-relevant candidate genes in MRKH patients. *Orphanet J.Rare Dis.* 6:32-1172-6-32.

Rao E, Weiss B, Fukami M, Rump A, Niesler B, Mertz A, Muroya K, Binder G, Kirsch S, Winkelmann M, Nordsiek G, Heinrich U, Breuning MH, Ranke MB, Rosenthal A, Ogata T, Rappold GA. (1997). Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nat.Genet.* 16:54-63.

Rao TP, Kuhl M. (2010). An updated overview on Wnt signaling pathways: a prelude for more. *Circ.Res.* 106:1798-1806.

Ravel C, Lorenco D, Dessolle L, Mandelbaum J, McElreavey K, Darai E, Siffroi JP. (2009). Mutational analysis of the WNT gene family in women with Mayer-Rokitansky-Kuster-Hauser syndrome. *Fertil.Steril.* 91:1604-1607.

Ravel C, Bashamboo A, Bignon-Topalovic J, Siffroi JP, McElreavey K, Darai E. (2012). Polymorphisms in DLGH1 and LAMC1 in Mayer-Rokitansky-Kuster-Hauser syndrome. *Reprod.Biomed.Online.* 24:462-465.

Reber M, Cereghini S. (2001). Variant hepatocyte nuclear factor 1 expression in the mouse genital tract. *Mech.Dev.* 100:75-78.

Reddy JC, Licht JD. (1996). The WT1 Wilms' tumor suppressor gene: how much do we really know? *Biochim.Biophys.Acta.* 1287:1-28.

Resendes BL, Sohn SH, Stelling JR, Tineo R, Davis AJ, Gray MR, Reindollar RH. (2001). Role for anti-Mullerian hormone in congenital absence of the uterus and vagina. *Am.J.Med.Genet.* 98:129-136.

Roberts CP, Haber MJ, Rock JA. (2001). Vaginal creation for mullerian agenesis. *Am.J.Obstet.Gynecol.* 185:1349-52; discussion 1352-3.

Roberts LM, Visser JA, Ingraham HA. (2002). Involvement of a matrix metalloproteinase in MIS-induced cell death during urogenital development. *Development.* 129:1487-1496.

Roos L, Brondum Nielsen K, Tümer Z. (2009). A duplication encompassing the SHOX gene and the downstream evolutionarily conserved sequences. *Am.J.Med.Genet.A.* 149A:2900-2901.

Rosenfeld JA, Coppinger J, Bejjani BA, Girirajan S, Eichler EE, Shaffer LG, Ballif BC. (2010). Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. *J.Neurodev Disord.* 2:26-38.

Schinzel A, Schmid W, Fraccaro M, Tiepolo L, Zuffardi O, Opitz JM, Lindsten J, Zetterqvist P, Enell H, Baccichetti C, Tenconi R, Pagon RA. (1981). The "cat eye syndrome": dicentric small marker chromosome probably derived from a no.22 (tetrasomy 22pter to q11) associated with a characteristic phenotype. Report of 11 patients and delineation of the clinical picture. *Hum.Genet.* 57:148-158.

Schnabel CA, Godin RE, Cleary ML. (2003a). Pbx1 regulates nephrogenesis and ureteric branching in the developing kidney. *Dev.Biol.* 254:262-276.

Schnabel CA, Selleri L, Cleary ML. (2003b). Pbx1 is essential for adrenal development and urogenital differentiation. *Genesis.* 37:123-130.

Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nat.Methods.* 7:575-576.

Selleri L, Depew MJ, Jacobs Y, Chanda SK, Tsang KY, Cheah KS, Rubenstein JL, O'Gorman S, Cleary ML. (2001). Requirement for Pbx1 in skeletal patterning and programming chondrocyte proliferation and differentiation. *Development.* 128:3543-3557.

Shawlot W, Behringer RR. (1995). Requirement for Lim1 in head-organizer function. *Nature.* 374:425-430.

Shears DJ, Vassal HJ, Goodman FR, Palmer RW, Reardon W, Superti-Furga A, Scambler PJ, Winter RM. (1998). Mutation and deletion of the pseudoautosomal gene SHOX cause Leri-Weill dyschondrosteosis. *Nat.Genet.* 19:70-73.

Shen Y, Yan Y, Liu Y, Zhang S, Yang D, Zhang P, Li L, Wang Y, Ma Y, Tao D, Yang Y. (2013). A significant effect of the TSPY1 copy number on spermatogenesis efficiency and the phenotypic expression of the gr/gr deletion. *Hum.Mol.Genet.* 22:1679-1695.

Shokeir MH (1978). Aplasia of the Mullerian system: evidence for probable sex-limited autosomal dominant inheritance. *Birth Defects Orig.Artic.Ser.* 14:147-165.

Simpson JL (1999). Genetics of the female reproductive ducts. *Am.J.Med.Genet.* 89:224-239.

Sparrow DB, McInerney-Leo A, Gucev ZS, Gardiner B, Marshall M, Leo PJ, Chapman DL, Tasic V, Shishko A, Brown MA, Duncan EL, Dunwoodie SL. (2013). Autosomal dominant spondylocostal dysostosis is caused by mutation in TBX6. *Hum.Mol.Genet.* 22:1625-1631.

Stark K, Vainio S, Vassileva G, McMahon AP. (1994). Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by Wnt-4. *Nature.* 372:679-683.

Stefansson H, Einarsdottir A, Geirsson RT, Jonsdottir K, Sverrisdottir G, Gudnadottir VG, Gunnarsdottir S, Manolescu A, Gulcher J, Stefansson K. (2001). Endometriosis is not associated with or linked to the GALT gene. *Fertil.Steril.* 76:1019-1022.

Steinkampf MP, Dharia SP, Dickerson RD. (2003). Monozygotic twins discordant for vaginal agenesis and bilateral tibial longitudinal deficiency. *Fertil.Steril.* 80:643-645.

Strachan T, Read A. (2004). Human Molecular Genetics. Third edition, published by Garland Science.

Strubbe EH, Thijn CJ, Willemsen WN, Lappohn R. (1987). Evaluation of radiographic abnormalities of the hand in patients with the Mayer-Rokitansky-Kuster-Hauser syndrome. *Skeletal Radiol.* 16:227-231.

Strubbe EH, Willemsen WN, Lemmens JA, Thijn CJ, Rolland R. (1993). Mayer-Rokitansky-Kuster-Hauser syndrome: distinction between two forms based on excretory urographic, sonographic, and laparoscopic findings. *AJR Am.J.Roentgenol.* 160:331-334.

Strubbe EH, Cremers CW, Dikkers FG, Willemsen WN. (1994). Hearing loss and the Mayer-Rokitansky-Kuster-Hauser syndrome. *Am.J.Otol.* 15:431-436.

Sundaram UT, McDonald-McGinn DM, Huff D, Emanuel BS, Zackai EH, Driscoll DA, Bodurtha J. (2007). Primary amenorrhea and absent uterus in the 22q11.2 deletion syndrome. *Am.J.Med.Genet.A.* 143A:2016-2018.

Tanner JM (1962). Growth at Adolescence. *Blackwell, Oxford.* 2nd ed.:

Taylor HS, Vanden Heuvel GB, Igarashi P. (1997). A conserved Hox axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the Hoxa cluster genes. *Biol.Reprod.* 57:1338-1345.

Theiler K, Varnum DS. (1985). Development of rib-vertebrae: a new mutation in the house mouse with accessory caudal duplications. *Anat.Embryol.(Berl).* 173:111-116.

Thomas NS, Harvey JF, Bunyan DJ, Rankin J, Grigelioniene G, Bruno DL, Tan TY, Tomkins S, Hastings R. (2009). Clinical and molecular characterization of duplications encompassing the human SHOX gene reveal a variable effect on stature. *Am.J.Med.Genet.A.* 149A:1407-1414.

Timmreck LS, Pan HA, Reindollar RH, Gray MR. (2003). WNT7A mutations in patients with Mullerian duct abnormalities. *J.Pediatr.Adolesc.Gynecol.* 16:217-221.

Torres M, Gomez-Pardo E, Dressler GR, Gruss P. (1995). Pax-2 controls multiple steps of urogenital development. *Development.* 121:4057-4065.

Tsang TE, Shawlot W, Kinder SJ, Kobayashi A, Kwan KM, Schughart K, Kania A, Jessell TM, Behringer RR, Tam PP. (2000). Lim1 activity is required for intermediate mesoderm differentiation in the mouse embryo. *Dev.Biol.* 223:77-90.

Vainio S, Heikkila M, Kispert A, Chin N, McMahon AP. (1999). Female development in mammals is regulated by Wnt-4 signalling. *Nature.* 397:405-409.

van Lingen BL, Reindollar RH, Davis AJ, Gray MR. (1998). Further evidence that the WT1 gene does not have a role in the development of the derivatives of the mullerian duct. *Am.J.Obstet.Gynecol.* 179:597-603.

Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, Falchi M, Chen F, Andrieux J, Lobbens S, Delobel B, Stutzmann F, El-Sayed Moustafa JS, Chevre JC, Lecoeur C, Vatin V, Bouquillon S, Buxton JL, Boute O, Holder-Espinasse M, Cuisset JM, Lemaitre MP, Ambresin AE, Brioschi A, Gaillard M, Giusti V, Fellmann F, Ferrarini A, Hadjikhani N, Campion D, Guilmatre A, Goldenberg A, Calmels N, Mandel JL, Le Caignec C, David A, Isidor B, Cordier MP, Dupuis-Girod S, Labalme A, Sanlaville D, Beri-Dexheimer M, Jonveaux P, Leheup B, Ounap K, Bochukova EG, Henning E, Keogh J, Ellis RJ, Macdermot KD, van Haelst MM, Vincent-Delorme C, Plessis G, Touraine R, Philippe A, Malan V, Mathieu-Dramard M, Chiesa J, Blaumeiser B, Kooy RF, Caiazzo R, Pigeyre M, Balkau B, Sladek R, Bergmann S, Mooser V, Waterworth D, Reymond A, Vollenweider P, Waeber G, Kurg A, Palta P, Esko T, Metspalu A, Nelis M, Elliott P, Hartikainen AL, McCarthy MI, Peltonen L, Carlsson L, Jacobson P, Sjostrom L, Huang N, Hurles ME, O'Rahilly S, Farooqi IS, Mannik K, Jarvelin MR, Pattou F, Meyre D, Walley AJ, Coin LJ, Blakemore AI, Froguel P, Beckmann JS. (2010). A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature.* 463:671-675.

Wang P, Zhao H, Sun M, Li Y, Chen ZJ. (2012). PAX2 in 192 Chinese women with Mullerian duct abnormalities: mutation analysis. *Reprod.Biomed.Online.* 25:219-222.

Warot X, Fromental-Ramain C, Fraulob V, Chambon P, Dolle P. (1997). Gene dosage-dependent effects of the Hoxa-13 and Hoxd-13 mutations on morphogenesis of the terminal parts of the digestive and urogenital tracts. *Development.* 124:4781-4791.

Watabe-Rudolph M, Schlautmann N, Papaioannou VE, Gossler A. (2002). The mouse rib-vertebrae mutation is a hypomorphic Tbx6 allele. *Mech.Dev.* 119:251-256.

Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ, Autism Consortium. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *N.Engl.J.Med.* 358:667-675.

Weksberg R, Shuman C, Caluseriu O, Smith AC, Fei YL, Nishikawa J, Stockley TL, Best L, Chitayat D, Olney A, Ives E, Schneider A, Bestor TH, Li M, Sadowski P, Squire J. (2002). Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum.Mol.Genet.* 11:1317-1325.

Wester T, Tovar JA, Rintala RJ. (2012). Vaginal agenesis or distal vaginal atresia associated with anorectal malformations. *J.Pediatr.Surg.* 47:571-576.

White PH, Farkas DR, McFadden EE, Chapman DL. (2003). Defective somite patterning in mouse embryos with reduced levels of Tbx6. *Development.* 130:1681-1690.

Willem M, Miosge N, Halfter W, Smyth N, Jannetti I, Burghart E, Timpl R, Mayer U. (2002). Specific ablation of the nidogen-binding site in the laminin gamma1 chain interferes with kidney and lung development. *Development.* 129:2711-2722.

Yoshimizu T, Miroglio A, Ripoche MA, Gabory A, Vernucci M, Riccio A, Colnot S, Godard C, Terris B, Jammes H, Dandolo L. (2008). The H19 locus acts in vivo as a tumor suppressor. *Proc.Natl.Acad.Sci.U.S.A.* 105:12417-12422.

Zenteno JC, Carranza-Lira S, Kofman-Alfaro S. (2004). Molecular analysis of the anti-Mullerian hormone, the anti-Mullerian hormone receptor, and galactose-1-phosphate uridyl transferase genes in patients with the Mayer-Rokitansky-Kuster-Hauser syndrome. *Arch.Gynecol.Obstet.* 269:270-273.

Zhao Y, Potter SS. (2001). Functional specificity of the Hoxa13 homeobox. *Development.* 128:3197-3207.