From THE DEPARTMENT OF MOLECULAR MEDICINE AND SURGERY
Karolinska Institutet, Stockholm, Sweden

CLINICAL AND MOLECULAR GENETIC CHARACTERIZATION OF CONGENITAL MALFORMATIONS

Johanna Winberg

Karolinska Institutet

Stockholm 2015

All previously published papers were reproduced with permission from the publisher.
Cover: Part of the print “Heartbeats” picturing heart from unknown animal, made after original woodcut. Courtesy of Edit & Björnen (editochbjornen.se).
Published by Karolinska Institutet. Printed by Eprint AB.

© Johanna Winberg, 2015
ISBN 978-91-7549-893-5

Clinical and molecular genetic characterization of congenital malformations
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Johanna Winberg

Principal Supervisor:
Professor Agneta Nordenskjöld
Karolinska Institutet
Department of Women's and Children's Health

Co-supervisor(s):
Associate professor Ann Nordgren
Karolinska Institutet
Department of Molecular Medicine and Surgery

Dr Peter Gustavsson
Karolinska Institutet
Department of Molecular Medicine and Surgery

Opponent:
Associate professor David Gisselsson Nord
Lund University
Department of Laboratory Medicine

Examination Board:
Associate professor Eva-Lena Stattin
Umeå University
Department of Medical Biosciences

Professor Filip Farnebo
Karolinska Institutet
Department of Molecular Medicine and Surgery

Associate professor Ann-Charlotte Thuresson
Uppsala University
Department of Immunology, Genetics and Pathology

To Dagmar, Rasmus and all of my family

ABSTRACT

Congenital malformations are important causes of perinatal mortality and morbidity, and around 4% of children are diagnosed with a malformation during their first year of life. Despite improved surgical treatment, several malformations are associated with lifelong sequelae requiring specialized health care. Important issues for these families are the etiology, prognosis and recurrence risk of the malformation in future pregnancies. Nowadays, around 50% of patients with malformations in combination with cognitive impairment receive an etiologic diagnosis after genetic evaluation. The aims of this thesis were to increase the knowledge of genetic causes behind congenital malformations, to improve clinical genetic investigations of these patients and at the same time to identify genes involved in normal and impaired organ development.

Study I Whole-body human chimerism is the result of two zygotes giving rise to one individual, and is a rarely detected condition that can cause congenital anomalies. We have studied the molecular background of a 46,XX/47,XY,+14 karyotype identified in clinical genetic investigation in a boy with disorder of sex development (DSD). Based on molecular findings, we suggest that the chimerism in our patient is the result of dispermic fertilization of a parthenogenetically activated oocyte. This study highlights chimerism as an underlying cause of distinct cell populations in an individual, and shows the difficulty of predicting the severity of associated phenotypes in mosaic or chimeric forms of genetic aberrations.

Study II Tetrasomy 14 is a rare condition associated with multiple malformations, cognitive impairment and mortality when present in non-mosaic form. We report on molecular genetic and mitochondrial studies in an 8-year-old girl with a marker chromosome 14. We showed that the marker chromosome originated from maternal meiosis and was present in all cells analyzed, providing evidence that survival beyond infancy is possible in non-mosaic forms of this condition. The results emphasize importance of updating existing data on clinical outcomes of patients with severe diseases to correspond to high standard pediatric care.

Study III VACTERL association is a condition with multiple malformations including vertebral (V) anorectal (A) cardiac (C) tracheoesophageal (TE) renal (L) and limb (L) anomalies, without a known common cause. We performed array comparative genomic hybridization (array CGH) and DNA sequencing of the candidate genes *PCSK5*, *HOXD13* and *CHD7* to investigate the role of copy number variants (CNV) and single gene defects in 39 patients and fetal cases with VACTERL association or a VACTERL-like phenotype. We identified pathogenic gene dose alterations in 2/39 patients (5%) and a pathogenic mutation in *CHD7* in one patient, while single nucleotide variants of unclear significance were detected in *PCSK5*. We concluded that copy number variants are not common causes of VACTERL association and that CHARGE and VACTERL syndromes represent important differential diagnoses.

Study IV In this study, we have investigated the hypothesis that genetic mosaicism in malformed organs could be an underlying cause of congenital malformations. Array CGH analyses using DNA derived from cardiac tissue in 23 patients with congenital heart malformations were performed, and findings of pathogenic or unclear variants were compared with presence in blood in the same individuals. We identified pathogenic gene dose alterations in 2/23 patients (9%) and did not find evidence for mosaicism. We concluded that identification of copy number variants are important in individual cases of congenital malformations and that genetic mosaicism warrants further study using other molecular genetic technologies.

Study V We have studied the presence of copy number variants in 25 patients with congenital anorectal malformations, esophageal atresia and hydronephrosis using array CGH analysis and identified pathogenic variants in 2/25 cases (8%). We describe a mosaic structural variant of tetrasomy 15 identified in a patient with syndromic esophageal atresia, and report a novel putative susceptibility region for esophageal atresia. Mosaicism for pathogenic or unclear variants was investigated in both tissue and blood in eight cases and did not reveal discrepancies. The study shows importance of copy number analysis in individual patients with gastrointestinal malformations.

General conclusions from our studies are that tissue-specific genetic mosaicism for copy number variants does not appear to be a common cause of congenital malformations and that CNV analysis is important in patients with congenital malformations as identification of high penetrance variants in some families markedly improves accuracy of recurrence risk estimations in these families.

LIST OF SCIENTIFIC PAPERS

I. **Winberg J**, Gustavsson P, Lagerstedt Robinson K, Blennow E, Lundin J, Iwarsson E, Nordenström A, Anderlid B-M, Bondeson M-L, Nordenskjöld A, Nordgren A. Chimerism resulting from parthenogenetic activation and dispermic fertilization.
Am J Med Genet Part A 152A:2277–2286. 2010.

II. **Winberg J**, Lagerstedt Robinson K, Naess K, Lesko N, Wibom R, Liedén A, Anderlid B-M, Graff C, Nordenskjöld A, Nordgren A, Gustavsson P. Partial tetrasomy 14 associated with multiple malformations.
Am J Med Genet Part A 161A:1284–1290. 2013.

III. **Winberg J**, Gustavsson P, Papadogiannakis N, Sahlin E, Bradley F, Nordenskjöld E, Svensson P-J, Annerén G, Iwarsson E, Nordgren A, Nordenskjöld A. Mutation screening and array comparative genomic hybridization using a 180K oligonucleotide array in VACTERL association.
PLoS One 9:e85313. 2014.

IV. **Winberg J**, Berggren H, Malm T, Johansson S, Johansson Ramgren J, Nilsson B, Liedén A, Nordenskjöld A, Gustavsson P, Nordgren A. No evidence for mosaic pathogenic copy number variations in cardiac tissue from patients with congenital heart malformations.
Eur J Med Genet 58 (3) 129-133. 2015.

V. **Winberg J**, Gustavsson P, Larsson M, Ehrén H, Fossum M, Wester T, Nordgren A, Nordenskjöld A. Pathogenic gene dose alterations in a subset of patients with congenital malformations.
Manuscript.

CONTENTS

1 Introduction 7
1.1 Embryology and congenital malformations 7
1.1.1 Gametogenesis and fertilization 7
1.1.2 Genetic errors of gametogenesis 10
1.1.3 Embryogenesis 10
1.1.4 Normal and impaired organogenesis 11
1.1.5 Descriptive terms 18
1.1.6 Temporal effect 19
1.1.7 Epidemiology 19
1.1.8 Causes of congenital malformations 20
1.2 Copy number variation 25
1.2.1 Definition 25
1.2.2 Phenotypic influence of CNVs 25
1.2.3 Mechanisms behind formation of CNVs 26
1.3 VACTERL association 27
1.3.1 Causes of VACTERL association 27
1.4 Mosaicism and chimerism 29
1.4.1 Definitions 29
1.4.2 Forms of mosaicism 29
1.4.3 Mosaicism in disease 32
1.4.4 Chimerism 32
2 Aims 33
2.1 Overall aims 33
2.2 Specific aims 33
2.2.1 Investigation of copy number variants using array CGH in patients with congenital malformations 33
2.2.2 Mosaicism 33
2.2.3 VACTERL association 33
3 Materials and methods 35
3.1 Patients 35
3.1.1 Study I and II 35
3.1.2 Study III 35
3.1.3 Study IV 35
3.1.4 Study V 35
3.1.5 Informed consent 36
3.2 Tissue samples 36
3.3 DNA extraction 36
3.4 DNA sequencing (Sanger method) 37
3.5 Microsatellite marker analysis and quantitative fluorescent PCR (QF-PCR) 37
3.6 Multiplex Ligation-dependent Probe Amplification (MLPA) 38

3.7 Chromosome analysis 38
3.8 Fluorescent in situ hybridization (FISH) 39
3.9 Array comparative genomic hybridization (array CGH) 39
3.9.1 Principle 39
3.9.2 Different types of microarray platforms 40
3.9.3 Interpretation of CNVs 42
4 Results and discussion 43
4.1 Genetic causes of congenital malformations detected by array CGH (studies III, IV and V) 43
4.1.1 Copy number variants in patients with congenital heart malformations 48
4.1.2 Copy number variants detected in patients with esophageal atresia 48
4.1.3 Copy number variants detected in patients with anal atresia 48
4.1.4 Copy number variants in patients with hydronephrosis 49
4.1.5 Interesting copy number variant findings in patients with congenital malformations 49
4.1.6 Summary 50
4.2 Mosaicism and chimerism in patients with congenital malformations (studies I, II, IV and V) 50
4.2.1 Aspects of mosaicism 52
4.2.2 Chimerism 54
4.2.3 Summary 55
4.3 Causes of VACTERL association (study III) 55
4.3.1 Copy number variants 55
4.3.2 Single gene defects 55
4.3.3 Patient selection 56
4.3.4 Summary 57
5 Concluding remarks and future perspectives 58
6 Sammanfattning på svenska 60
7 Acknowledgements 62
8 Referenser 65

LIST OF ABBREVIATIONS

AA	Anal atresia
AER	Apical ectodermal ridge
Alu	*Arthrobacter luteus*
Array CGH	Array comparative genomic hybridization
ART	Assisted reproductive technologies
ASD	Atrial septal defect
BAC	Bacterial artificial chromosome
BASE	Bio Array Software Environment
BAV	Bicuspid aortic valves
Bp	Base pair
CAKUT	Congenital Anomalies of the Kidney and Urinary Tract
CEP	Chromosome enumeration probe
CHARGE	Coloboma of the eye Heart defects Atresia of the choanae Retardation of growth Genital and/or urinary defects Ear anomalies and/or deafness
CNS	Central nervous system
CNV	Copy number variation
CoA	Coarctatio aortae
Cy3	Cyanine 3
Cy5	Cyanine 5
DD	Developmental delay
DECIPHER	DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources
DNA	Deoxyribonucleic acid
DSD	Disorder of sex development
EA	Esophageal atresia
ECARUCA	European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations
FISH	Fluorescent *in situ* hybridization
FoSTeS	Fork stalling and template switching

HIV	Human immunodeficiency virus
ID	Intellectual disability
ISCA	International Standards for Cytogenomic Arrays
IVF	In vitro fertilization
Kb	Kilobase pair
LH	Luteinizing hormone
LINE	Long interspersed nuclear element
LTR	Long terminal repeat
mGy	Milligray
Mb	Megabase pair
MLPA	Multiplex Ligation-dependent Probe Amplification
MMBIR	Microhomology-mediated break-induced replication
mRNA	Messenger ribonucleic acid
MS-MLPA	Methylation-specific MLPA
NAHR	Non-allelic homologous recombination
NHEJ	Non-homologous end-joining
PCR	Polymerase chain reaction
PGC	Primordial germ cell
QF-PCR	Quantitative fluorescent PCR
RNA	Ribonucleic acid
SHH	Sonic hedgehog gene
SNP	Single nucleotide polymorphism
SBDR	Swedish Birth Defects Registry
TEF	Tracheoesophageal fistula
UPD	Uniparental disomy
VACTERL	Vertebral Anal Cardiac TracheoEsophageal Renal Limb
VOUS	Variant of unclear clinical significance
VSD	Ventricular septal defect
ZPA	Zone of polarizing activity

1 INTRODUCTION

1.1 EMBRYOLOGY AND CONGENITAL MALFORMATIONS

1.1.1 Gametogenesis and fertilization

The formation of a human embryo begins with the fusion of a female and a male gamete, an oocyte and a spermatozoon, and the development of these germ cells start in the early embryo. Precursor cells, primordial germ cells (PGCs), are formed in the epiblast layer of the embryo and migrate to the developing gonads which they reach 4-6 weeks after fertilization. During migration, and as resident cells of the gonads, the PGCs proliferate by mitosis and increase in number. The differentiation into female or male gametes depends on the surrounding gonadal tissue; PGCs in the developing ovary become oogonia while PGCs surrounded by testicular tissue become spermatogonia. Female and male gametogenetic processes have different characteristics, described below.

1.1.1.1 Oogenesis

In female gametogenesis (oogenesis), a mitotic expansion of the germ cells lead to a peak amount of almost 7 million germ cells around the 5th fetal month, after which apoptosis starts, leaving around 2 million cells at birth (1). In the developing ovary, PGCs differentiate into oogonia and enter the first reduction division (meiosis I) becoming primary oocytes, the cell type that resides in the ovary until ovulation. Simultaneously, follicle formation starts and the oocytes become surrounded by somatic cells and form primordial follicles. During the remainder of fetal development, birth and childhood, a continuous reduction of the ovarian reserve through apoptosis of oogonia lead to a remainder of 200 000 follicles at the start of puberty (1). From the onset of puberty, hormonal stimulation leads to recruitment of groups of primordial follicles that mature in a process termed folliculogenesis. Primordial follicles are activated and with oocyte growth, granulosa cell proliferation and recruitment of stromal theca cells, they mature through different stages; primary follicles, secondary follicles, early antral follicles and preovulatory follicles, a process that takes around 6 months (2). Ovulation is finally induced by the peak increase of LH, when the follicle ruptures and the oocyte covered by granulose cells (the oocyte-cumulus complex) is released into the oviduct while the theca cells and remaining granulose cells of the follicle are converted to the progesterone and estradiol producing cells of the corpus luteum. During maturation of the oocyte, the volume increases about 100 fold and RNA and protein molecules required for growth of the cell accumulate in order to sustain the early development of an embryo in the event of fertilization (2). During maturation, the competence to complete meiosis is also acquired. This process is induced by the preovulatory LH peak, so that after ovulation, the released oocyte has entered into meiosis II and remains arrested at the metaphase stage.

1.1.1.2 Spermatogenesis

In contrast to oogenesis, much of the differentiation and growth in spermatogenesis take place after birth. In utero, PGCs entering testis migrate to the basal membrane of the

primitive sex cords where they rest until after birth, surrounded by supportive Sertoli cells that provide nutrients and a protective environment. Around six months after birth they differentiate into spermatogonia (3). Spermatogenesis starts in early puberty when the sex cords acquire a lumen and develop into the seminiferous tubules of the testis. In the lumen of the tubule, spermatogonia with retained stem cell function give rise to daughter cells committed to sperm development. These expand in number by mitotic divisions and differentiate from spermatogonia to primary and subsequently secondary spermatocytes, passing through meiosis I and II. Subsequently the spermatids develop acrosomal heads covering the nucleus, carrying enzymes for breakdown of the oocyte zona pellucida, and also the middle piece and tail structures. Mature spermatozoa are released into the lumen of the seminiferous tubules, leaving most of the cytoplasm behind, and are transported to the epididymis where they acquire full motility (4).

1.1.1.3 Meiosis

Differentiation of germ cells includes two reduction divisions, meiosis I and II, during which each daughter cell receives half the chromosomal content of the original cell. Mature human germ cells contain haploid genomes comprising 23 chromosomes (n) in contrast to diploid somatic cells that contain 46 chromosomes ($2n$). Meiosis differs from mitotic division, in which daughter cells normally receive the same amount of DNA material as that of the original cell, and also by the exchange of genetic material between chromosomes in a process called recombination. This exchange takes place during prophase of the first meiotic division, which comprises five phases that can be identified when dividing cells are studied in the microscope; leptotene, zygotene, pachytene, diplotene and diakinesis. Homologous chromosomes pair during the first meiotic division and are temporarily bound by synaptonemal complexes. Breaks in the DNA strands are introduced and the strands exchange their original “stem” chromosomes, resulting in new combinations of the two original homologous chromosomes. In this way the chromosomal content of the haploid genome of the mature germ cells become a mixture of the maternal and paternal homologues, and if fertilized results in contribution of a mixture of maternal and paternal genes to the new embryo. In a male germ cell approximately 52 chiasmatas, or crossings of chromosome arms, form per cell and are distributed with at least one chiasmata per chromosome arm (1). In meiosis, the sorting of chromosomes to the daughter cells adds an additional layer of genetic variation, as this occurs randomly and results in a mixture of maternal and paternal chromosomes in each germ cell. The timing and results of the meiotic divisions differ in oogenesis and spermatogenesis (Fig. 1).

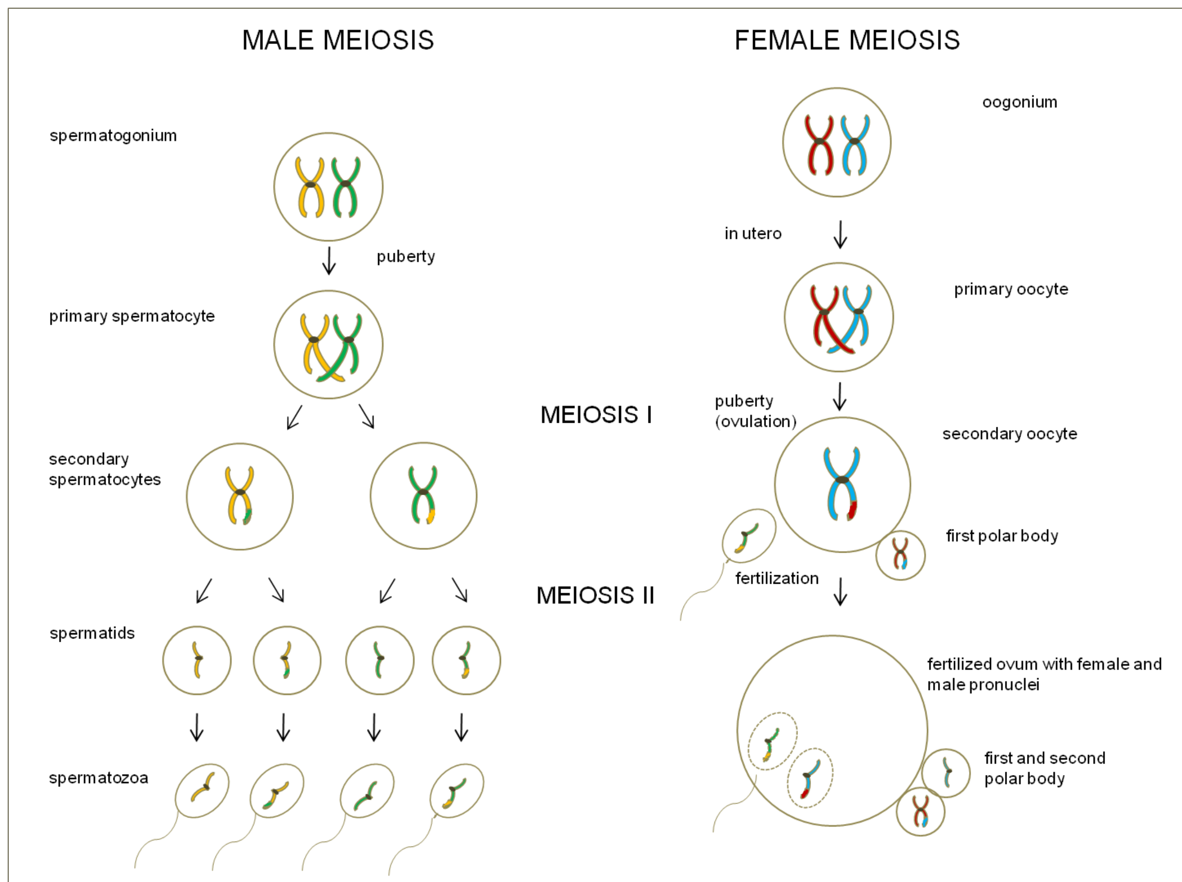

Figure 1. Female and male meiosis. In oogenesis, oocytes enter the first meiotic division in fetal life. They arrest in a prolonged diplotene phase called dictyotene, where they remain until they are activated for ovulation, resulting in a prophase of meiosis I lasting for up to several decades. Meiosis I is completed during the preovulatory phase in the menstrual cycle and the cell enters into meiosis II where it once again arrests, this time in metaphase, and is completed only in the event of fertilization. The meiotic divisions of the oocyte take place in an asymmetric fashion, so that one of the daughter cells, the secondary oocyte, receives the majority of the cytoplasm and its contents. The other cell remains as a small polar body between the cell membrane and the zona pellucida and sometimes undergoes the second division but is not known to take part in fertilization. In this way all protein and RNA molecules accumulated in the developing oocyte are retained in one oocyte that has the potential to contribute to the embryo.
Spermatogenesis starts at the onset of puberty, and from then on occurs continuously in the testes. Both meiotic divisions are completed in the testis, and mature sperm that reach the female genital tract contain haploid genomes. In male meiosis distribution of cytoplasm is symmetric, with equal amounts of cytoplasm being divided between the daughter cells, resulting in four functioning sperm from a single spermatocyte.

1.1.1.4 Fertilization

The spermatozoon is transported by movements of the uterus and oviduct as well as its own movement into the ampulla of the oviduct where fertilization often takes place (4). In the female genital tract, spermatozoa undergo a process called capacitation which is required for passage through the corona cells of the oocyte-cumulus complex. Upon interaction between spermatozoon and oocyte-cumulus complex, the acrosome reaction takes place, with release of enzyme-containing granules that aid the spermatozoon in the passage through the zona pellucida. Interaction with the zona pellucida also induces a protein modification reaction that inhibits additional sperm to fuse with the oocyte, as well as zona pellucida protein breakdown through protease activity (5), enabling the spermatozoon to come into contact with the oocyte membrane. After sperm adhesion to the oocyte, their germ cell membranes fuse and the head

and tail of the spermatozoon enter the oocyte cytoplasm (4). After sperm entry, the oocyte completes meiosis II, ending with expulsion of the second polar body. The remaining chromosomes are arranged in the female pronucleus, while the sperm nucleus forms the male pronucleus and the DNA of both pronuclei is replicated. During the earliest stages of embryo development, no zygotic gene transcription occurs and events are controlled by mRNA and proteins contributed by the oocyte, accumulated during oogenesis (6). The first mitoses of the embryo, the cleavage stage, take place without cell growth and result in reduction of cell sizes. Subsequently, zygotic gene transcription is initiated and maternal control of the embryo diminished.

1.1.2 Genetic errors of gametogenesis

Oogenesis and spermatogenesis render female and male gametes susceptible to distinct genetic abnormalities that can be transferred to the embryo and result in diseased offspring. The prolonged meiosis I of the oocyte is believed to render it susceptible to failed separation of homologous chromosomes at completion of meiosis I, so called *non-disjunction*, resulting in trisomy conceptions upon fertilization, a risk that increases distinctly at a maternal age of 35 years. Consequently, the majority of the extra chromosomes 21 found in Down syndrome are of maternal origin. In contrast, spermatogenesis is characterized by a high number of replication cycles due to the many cell divisions required for lifelong sperm production, which increases the likelihood for introduction of replication based errors such as point mutations and CNVs (7). Indeed, 80% of *de novo* point mutations in patients with autism spectrum disorders and schizophrenia have been shown to be paternal in origin, and the frequency of total *de novo* mutations in offspring, unrelated to phenotype, increase with paternal age (8, 9).

1.1.3 Embryogenesis

During the development of the fertilized egg from a zygote to a full-grown fetus, different developmental periods can be defined. The *embryonic period* lasts from fertilization until the eighth week. *Blastogenesis*, ending with the formation of the three-layered embryo, takes place during the first two weeks. From the end of the second week until the eighth week, precursors of all organ systems are formed during *organogenesis*. The *fetal period*, which lasts from the ninth week until birth, is characterized by maturation and growth of the organs.

After fertilization, the embryo develops into a morula at the 16-cell stage, when the cells separate into an inner cell mass (the embryoblast) giving rise to the embryonic structures, and an outer cell mass (the trophoblast), that contributes to the placenta. The morula develops into a blastocyst, visually characterized by a fluid-filled cavity, and after hatching from the zona pellucida, implantation into the uterine cavity starts. Cells of the trophoblast, called syncytiotrophoblast, invade the endometrial stroma while the embryoblast differentiates into two layers: the epiblast and the hypoblast. The syncytiotrophoblast invade the maternal capillaries in the uterine stroma and establishes a connection with the maternal blood flow that result in the first uteroplacental circulation. An important process in embryo development

that takes place during this period is the determination of anterio-posterior, dorso-ventral and left-right body axis, which is a tightly coordinated event regulated by gene families important in embryological development. The embryo subsequently undergoes a process termed gastrulation, in which cells of the epiblast proliferate, migrate and invaginate through the primitive streak. They subsequently form the three germ layers from which the different tissues and organ systems in the embryo develop during organogenesis; ectoderm, mesoderm and endoderm. The derivatives of the different germ layers are listed in Table 1.

Table 1. Embryological origin of human tissues (10)

Ectodermal derivatives	Mesodermal derivatives	Endodermal derivatives
nervous tissue	skeletal muscle	epithelium of digestive tract (except oral and anal cavities)
epidermis	smooth and cardiac muscle	glandular derivatives of digestive tract (liver and pancreas)
epidermal derivatives (hair, hair follicles, sweat glands, nails)	connective tissue (cartilage, bones)	epithelium of respiratory tract
cornea, lens of eye	blood, bone marrow, lymphoid tissues	thyroid, parathyroid, and thymus glands
epithelium of oral and nasal cavities, paranasal sinuses, anal canal	endothelium of blood vessels and lymphatics	epithelium of reproductive ducts and glands
tooth enamel	lining of body cavities	epithelium of urethra and bladder
epithelium of pituitary and pineal glands, and adrenal medulla	organs of urogenital system (kidneys, ureters, gonads, reproductive ducts)	

1.1.4 Normal and impaired organogenesis

During the third week the notochord is formed in the mesodermal layer of the embryo. This midline signalling structure is crucial for folding and patterning of the embryo,. The notochord functions transiently during embryogenesis and produces developmentally important morphogens, such as the sonic hedgehog protein (SHH), that regulate development of surrounding tissues.

During organogenesis the embryo undergoes a series of foldings in cephalo-caudal and lateral direction, resulting in formation of internal cavities and body wall closure. Development of some organs will be described briefly below.

1.1.4.1 Development and malformations of the heart

The heart is mainly formed from two different populations of mesodermal cells; the first and second heart field. Cells from the first heart field form the cardiac crescent in the cranial end of the embryonic disc. As a result of the cephalocaudal and lateral foldings of the embryo, the cardiac crescent fuses and forms the primitive heart tube. The second heart field is located medial to the cardiac crescent and contributes to most parts of the heart while the first heart field contributes to the left ventricle and atrial appendages (11). In addition, cells from the neural crest migrate into the region of the developing heart and contribute cells to the outflow

tract and the great arteries. The epicardium is formed from cells of the proepicardial organ (11). Formation of the heart is a complex three-dimensional morphogenetic process that results in the formation of two parallel circulations from the original “serial” circulation. The process involves looping of the cardiac tube, growth of endocardial cushions and formation of muscular contributions that participate in septation of atria, ventricles and great arteries and formation of the heart valves. The formation of the heart is regulated by temporally and spatially restricted gene expression, and biomechanical forces on the developing heart tissue, exerted by blood flow, also affects morphogenesis (12).

Heart malformations are the most commonly diagnosed malformations, with an incidence of approximately 0.8% and can vary in clinical importance from subclinical to life-threatening (13). The most sensitive period in heart development is during the fifth week after fertilization (14). Malformations include different types of atrial and ventricular septal defects (ASDs and VSDs) that result from disturbances of the septation processes, and atrioventricular septal defect in which a common atrioventricular canal persists together with atrial and septal defects due to failure of fusion of the superior and inferior endocardial cushions. Failure of the outflow tract to form properly can result in a common arterious trunc (truncus arteriosus) instead of separation into aorta and the pulmonic trunc, or in transposition of the great arteries when the spiralling of the conotruncal septum is impaired causing the aorta to arise from the right ventricle and the pulmonary artery from the left. Tetralogy of Fallot is a sequence of structural anomalies caused by the unequal division of the conus region due to anterior displacement of the conotruncal septum, resulting in right ventricular outflow tract obstruction, a ventricular septal defect, displacement of the aorta to the right and right ventricular hypertrophy. Impaired development of the right atrioventricular valves may lead to tricuspid atresia, which can also be a part of the more complex Ebstein anomaly. Valvular stenosis of the semilunar valves in the aorta and pulmonary artery results from varying degrees of fusion of the valves, and depending on the degree of fusion can result in underdevelopment of the great vessel associated with the defect valves. Hemodynamic alterations leading to reduced aortic blood flow from the left ventricle are believed to affect development of the aorta and result in constriction of an aortic segment, coarctation of the aorta (Fig. 2).

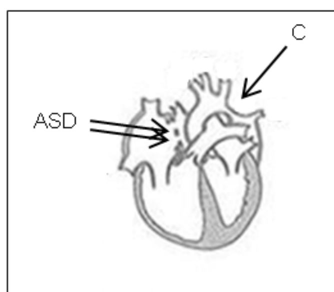

Figure 2. Heart malformations. Figure showing combined heart malformation with two atrial septal defects (ASDs) and coarctation of the aorta (C). Modified from Stevenson and Hunter, *Mol Syndromol* 2013 (14).

Many gene families are known to be involved in cardiac morphogenesis, and some of the most important known transcription factor families involved in heart development are *NKX2.5*, *GATA4* and *TBX1*.

Treatment of different malformations varies from surgical repair within the first day after birth to none. Closure of ASD and dilatation of pulmonary valve stenosis can often be performed endoscopically, while repair or replacement of heart valves, dilatation of arteries and outflow tracts and repair of complex defects is performed with open heart surgery.

1.1.4.2 Development and malformations of the gastrointestinal tract

The primitive gut is formed when the embryonic disc undergoes cephalocaudal and lateral folding, by enclosure of part of the yolk sac cavity, lined by endoderm, into the developing embryo. The primitive gut stretches between the anterior oropharyngeal membrane and the posterior cloacal membrane and is divided into the foregut, midgut and hindgut. The endoderm gives rise to the epithelial lining of the intestine, lung, pancreas and thyroid and to parenchymal cells of the liver, while smooth muscle, connective tissue, the peritoneum and stromal components of the glands arise from visceral mesoderm. The foregut gives rise to the esophagus and future lung, with a tracheoesophageal septum gradually separating the future respiratory system from the esophagus. Both the stomach and the primary intestinal loop, arising from the midgut, undergo rotation during development. The hepatic diverticulum, or liver bud, arises from the end of the foregut. The hindgut extends into the cloaca, where the urorectal septum separates the dorsal anorectal canal from the ventrally located primitive urogenital sinus. The urorectal septum subsequently forms the perineal body between the openings of the urogenital sinus and the cloaca, the future urinary and fecal excretion systems. Gut connection with the surroundings is created by breakdown of the oropharyngeal and cloacal membranes later during development.

Disturbances of gastrointestinal development can affect different parts of the system and result in malformations such as body wall defects (gastroschisis and omphalocele), malrotation of the primary intestinal loop and atresia of different parts of the gut. Esophageal and anal atresia will be described further below.

Esophageal atresia (EA)

Esophageal atresia is a malformation that occurs in 1/3 500 births and develops during the fifth week after fertilization (14, 15).

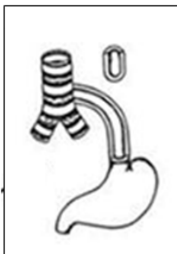

Figure 3. Esophageal atresia. Figure showing the most common form of esophageal atresia (~87%), where the proximal esophagus ends blindly while the distal part of the esophagus has a connection to the trachea, termed tracheoesophageal fistula (TEF). Other types of esophageal atresia with or without TEF occur in a lower frequency (17). Modified from Stevenson and Hunter, *Mol Syndromol* 2013 (14).

The knowledge of the mechanisms that result in EA/TEF in these malformations is still very limited, although pathogenesis has recently been shown to involve disturbance in molecular specification of the ventral and dorsal walls of the foregut (16).
Genes in which mutations are known to cause esophageal atresia include *SOX2*, *MYCN* and *CHD7* (16).

The malformation is surgically repaired with creation of an esophageal anastomosis within the first day after birth. The reported mortality is approximately 10% and is related to associated anomalies, compared to 100% mortality without surgery. Lifelong gastric and pulmonary complications are common in these patients (17).

Anorectal malformations (ARM)

Anorectal malformations occur in around 1/5 000 births and can present in a spectrum from imperforate anus with or without fistula (rectoperineal, vestibular, urethral or vesical) to cloacal malformation (18). The malformations result from abnormal development of the cloaca and urorectal septum during the seventh week after fertilization, but the pathogenic mechanisms behind the malformations are largely unknown.

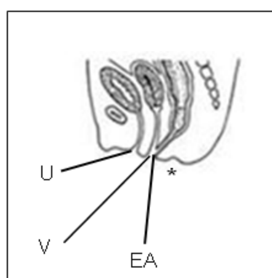

Figure 4. Anorectal malformations. Figure showing ectopic, underdeveloped anal opening at the base of the vagina. U urethra, V vagina, EA ectopic anus, * marking normal location of anus. Modified from Stevenson and Hunter, *Mol Syndromol* 2013 [Stevenson and Hunter, 2013].

Genes in which mutations are known to cause syndromic ARM in humans include *MNX1*, *SALL1* and *GLI3*, while mutations in *HOXD13* have been described in a single patient (19, 20).

Less severe anorectal malformations can be repaired in a single procedure after birth while patients with more severe forms receive a colostomy in the neonatal period until final repair can be performed at 1-2 months of age (18). The technique used for repair is called posterior sagittal anorectoplasty and involves division of the fistula, release of the gut and placement in normal position as well as creation of the pelvic floor muscles and external anal sphincter. Obstipation and urinary or fecal incontinence are common sequelae after surgery because of impaired development of the inner sphincter muscle and associated nerves.

1.1.4.3 Development and malformations of the renal system

The urological system originates from mesoderm, and two different transient systems (the pronephros and mesonephros) develop and regress before the permanent system is formed from the metanephros. The permanent collecting system including ureter, renal pelvis and calyces, develops from the ureteric bud which springs from the mesonephric (Wolffian) duct, invades the metanephric mesoderm and induces differentiation into renal parenchyma.

Congenital Anomalies of the Kidney and the Urinary Tract (CAKUT) have an incidence of 1/500 births (21). Disturbance of renal development during the fifth to eighth week can lead to uni- or bilateral renal agenesis with absence of ureter and kidney, if the ureteric bud fails to grow from the nephric duct. Inversely, duplex ureter and kidney, or duplex ureter and collecting system can result from growth of supernumerary ureteric buds. Impaired

interaction between the ureteric bud and the metanephric mesenchyme can result in renal hypodysplasia. Kidneys that are located more inferiorly than normally and are fused in the caudal lobes are called horseshoe kidneys. Impaired closure of the urinary bladder, bladder exstrophy, is a severe and very rare type of malformation (incidence of 1/10 000 to 1/50 000 in live births) (22).

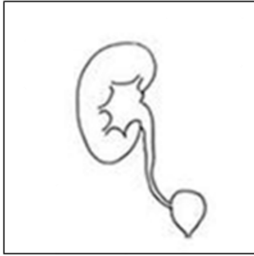

Figure 5. Renal malformations. Figure showing left unilateral renal agenesis and hydronephrosis of the right kidney. Modified from Stevenson and Hunter, *Mol Syndromol* 2013 (14).

Genes known to be involved in renal development and associated with human disease are the *HNF1β*, *PAX2* and *RET* genes (21).

While CAKUT can be subclinical, the more severe forms account for around 30% of renal failure in children, all of whom need to undergo renal replacement therapy (21).

1.1.4.4 Development and malformations of the reproductive system

Early gonads form as genital ridges in close relation to the mesonephros of the developing kidney, and primordial germ cells migrate to, and invade the ridges where they induce further gonadal development. Identical primitive genital structures develop in female and male embryos; bipotential gonads, genital ducts (Müllerian and Wolffian ducts) and external genitalia (genital tubercle, urethral folds and genital swellings) and all structures initially retain the possibility to develop into both female and male systems, depending on molecular signalling. In the event of female development, gonads differentiate into ovaries which produce estrogens that stimulate the Müllerian ducts to fuse and form female internal genitalia (the uterus, the oviducts and the upper part of the vagina) while the genital tubercle becomes the clitoris, the urethral folds become the labia minora and the genital swellings form the labia majora. In male development, gonads differentiate into testes that produce several important hormones; anti-Müllerian Hormone (AMH), which leads to regression of the Müllerian ducts and more importantly testosterone, which stimulates the Wolffian ducts to develop into the vas deferens, the rete testis, the epididymis and the seminal vesicle. External genitalia are stimulated by the testosterone metabolite dihydrotestosterone, which leads to elongation of the genital tubercle into the phallus, the formation of the penile urethra from the urethral folds and scrotal development from the genital swellings. The testes are formed in the abdomen, and during fetal development descend into the scrotum in two phases, both hormonally regulated from the testes. The transabdominal phase takes place between week 8 and 15 and is regulated by insulin-like hormone 3, while the second inguinoscrotal phase occurs between week 25 and 35 and is androgen-dependent (23).

Genital malformations in females

Uterine malformations are reported in 4% of reproductive women, and are caused by impaired development or fusion of the Müllerian ducts in the female fetus (24). This can lead to complete uterine agenesis, duplications of the whole or part of the uterus (uterus didelphys or uterus bicornis) or failed reabsorption of the midline septum (septate uterus or arcuate uterus). Uterine malformations are associated with recurrent pregnancy loss, infertility and obstetric complications. For septate uterus, the delivery rate is as low as 40% untreated (24). Atresia of part of both Müllerian ducts can result in cervical atresia, and failed development or impaired fusion of the sinovaginal bulbs can cause vaginal atresia or duplication of the vagina, respectively.

In many cases anatomically less complicated congenital uterine malformations, such as septate or arcuate uterus, may be treated surgically by hysteroscopy with drastic reduction of the risk for pregnancy loss. More complex malformations, such as uterus didelphys, may be associated with better pregnancy outcomes if not treated (25).

Genital malformations in males

The most common malformation of male genitalia is hypospadias, which occurs with an incidence of 8/1000 live-born boys in Sweden (26). Hypospadias results from incomplete fusion of the urethral folds that causes abnormal location of the urethral meatus at any point from the glans penis to the perineum along the ventral midline of the penis. Other features associated with hypospadias are ventral penile curvature (chordee) and cleaved ventral prepuce (27). Surgical repair, urethroplasty, is often performed between 6 and 18 months of age and can also include chordee excision, penile straightening, urethral, glanular and meatal reconstruction as well as skin reconstruction (27).
A severe malformation is epispadias in which the urethral meatus is located on the dorsal side of the penis. This malformation is often associated with exstrophy of the bladder with incomplete closure of the body wall. Other penile malformations include micropenis, defined as 2.5 standard deviations below mean dorsal penile length, and bifid or duplicated penis as the result of a split genital tubercle.

Disorders of sex development (DSD)

DSD are defined as conditions when chromosomal, gonadal or anatomical sex is atypical, and the incidence of genital anomalies, when genitalia are not easily classified as male or female, is approximately 1/4 500 births (28). These conditions are classified into sex chromosome DSD, 46,XX DSD and 46,XY DSD. Chromosomal DSD includes 47,XXY (Klinefelter syndrome), 45,X (Turner syndrome), 45,X/46,XX and other mosaic or structural chromosomal aberrations leading to monosomy of whole or part of the X-chromosome, 45,X/46,XY (mixed gonadal dysgenesis) and 46,XX/46,XY (mosaicism or chimerism) karyotypes (29). Both 46,XX and 46,XY DSD can be further classified into disorders of

gonadal development (gonadal dysgenesis, ovotesticular DSD as well as testicular DSD in 46,XX and testis regression in 46,XY) disorders related to excess (46,XX) or deficit (46,XY) of androgens and other causes. DSD can lead to sterility, but in the most common form (under-masculinization of male fetuses) fertility is often unaffected. Importantly, gonadal dysgenesis is associated with a high risk for gonadal tumors (30). The treatment of individuals with DSD is complex with respect to gender assignment, hormone replacement, surgical as well as psychosocial management and is ideally handled by multidisciplinary teams at specialized units.

Genes important in female development include *WNT4* and *DAX1* while and in male development *SRY*, *WT1, SOX9*, *FGF9* and *SF1* (4).

1.1.4.5 Development and malformations of the vertebrae

The vertebrae develop from somites, segments that form from mesoderm on either side of the notochord. Sclerotome cells from the somites migrate to meet cells from the opposing somite, surrounding the neural tube. The vertebrae are formed through resegmentation in which the caudal part of each somite grow into and fuse with the cranial part of its caudal neighbour while the mesenchyme in the middle of the initial segments remain and form part of the intervertebral discs.

Vertebral malformations result from an abnormal formation or segmentation process of the somites during the fourth and fifth weeks after fertilization, and are reported in <1/1000 live births, but the true frequency is likely higher because of asymptomatic forms (31). Unilateral complete failure of formation results in different types of hemivertebrae, while partial failure can give rise to wedge vertebrae. Failure of midline fusion of the paired somite halves of the vertebrae can lead to development of so called butterfly vertebrae. Unilateral segmentation defects can result in unilateral unsegmented bars, while bilateral segmentation defects can result in block vertebrae. Vertebral anomalies are associated with development of scoliosis, kyphosis and back pain, especially the combination of a unilateral unsegmented bar and contralateral multiple segmented hemivertebrae. The simultaneous presence of rib anomalies can impair lung development (31).

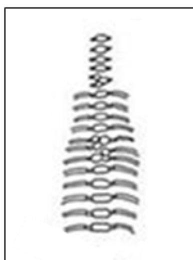

Figure 6. Vertebral malformations. Figure showing multiple segmentation defects of the cervical and thoracic vertebrae including 14 thoracic vertebrae and ribs. Modified from Stevenson and Hunter, *Mol Syndromol* 2013 (14).

Genes important in vertebral development and associated with malformations in humans include *HES7*, *DLL3* and *MESP2* (31).

Surgical procedures for congenital scoliosis include fusion procedures which may have side effects of restrictive lung disease, or procedures involving fixation material that allows

growth of the spine, with the downside of requirement for multiple surgical procedures and other complications (31).

1.1.4.6 Development and malformations of the limbs

The limbs develop from the limb buds that arise from the body wall, with a core forming from mesoderm and a cover of ectodermal cells. The forelimbs form before the hindlimbs. The growth of the limb buds is stimulated from the apical ectodermal ridge (AER), a signalling center at the distal end of the limb bud, and differentiation of the tissue into cartilage and muscle takes place in a proximal-to-distal direction. Development of the hands and feet involves apoptosis in the AER, leading to separation of tissue into the fingers and toes. Another signalling center, the zone of polarizing activity (ZPA), determines antero-posterior patterning of the limb. Ossification of bones in the extremities takes place from ossification centers and starts at the end of the embryonic period. After birth, the diaphyses are often ossified but the epiphyseal plate is still cartilaginous and is responsible for growth of the long bones.

Limb malformations occur during the sixth week after fertilization (14). The incidence of congenital upper limb malformations is 1-2/500 live births (32, 33), which is higher than for lower limb anomalies. Limb malformations can be classified into defects that result from failure of formation, failure of differentiation, duplication, over- or undergrowth, constriction ring syndrome or as a part of generalized abnormalities and syndromes (33). Limb malformations can affect a whole extremity as in amelia (absence of limb) or micromelia (underdevelopment of a limb) or only the long bones as in phocomelia (long bones shortened or absent in one or more extremities), hemimelia (shortening or absence of one of the long distal bones radius, ulna, tibia or fibula) or sirenomelia (fusion of the legs). Hand and foot anomalies include for example syndactyly (fusion of digits), poly- or hypodactyly (more or less than five digits), brachydactyly (short digits) and ectrodactyly (deficiency or absence of middle digits) (34).

Genes known to be important in limb development, and that cause malformations in humans when mutated, include *HOXD13*, *TBX4*, *TBX5* and *FGF10* (4).

Treatment options include correcting surgical procedures and custom-made prosthetic limbs in some cases, with the goal to improve function and life quality (34).

1.1.5 Descriptive terms

Congenital *malformations* are defined as primary structural defects that result from errors of morphogenesis (35). In comparison, *secondary defects* occur due to disruptions of normal morphogenesis, for example by teratogenic agents or trauma. Further, *deformations* are defined as alterations in shape or structure of normally formed fetal parts. In this thesis, the term malformation will be used for both primary and secondary defects, since the underlying causes of observed malformations are often unknown.

Malformations can occur isolated or in combination with other congenital abnormalities. A *malformation sequence* is a condition where one initial malformation leads to anomalies in other structures, whereas a *malformation syndrome* is a recognized pattern of malformations that occur as a response to one common cause. In cases with recognized malformation patterns where no common cause is known, the term *malformation association* is used.

Description of congenital malformations includes different terms for failed or improper development. *Agenesis* is used to describe the absence or failed development of an organ due to absence of the embryologic primordium of the organ, for example renal agenesis. *Atresia* refers to the congenital absence of body openings or closure of tubular structures, e.g. anal and pulmonary atresia.

1.1.6 Temporal effect

Congenital malformations occur as a response to events that affect the embryonic or fetal period of development. The first two weeks after fertilization, are considered an "all-or-nothing" period, when the cells of the embryo retain a high degree of potency and one cell can replace another in case of loss. However, if an insult occurs that leaves too many cells damaged, the embryo may be lost, although in some cases the result may not be lethal and may result in malformations (36, 37). During the period of organogenesis, when all organ systems are founded and the degree of cell differentiation is higher, the risk for organ malformations developing from embryonic insults is high. Early events before the end of the fourth week can result in severe neural, cardiac, gastrointestinal, renal and extremity defects while events in the second half of the organogenesis lead to milder defects. Insults during the fetal period, from the ninth week until birth, have less impact on organ development and are known to affect facial features, which can result in dysmorphism (38).

1.1.7 Epidemiology

Advances in pediatric surgery and intensive care have markedly increased survival among children with congenital malformations during the last decades, but congenital malformations are important causes of perinatal mortality and morbidity, and in Sweden, malformations or chromosomal aberrations are found in 30% of infants that die in the perinatal period (38). Severe birth defects in infants and aborted fetuses that are diagnosed within one month after birth are reported by neonatologists/pediatricians, obstetricians/gynecologists, pediatric cardiologists and cytogenetic laboratories to the Swedish Birth Defects Registry (SBDR), held by the National Board of Health and Welfare. In 2012, 2053 children (1.9%) were born with and 568 fetuses were aborted due to congenital malformations, though it should be noted that underreporting to this registry is a known problem. In addition, 183 children (0.2%) were born with and 408 fetuses were aborted due to chromosomal aberrations. Reported malformations were isolated (affecting one organ system) in 97% of children, and 60% of the reported children were male. Table 2 shows the reported number of cases of malformations in different organ systems in 2012 (39). Including less severe malformations, it is estimated that

around 4% of children have a malformation that is discovered during their first year of life (38).

Table 2. Congenital malformations reported to the SBDR during 2012 (39)

Type of malformation	Live births[1]	Aborted fetal cases
Heart malformation	745	34
Hypospadias	289	0
Cleft lip and palate	95	3
Congenital talipes equinovarus	67	5
Cleft palate	59	2
Limb reduction defect	44	7
Eye malformation	39	0
Cystic diseases of the kidney	34	6
Severe ear malformation	31	0
Polydactyly	31	1
Anorectal malformation	25	5
Hernia of the body wall	22	15
Esophageal atresia/stenosis	20	1
Gastroschisis	17	4
Renal agenesis/hypoplasia	15	7
Diaphragmatic hernia	13	7
Other small intestine anomalies	12	0
Hydrocephalus	10	15
Duodenal atresia	10	1
Spina bifida	9	23
Hirschprung's disease	8	0
Generalized skeletal dysplasia	7	13
Omphalocele	5	11
Disorder of sex development	4	0
Anencephaly	3	21
Microcephaly	3	0
Encephalocele	1	3

SBDR Swedish Birth Defects Registry.
[1]Total number of births in 2012 was reported as 109 848.

1.1.8 Causes of congenital malformations

Following technical advances in molecular biology, the knowledge of the etiology behind congenital malformations has increased considerably. Still, the etiology can be determined in just over 50% of patients with congenital malformations in combination with developmental delay, and in a considerably lower proportion of patients with isolated malformations.

Known causes of congenital malformations include genetic and environmental factors that influence organogenesis and cause fetal abnormalities. It has also long been suggested that combinations of genetic and environmental factors are responsible for disease.

1.1.8.1 *Genetic causes*

Genetic aberrations of different size and type can cause congenital malformations and developmental delay.

Chromosomal aberrations

Losses and gains of chromosomal material that can be detected by chromosome analysis are large (≥~ 5 Mb) and are believed to affect embryonic development by changing the gene dose of many genes simultaneously, including developmentally important genes that are sensitive to gene dose alterations.

Chromosomal aberrations usually affect neurological development and cause developmental delay and often malformations. Chromosomal aberrations include aneuploidies, with loss or gain of whole chromosomes, and deletions and duplications that affect parts of chromosomes. Constitutional aneuploidies seen in live-born children are trisomy 13, 18 and 21 and sex chromosome abnormalities, while other aneuploidies are not compatible with postnatal survival in constitutional form. Chromosomal translocations lead to different types of deletions and/or duplications that are most often unique to each family while other deletions or duplications are recurrent, such as Wolf-Hirschhorn syndrome (deletion of 4p) and cri du chat syndrome (deletion of 5p). Chromosomal aberrations can be caused by non-disjunction of homologous chromosomes or sister chromatids in the first or second meiotic divisions in parental germ cells in the case of aneuploidies, and from malsegregation of chromosomes resulting in unbalanced germ cells in translocation carriers.

Microdeletions and microduplications

Deletions and duplications smaller than 5 Mb can be detected by screening analyses such as array CGH or whole genome/exome sequencing, or by specific analyses such as FISH analysis, MLPA or quantitative PCR. Microdeletions and microduplications result in dose change for a lower number of genes compared to chromosomal aberrations and can affect development through several mechanisms: gene dose effect, position effects, gene disruptions or gene fusions. Recurrent microdeletions and microduplications are known causes behind malformation syndromes such as 22q11 deletion syndrome, Williams syndrome (deletion of 7q11.2) and Potocki-Lupski syndrome (duplication of 17p11.2), while the effect of rare deletions and duplications are more difficult to predict. Microdeletions and microduplications that cause disease often occur *de novo* in affected individuals, but some pathogenic gene dose alterations are inherited from parents with the same or milder phenotypes (40). Gene dose alterations, or copy number variations, are further discussed in chapter 1.2.

Single gene defects

Mutations in single genes (insertions, deletions and base substitutions) can be detected by sequence analysis of specific genes and also by screening analysis using whole genome/whole exome approaches. Single gene mutations that lead to loss of function, gain of function or altered function can cause disease including malformations for example in

CHARGE syndrome (*CHD7* mutations), Kabuki syndrome (*KMT2D* or *KDM6A* mutations) and Fanconi anemia (mutations in one of at least 15 different genes can cause disease). Single gene defects occur *de novo* or are inherited in a dominant, recessive or X-linked fashion.

Imprinting

Imprinting refers to a normal process that results in inactivation of alleles or regions of homologous chromosomes in a determined pattern dependent on the parental origin of the allele/homologous chromosome, and is thought to affect around 1% of human genes (38). Imprinted genes or regions are susceptible to disease in the case of deletion or inactivation of the active allele or gene region. In the 15q11-q12 region, some genes are imprinted on the maternal homologue while others are imprinted on the paternal homologue, and therefore deletion of the same genetic region can cause different phenotypic effects in individuals depending on the parental origin of the homologue that is affected by the deletion (Prader-Willi and Angelman syndromes). Other examples of imprinting disorders are Beckwith-Wiedemann and Silver-Russell syndromes. Imprinting disorders are sometimes caused by the phenomenon of uniparental disomy (UPD), in which case both homologues in a chromosome pair originate from the same parent.

The de novo mutation concept

A de novo mutation has been defined as “a genetic variation that is present for the first time in one family member (child) as a result of a mutation in a germ cell (egg or sperm) of one of the parents, or has occurred in the fertilized egg itself” (9) and are associated with sporadic disease as opposed to familiar disease caused by inherited mutations. Conclusions from massive parallel sequencing analysis in patient-parent trios in patients with neurodevelopmental phenotypes, is that *de novo* mutations in many different genes collectively represent an important role in phenotypes such as intellectual disability, autism spectrum disorders and schizophrenia (9). The importance of *de novo* mutations to a specific phenotype is proposed to correlate with the size of the mutational target (Fig. 7), the target mutability (highly mutable regions include CpG rich areas for point mutations and segmental duplications for CNVs) and paternal age at conception (point mutations and CNVs caused by replication-based mechanisms increase with paternal age) (7).

With the new sequencing techniques, the estimation of *de novo* mutation frequencies per generation predicts 74 single nucleotide variants, 3 small insertions or deletions and 0.02 CNVs per genome and one *de novo* mutation per exome (7). Overall mutation rates vary between and within families and are influenced by parental sex, age, ethnicity and predisposing genomic characteristics such as inversions, duplications, translocations and mutations in genes involved in DNA repair or recombination (7). Varying incidence rates for different recurrent genomic rearrangements mediated by NAHR can vary between ethnic groups because of ethnic-specific predisposing structural characteristics such as number of segmental duplications and their orientation in a specific region (7).

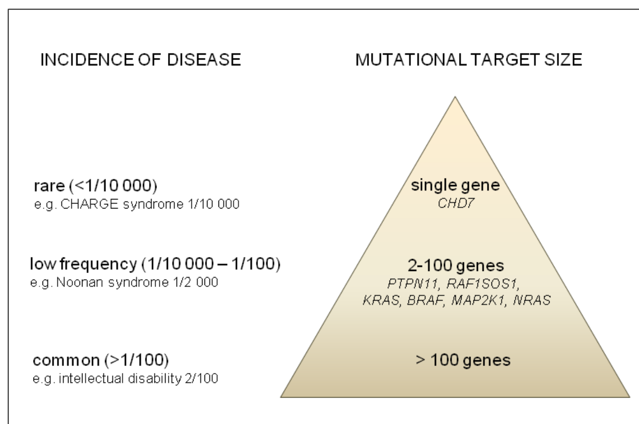

Figure 7. Impact of mutational target size on frequency of disorders caused by *de novo* mutations. Estimation of the number of disease-associated genes for phenotypes of varying incidence, exemplified by CHARGE syndrome, Noonan syndrome and intellectual disability. Modified from Veltman and Brunner, *Nat Rev Genet* 2012 (7).

1.1.8.2 Environmental factors

Embryological or fetal exposure to some environmental factors during sensitive time points in development can disrupt normal morphogenesis and result in secondary defects. Examples of teratogens, for which exposure during pregnancy increases the risk for malformations include maternal disease, drugs, radiation and industrial or agricultural chemicals.

Maternal infection

Several maternal infections are known to disturb embryonic and fetal development and cause abnormalities. Primary rubella infection in a non-vaccinated mother during the embryonic period (≤ 8 weeks) leads to malformations in 80% of cases and primarily causes heart defects, eye defects (congenital cataracts or retinopathy) and ear defects. Infection after 20 weeks of gestation is not associated with birth defects (41).
Infection with cytomegalovirus (CMV), either primary or secondary, may cause a fetal infection that can lead to microcephaly, sensorineural hearing impairment and visual impairment due to chorioretinitis, optic atrophy or damage to the occipital cortex (42). Toxoplasma infection may also cause hearing and visual impairments in the fetus (43).

Maternal chronic disease

It is known that the risk for birth defects in pregnancies where the mothers have diabetes mellitus (including type I, II and gestational diabetes) is increased, especially the risk for heart malformations and neural tube defects. It has been proposed that increased glucose levels may affect the establishment of epigenetic patterns, leading to changed gene expression and abnormalities in the embryo (44).
Other chronic diseases that seem to be associated with malformations independently of treatment are obesity (in absence of diabetes), migraine and paroxysmal supraventricular tachycardia (44, 45).

Drug treatment

Drugs are known teratogens that are believed to cause malformations through a number of disruptive mechanisms e.g. neural crest, endocrine or vascular disruptions, folate antagonism, oxidative stress or specific receptor- or enzyme-mediated teratogenesis (46). Different types

of drugs in clinical or recreational use are associated with fetal abnormalities, primarily cardiac abnormalities, for example chemotherapeutic agents (methotrexate), anticonvulsants (valproate), mood stabilizers (lithium), antibiotics (erythromycin), anticoagulants (warfarin) and retinoids (isotretionin) (35, 47-49). The knowledge of some such associations is used to choose alternative treatments if possible, although in severe cases the health of the pregnant woman is prioritized. Regarding recreational drugs, alcohol is a known teratogen and associations between maternal smoking and several different birth defects have been reported (50).

Associations between drugs and birth defects are difficult to establish due to a number of inherent factors. Malformations are rare events, and associations are therefore most efficiently studies by case-control studies, which may introduce selection and exposure information may be subject to observation bias. Also, many factors may affect outcomes, for example multiple drug treatment, timing of fetal exposure, dosage level and underlying genetic predisposition. In addition, the disease for which the pregnant woman is treated may in itself increase the risk for fetal abnormalities.

Assisted reproductive technologies (ART)

Conflicting evidence has been presented as to whether assisted reproduction techniques using in vitro fertilization (IVF) increase the risk for birth defects (51). It is proposed that the observed slightly elevated risk for major malformations seen in ARF pregnancies that has been presented in recent studies (3-4% as opposed to 2-3% for a normal population) may be explained by an inherent higher risk for birth defects in couples receiving ART treatment, and not related to the procedure (51).

Radiation

Exposure of the fetus to high doses of radiation increases risk for microcephaly and other birth defects. However, fetal anomalies are only associated with exposure to more than 50 mGy, which is a dose markedly higher than the radiation exposure for the fetus at singular examinations using radiography, computed tomography or nuclear medicine (52).

Industrial and agricultural chemicals

An example of a teratogenic agricultural chemical is methyl mercure, a metal compound used to preserve seed grain, and methyl mercury poisoning in pregnant women is known to increase the risk for birth defects (35).

1.1.8.3 Multifactorial etiology

A multifactorial model of inheritance, in which genetic and environmental factors add up to a threshold over which morphogenesis results in a malformation, has been proposed for many isolated congenital malformations, on the basis of statistical evidence (53, 54). This mechanism is proposed to explain the occurrence of phenotype in families with non-Mendelian inheritance and sporadic disease. Still it is possible that di-, oligo- or polygenic

causes can explain familial occurrence, and sporadic disease may be caused by *de novo* mutations.

1.2 COPY NUMBER VARIATION

1.2.1 *Definition*

Structural variation in the human genome includes balanced (inversions and balanced translocations) and unbalanced forms (deletions, duplications and insertions) in terms of copy number, which refers to the number of times a specific genomic region occurs in a genome (55). Unbalanced forms, copy number variants (CNVs), were initially defined as stretches of DNA more than 1000 bp (1 kb) in size that differ in copy number in relation to a reference genome (56), but since refined methods, including high-throughput paired-end sequencing, have shown copy number variants of smaller sizes, a later definition refers to CNVs as ≥ 50 bp in size (55).
Copy number variation was first described in 2004, when technologies allowing comparison of whole genome copy number patterns between different genomes emerged. Two different studies described that segments of DNA, in size between small variants affecting base pairs and large variants such as heteromorphisms, were scattered over large regions of the human genome and contributed to genetic variation (57, 58). It is believed that CNVs cover 5-12% of the human genome and they probably account for 0.5-1% of the genetic variation between two individuals (59, 60). In an evolutionary perspective, it is likely that copy numbers with genetic content have inferred evolutionary benefits that have lead to enrichment in the human genome. Different types of CNVs are shown in figure 8.

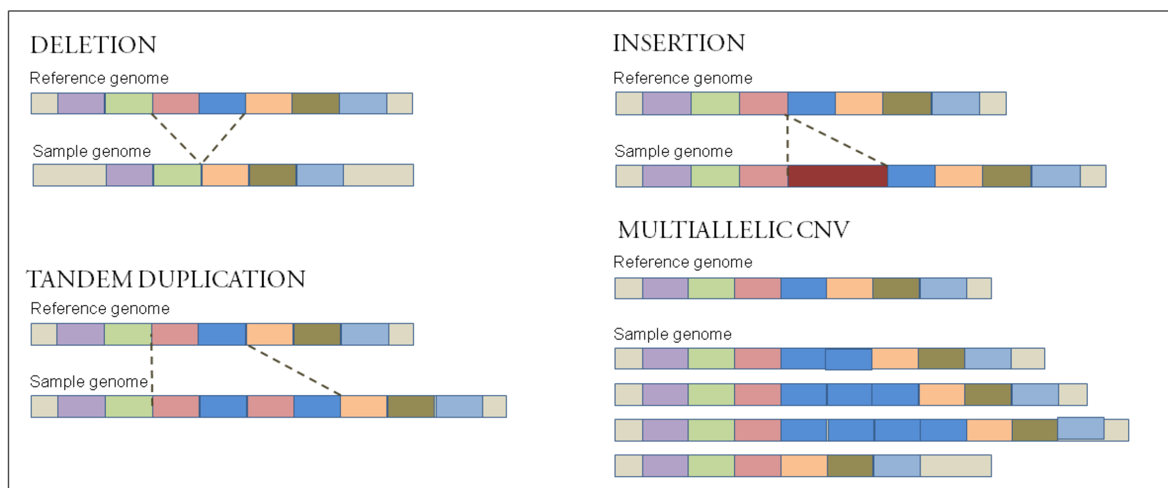

Figure 8. Forms of copy number variation. Figure showing different forms of copy number variation. Duplicated segments can be localized adjacent to the original region (in tandem) or be localized at a distant region of the genome (insertion). Multiallelic CNVs exist in several different alleles with varying copy numbers. Modified from Lee and Scherer, *Expert Rev Mol Med* 2010 (40).

1.2.2 *Phenotypic influence of CNVs*

CNVs are believed to affect phenotypes by changing gene expression, and studies have shown that expression of genes within and in the vicinity of CNVs are affected by copy

number variants (61). More than 50% of the effects of known CNVs are caused by gene disruption or disruption of the regulatory units associated with genes, rather than by gene dosage change (40). Changing the regulatory landscapes of genes can lead to a variety of effects since changes can be tissue- or developmental stage-specific (55). Interestingly, it seems that changed mRNA levels correspond poorly with protein levels, due to the several levels of regulation that determine protein levels (55). In general, it seems that deletions have more severe effect on human phenotypes than duplications, since 80% of CNVs reported in the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER, https://decipher.sanger.ac.uk/) are deletions and that the duplications listed are correlated to less severe symptoms (62).

Recurrent microdeletions and microduplications were first identified as causes of developmental disease, and rare CNVs are also known to be responsible for disease phenotypes (63-65). Lately, CNVs have been implicated in complex diseases, for example Parkinson disease, Alzheimer disease, schizophrenia, HIV susceptibility and Crohn’s disease (63).

Importantly, the term CNV is a structurally descriptive term that implies presence of varying copy number for a specific region, but does not confer information about clinical significance. In order to specify clinical impact, CNVs should be defined as benign or pathogenic (66).

1.2.3 Mechanisms behind formation of CNVs

Genome-wide studies have estimated that *de novo* CNVs > 100 kb occur in 1/50 individuals (67). Different mechanisms for formation of CNVs have been described, with different types of mechanisms responsible for recurrent and rare CNVs.

Non-allelic homologous recombination (NAHR)

This mechanism involves recombination between non-homologous sequences of high similarity present in the genome, mostly segmental duplications or low-copy repeats which are > 10 kb in length and show 95-97% sequence similarity and therefore confer risk for erroneous pairing during recombination. Repeats oriented in the same direction mediate formation of duplications and deletions of the interval between the repeats, while repeats oriented in opposite directions mediate formation of inversions. NAHR can occur during meiosis and mitosis and lead to constitutional or somatic, often recurrent, CNVs (63).

Non-homologous end-joining (NHEJ)

This mechanism is used to repair double-stranded breaks in chromosomes caused by environmental factors, and physiologically in B and T cells for creation of receptor and immunoglobulin diversity. CNVs resulting from NHEJ are non-recurrent and often found in shorter repeat sequences such as LINEs, *Alu* repeats and LTRs. A molecular “scar” with

inserted nucleotides can often be identified. This mechanism can cause deletions and translocations, but not duplications.

Fork Stalling and Template Switching (FoSTeS)/Microhomology-mediated break-induced replication (MMBIR)

The FoSTeS mechanism is active during DNA replication and involves fork stalling, disengagement of the lagging strand and association with a new replication fork, induced by microhomology sequences. This process may be repeated several times which can lead to complex rearrangements, and is responsible for deletions, duplications and inversions. Another replication-based mechanism that involves template-switching is MMBIR which is activated by breakage of a single DNA strand at the replication fork (63).

1.3 VACTERL ASSOCIATION

VACTERL association is a congenital heterogeneous condition with multiple malformations affecting different organ systems in affected individuals, and has an estimated incidence of 1/10 000-1/40 000 in different studies (68). The component features of VACTERL are vertebral defects (V), anal atresia (A), cardiac malformations (C), tracheo-esophageal fistula (T) with esophageal atresia (E), renal malformations (R) and limb defects (L). The condition was first described by Quan and Smith more than 40 years ago (69), and was then termed VATER and included vertebral, anorectal, tracheoesophageal and radial anomalies while cardiac and renal anomalies were later added. Neurocognitive disability is not considered part of the phenotypic spectrum, although a specific form of VACTERL, VACTERL-H, includes hydrocephalus due to aqueductal stenosis (OMIM:#276950) (68). VACTERL association usually occurs sporadically, although familial cases exist (70).
Diagnostic criteria, which are debated, include malformations in three of the organ systems, on the prerequisite that other malformation syndromes are not more likely causes. Due to the unknown cause, the heterogeneous phenotype and the phenotypic overlap with many other malformations syndromes, it is considered a diagnosis of exclusion (71). The clinical picture varies with the severity of the malformations seen in different individuals, and can differ from subclinical to life-threatening. Due to improvements in health care during the last decades, prognosis is rather good with present surgical techniques and intensive care, although sequelae cause substantial morbidity.

1.3.1 Causes of VACTERL association

There is no known, unifying cause of VACTERL association. It is generally believed to be a causally heterogeneous condition, although it is still possible that a single cause may explain the majority of cases, as in the case with *CHD7* in CHARGE syndrome. Genetic, environmental and multifactorial causes have been proposed as causative.

1.3.1.1 Genetic factors

Evidence for one or more genetic causes includes (14):

1) an increased incidence of one or more component features in first-degree relatives of patients with VACTERL association (72) which implies genetic components in at least a subset of cases
2) second-degree relatives show component features of VACTERL association
3) several chromosomal aberrations have been reported in single patients (70)
4) mutations in single genes have been described in individual patients
5) overlap between the phenotype seen in VACTERL association and several other malformation syndromes caused by single gene defects
6) VACTERL-like phenotypes are observed in mouse studies with mutations in genes (*Shh*, *Gli2*, *Gli3*, *Pcsk5*) involved in the embryologically important sonic hedgehog (*SHH*) signalling pathway (73, 74)

It is worth noting that despite overlap with other syndromes, there is no single syndrome that encompasses all VACTERL component features except for VACTERL-H, which in some cases is caused by mutations affecting the *FANC* genes and has once been reported in patient with a heterozygous mutation in the *PTEN* gene (75, 76). A mutation in the *ZIC3* gene, normally associated with heterotaxy, has also been reported in a patient with all component features of VACTERL association (14).

Causative genetic factors that are compatible with findings of low monozygotic concordance in twins and sporadic occurrence are for example single gene defects with reduced penetrance or variable expressivity, epigenetic/regulatory factors, *de novo* mutations, mosaic occurrence or di-/oligogenic inheritance.

1.3.1.2 Environmental factors

One hypothesis that explains the occurrence of malformations originating in different time windows during development (which is the case in VACTERL association) is that the embryo is chronically exposed to a teratogen during this time. It is known that type I, II or gestational diabetes in pregnant women increases the risk for vertebral, cardiac and limb component features of VACTERL association in the fetus (77). This is suggested to be caused by a combination of hyperglycemia, oxidative stress and mitochondrial dysfunction (14).

Teratogens that can cause single component features in VACTERL association include anticonvulsants, retinoic acid antagonists and alcohol. Other suggested teratogenic causes are for example infertility treatment and contraceptive use in pregnant mothers, but up to date, the evidence is weak (14).

1.3.1.3 *Developmental pathways*

Disturbance in developmental processes that affect formation of all organs involved in VACTERL association, for example disturbance in mesodermal proliferation and migration, epithelial-to-mesenchymal transition or apoptosis have been proposed. Signalling pathways that have been proposed as causative are the SHH, NOTCH and FGF pathways (14, 78).

1.4 MOSAICISM AND CHIMERISM

Historically, all cells of an individual have been considered to carry the same genetic content. Mosaicism and chimerism represent different variants of a phenomenon in which cells of an individual differ in genetic content, with presence of two or more genetically distinct cell populations. Mosaicism is known to occur physiologically in development of lymphocytes where somatic rearrangements in individual B- and T-cells are responsible for the diversity of T-cell receptors and immunoglobulins, fundamental for the adaptive immune system. Mosaicism and chimerism are also seen in some genetic diseases and unusual phenotypes.

1.4.1 Definitions

Mosaicism refers to the presence of cell populations differing in genetic content, which arise after fertilization (post-zygotically) involving one fertilized egg (Fig. 9). Chimerism is defined by the fusion of two fertilized eggs into one embryo (Fig. 9).

1.4.2 Forms of mosaicism

Mosaicism occurs as the result of a genetic abnormality being introduced after fertilization, and since this can happen at any time point in development or postnatal life, the proportions and distributions of affected cells vary accordingly. A mutation that occurs in one of the first cell divisions after fertilization results in the mutation being present in a large proportion of cells of the growing embryo, and may affect all or several different tissues. A mutation introduced at a later stage of differentiation may affect fewer, more differentiated cells, for example skin or brain tissue. This type of mosaicism is termed somatic mosaicism, while mutations occurring only in germ cells are termed germline mosaicism. Germline mosaicism is a phenomenon that explains the occurrence of multiple offspring affected by the same genetic condition when parents are not carriers of the mutation. Tissue-specific mosaicism, with restriction of CNVs to different tissues, has been reported in apparently healthy individuals (79).

The introduction of assisted reproduction techniques, have created possibilities to study human embryos at a cellular and molecular level. Studies using single-cell array CGH analysis in blastomeres from pre-implantation embryos have revealed a high degree of mosaic chromosomal and structural variation at this stage (80, 81), although it seems that a proportion of genetically abnormal embryos undergo "self-correction" during the first seven days after fertilization.

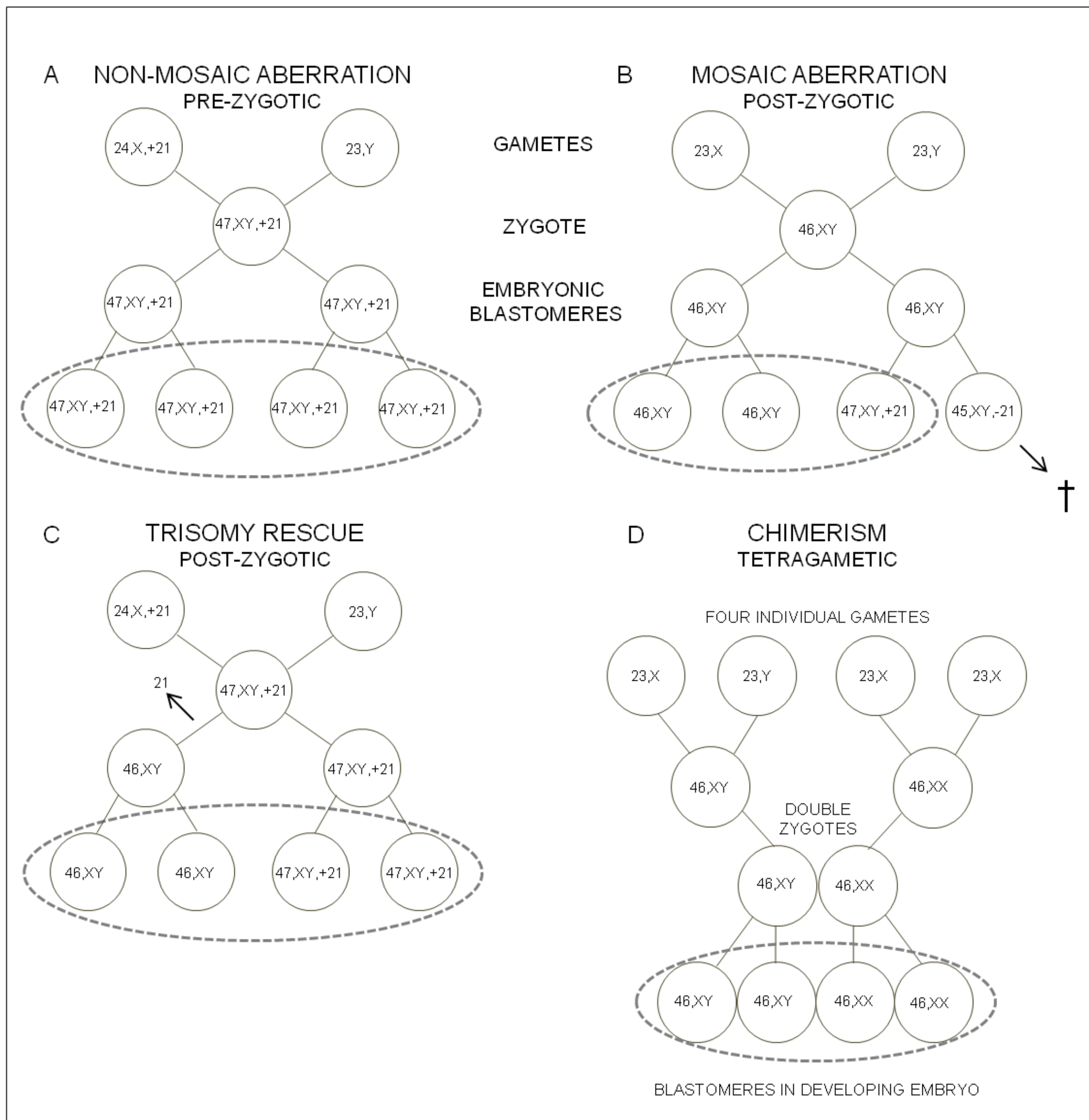

Figure 9. Mechanisms of formation of chromosomal mosaicism and chimerism. Figure A-C illustrating mechanisms of formation in mosaic and non-mosaic forms of aneuploidy, exemplified by trisomy 21. Non-mosaic aberrations are formed by pre-zygotic events (A), while mosaic aberrations are formed by postzygotic introduction of an aberration (B). Dagger marks extinction of the monosomic cell population, resulting in two genetically distinct cell populations in the embryo. In some cases, a pre-zygotic aberration is post-zygotically corrected in a population of cells, so called trisomy rescue (C). Figure D showing formation of tetragametic chimerism by fertilization of two oocytes by two separate spermatozoa, and subsequent fusion of the double zygotes into one embryo. Fertilization by stwo permatozoa carrying X and Y chromosomes, respectively, result in cell populations with discordant sex chromosome complement, which may result in disorders of of sex development.

One theory explaining this observation is that cell cycle control mechanisms are not fully activated during the first cell divisions after fertilization, and that after activation of these mechanisms, abnormal cells are mitotically arrested (82). It is also known that self-correction, or “rescue”, can occur through loss of abnormal chromosomes, especially for trisomies and possibly triploid conceptions, sometimes resulting in uniparental disomy (Fig. 9) (83, 84). Another form of rescue is revertant mosaicism, when primary mutations can be self-corrected in individual cells, for example in Fanconi anemia and some cutaneous diseases (76, 85).

Confined placental mosaicism is a specific type of mosaicism originating in the early separation of cells carrying a genetic aberration from normal cells in the embryonic tissues, so that the genetically abnormal cells become restricted to placental tissues. This may affect fetal growth by placental dysfunction and is also clinically important in chorionic villus sampling for prenatal diagnostics, since analysis may show genetic abnormalities not present in fetal tissues (86).

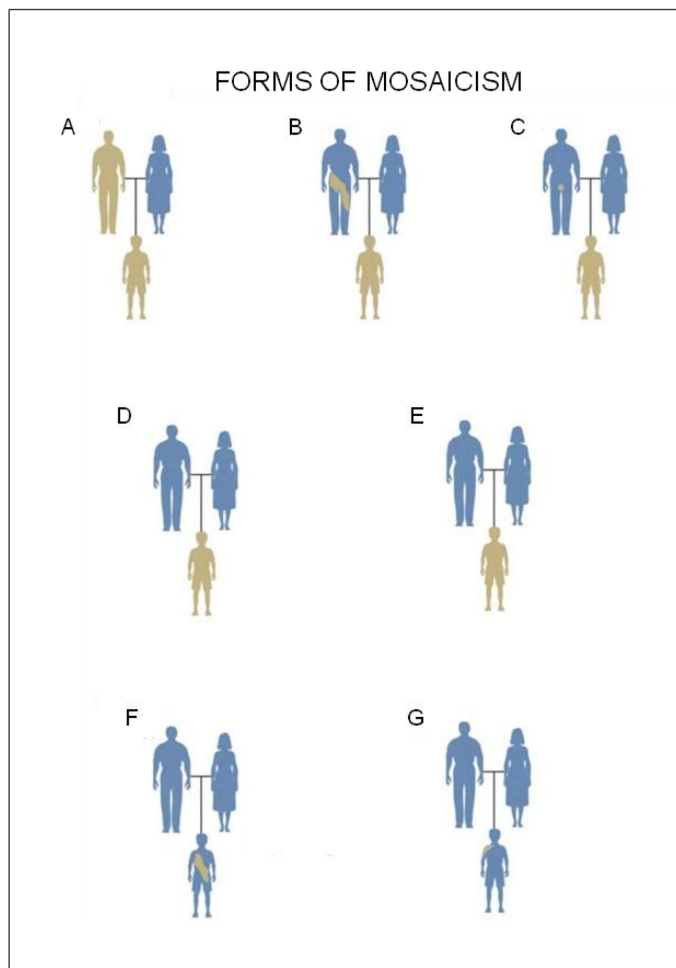

Figure 10. Forms of mosaicism. Figure showing different forms and origins of mosaicism. A) shows inheritance of mutation present in all paternal cells (non-mosaic mutation). B-C show inheritance of mutations from B) father with mosaic mutation present in several tissues including germ cells, and C) father with mutation present only in germ cells. D-E show two clinically indistinguishable variants of *de novo* mutations, with mutation arising in a single germ cell (D) or formed in one of the first cell divisions of the embryo (E). In A-E, the proband is heterozygous for the mutation in all cells. F-G show somatic mosaic mutations in the proband. In F), the mutation has arisen early in post-zygotic development, while in G), the mutation has occurred later in development and is restricted to a certain cell type or tissue. Modified from Freed et al, *Genes (Basel)* 2014 (89).

It has recently been shown, using breakpoint-specific sequencing in parents of probands with apparent *de novo* CNVs, that in a subset of these families the parents have low-level mosaicism for the same CNVs (87).This finding implies a higher recurrence risk than what would be estimated if a *de novo* mutation was present in a single parental germ cell or formed in the first cells of the embryo, and the recurrence risk depends on the frequency of the mutation in the germline. If a mutation confers a surviving advantage to the cell, the result can be an expansion of a mutation-carrying clone in the germline, and an increased recurrence risk, possibly as high as a dominant condition (87). Also, low-level parental mosaicism for single gene defects are being detected with high-resolution techniques (88).

1.4.3 Mosaicism in disease

The implication of genetic mosaicism for disease phenotypes is difficult to predict in many cases, since it is dependent upon timing of the mutation event, cell type/types affected and distribution of genetically aberrant cells. However, mosaic forms of disease are generally associated with milder phenotypes compared to non-mosaic aberrations.

The most commonly described form of mosaicism comprises chromosomal aberrations, since cytogenetic techniques have enabled detection of these aberrations for a long time. Mosaic forms of aneuploidies such as trisomy 13, 18 and 21 usually have less severe phenotypic consequences, and mosaic forms of other aneuploidies not seen in constitutional form such as trisomy 8, 9, 14, 17 and 22 are found. Mosaic forms of other structural variants such as translocations, inversions, ring chromosomes and supernumerary chromosomes are not as commonly reported (86). Several mosaic monogenic disorders have been reported during recent years often affecting the skin, skeletal and vascular systems and some occurring only in mosaic form, such as McCune-Albright, CLOVES and Proteus syndromes (89). Mosaicism is also reported in neurodegenerative diseases, normal aging and cancer (89).

1.4.4 Chimerism

Whole-body chimerism occurs as the result of a fertilization error when two separate zygotes give rise to one embryo (Fig. 9). Other types of chimerism include blood chimerism, when blood cells are exchanged between twins, or between mother and fetus and as a result of blood transfusion. Chimerism formed from two embryos with the same sex chromosome content is not known to cause a phenotype, and may rarely be discovered when blood group testing shows evidence of more than one blood group. When two embryos of different sex chromosome content form one embryo, genital development may be affected and result in ambiguous external genitalia which will lead to clinical genetic investigations. The phenotype varies in severity from normal male or female phenotype to different degrees of ambiguous genitalia, and is often associated with sterility. Chimerism explains a proportion of individuals with ovotesticular DSD, when both ovarian and testicular tissues are present in the same individual. Chimerism is proposed to form through different mechanisms; tetragametic chimerism, parthenogenetic chimerism, androgenetic chimerism and fertilization of the second polar body. In tetragametic chimerism two separate zygotes, formed from four gametes with equal quantitative contributions from the mother and the father, fuse. Parthenogenetic and androgenetic chimerism involve unequal quantitative contributions from mother and father; in parthenogenetic chimerism three of the four genomic contributions are of maternal origin while the corresponding proportions have paternal origin in androgenetic chimera. Fertilization of the second polar body has not been shown, but is suggested as an additional mechanism to give rise to chimeric embryos.

2 AIMS

2.1 OVERALL AIMS

The overall aims of this thesis were to improve clinical genetic investigations for patients with congenital malformations, and at the same time identify genes important in normal development of organs. Due to the limited knowledge of genetic factors in the development of malformations, clinical genetic counselling for couples with a previous affected child is unsatisfactory for families and doctors. An increase in patients with malformations who survive and reproduce, due to improvements in surgical and intensive care possibilities, also motivates increased research so that better prenatal diagnostics and more accurate estimations of recurrence risk can be given to these individuals.

2.2 SPECIFIC AIMS

More specifically we wanted to investigate the role of copy number variation in development of congenital malformations, and this was attempted by analyzing copy number variations in patients with different malformations.

2.2.1 Investigation of copy number variants using array CGH in patients with congenital malformations

We wanted to systematically investigate copy number variants in patients with congenital malformations by analysis with array CGH. Our intent was to investigate if pathogenic copy number variants were common causes in patients with malformations and also to identify candidate genes important in development of specific organs.

2.2.2 Mosaicism

It is known that mosaicism for chromosomal abnormalities often result in milder and more varied clinical phenotypes compared to non-mosaic cases. Until the introduction of array CGH, the possibilities to investigate smaller aberrations and different tissues have been limited due to requirement for technologies with higher resolution and need for tissue culture. We wanted to use array CGH analysis in malformed tissue to investigate if genetic mosaicism in general, and specifically tissue-restricted mosaicism in malformed tissue, was an overlooked cause of congenital malformations due to the difficulty of detection. To investigate this possibility we wanted to collect malformed tissue and blood samples from individuals with congenital malformations to compare the presence of pathogenic CNVs between healthy and malformed tissue.

2.2.3 VACTERL association

Chromosomal aberrations and single gene defects have been described in individual patients with VACTERL association, but no common cause has been identified in this condition. A specific aim was to investigate if copy number variation represented a common cause in individuals with VACTERL association, since there is strong belief that this condition with multiple malformations has a genetic background. In addition, we wanted to screen for

mutations in three specific genes (*PCSK5*, *HOXD13*, *CHD7*) in which mutations had been reported in single studies shortly before the start of this study (*PCSK5* and *HOXD13*) or where the associate phenotype was overlapping with VACTERL, but where the extent of mutations found in patients with VACTERL association was not known (*CHD7*).

3 MATERIALS AND METHODS

3.1 PATIENTS

3.1.1 Study I and II

The two patients described in studies I and II were identified and chosen for research analysis after clinical genetic investigations were performed at the Department of Clinical Genetics, Karolinska University Hospital, Sweden.

3.1.2 Study III

Patients with VACTERL association in study III were identified through the Swedish VACTERL Society and from the Pediatric Surgery Department at Astrid Lindgren Children's Hospital, Stockholm, Sweden. Fetal cases fulfilling VACTERL criteria were identified from the registries at the section for Perinatal pathology unit at the Karolinska University Hospital in Huddinge, Stockholm, Sweden. Patients were included if they fulfilled diagnostic criteria for VACTERL association (≥ 3 component features).

A subgroup of patients with VACTERL-like phenotypes was identified from a previously established biobank including DNA from patients with esophageal atresia. Patients with esophageal atresia and a minimum of one more component feature of VACTERL association were included.

3.1.3 Study IV

Patients surgically treated for congenital heart malformations were identified and included from the two Swedish centers performing pediatric cardiac surgery; the section for Pediatric Cardiology, Queen Silvia Children's Hospital in Gothenburg and the Pediatric Cardiac Surgical Unit, Children's Hospital at the University Hospital in Lund. The only inclusion criterion was that tissue removal was a planned part of the surgical procedure. Patients with isolated as well as syndromic phenotypes were included. In total, we received samples from 33 patients. Ten patients were excluded for different reasons: three patients were excluded due to a known syndromic diagnosis (two patients with Down syndrome and one patient with CHARGE syndrome), six patients were excluded due to insufficient or inadequate quality of material for analysis, and one patient was excluded after analysis when it was revealed that he had a brother with the same type of malformation (familial case). Altogether, 23 patients were included in the study. One of the patients in this study was also reported in study III (V11 in study III/P7 in study IV, array CGH analysis was performed once using DNA from heart tissue).

3.1.4 Study V

Patients surgically treated for congenital malformations were identified at the Department of Pediatric Surgery at Astrid Lindgren Children's Hospital, Stockholm, Sweden. We used the same inclusion and exclusion criteria as in study IV, that patients were included if tissue

removal was a planned part of the surgical procedure and that patients with known causative diagnoses were excluded, respectively. In total we collected 54 samples from patients with varying diagnoses. We received samples from patients with the following diagnoses: anorectal malformations (19), esophageal atresia (12), urological malformations (13), biliary atresia (3), vascular malformations (3), gastroschisis (1), omphalocele (1), diaphragmatic hernia (1) and ovotesticular DSD (1).

We excluded thirteen patients for whom written consent could not be obtained, five patients that had normal array CGH analysis results in previous clinical investigations, one patient due to insufficient or inadequate quality of material for analysis and one patient due to previously known diagnosis (Down syndrome). After excluding these patients, nine patients were not prioritized for analysis, due to low numbers of patients in each group (diaphragmatic hernia, vascular malformation, polycystic kidneys, stenosis of the ureter, kidney duplication and biliary atresia). Altogether, 25 patients were included in the study.

3.1.5 Informed consent

All patients were included after informed consent was given from parents. For fetal cases, these were included from a biobank to which parents had given permission for research purposes.

3.2 TISSUE SAMPLES

Tissue samples from patients with congenital malformations were collected at surgery so that tissue removed from the malformed organ as part of the surgical procedure was preserved in physiologic saline solution (9 g/L).

Tissue samples from fetal cases were collected from the biobank of fresh frozen tissue samples routinely collected during autopsy proceedings at the Section for Perinatal Pathology at Karolinska University Hospital in Huddinge.

3.3 DNA EXTRACTION

For DNA extraction, tissue preserved in saline solution was placed on a cell culture dish and a piece of desired size (20-40 mg) was cut and weighed. The sample was placed in Cell Lysis Solution from the Gentra Puregene Blood Kit (QIAGEN Sciences, Maryland, USA) followed by incubation at 65º for 45 minutes and subsequently a 10 minute incubation in a 56º heating block after which Proteinase K was added and the sample vortexed vigorously and left over night in 56º, after which the regular protocol was followed.

Genomic DNA was isolated from peripheral blood samples according to standard procedures using the Gentra Puregene Blood Kit (QIAGEN Sciences, Maryland, USA).

3.4 DNA SEQUENCING (SANGER METHOD)

The Sanger method for determining a target sequence is a polymerase chain reaction (PCR)-based method that works by stepwise prolongation of a specific DNA target sequence, using target-specific primers and a thermostable DNA polymerase in a mixture of regular deoxyribonucleotides and fluorescently labelled dideoxyribonucleotides. The latter molecules do not have the hydroxyl group of the carbon at position 3' of the deoxyribose sugar ring, interrupting further prolongation of the chain. The result of the reaction is a multitude of fragments of all possible different lengths, with a ddNTP located at the last position. The DNA sequence is determined through size-based separation of the fragments and identification through detection of wavelengths of emitted fluorescence.

Identified variants are compared to variants reported in databases with sequence data from healthy individuals and characterized according to the effect of the mutation on protein amino acid sequence (synonymous/non-synonymous, missense, nonsense or frameshift mutation) and splicing. Other factors that influence assessment of a detected single nucleotide variant are the degree of conservation of the nucleotide across species and location in functionally important sites or domains of the protein.

DNA sequencing was used in study III to screen the *PCSK5*, *CHD7* and *HOXD13* genes for mutations in patients and fetal cases with VACTERL association or a VACTERL-like phenotype. Primers for *PCSK5* screening were designed for the 39 exons found in the Ensembl genomic sequence ENSG00000099139, while primers designed at the Department of Clinical Genetics, Karolinska University Hospital, were used in sequencing of *HOXD13* (exons 1 and 2) and *CHD7* (exons 2-38).

Assessment of detected variants was performed with comparison to single nucleotide variants listed in dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) and by using pathogenicity prediction algorithms such as the Alamut v 2.0 software (Interactive Biosoftware, http://www.interactivebiosoftware.com). Alignment of protein sequences for PCSK5 was performed using the Uni-ProtKB database (http://www.uniprot.org/) and included sequences from six vertebrates.

3.5 MICROSATELLITE MARKER ANALYSIS AND QUANTITATIVE FLUORESCENT PCR (QF-PCR)

The possibility to effectively distinguish alleles and homologous chromosomes, and trace them within pedigrees is enabled by the existence of close to seven hundred thousands of short tandem repeats, or microsatellites, dispersed throughout the human genome (90). The repeats consist of a varying number of contiguous copies of 1-6 bp-units. Due to inexact copying of these sequences, there is a high degree of repeat length polymorphisms, with around 70% of individuals having varying number of copies in a specific location. To be able to distinguish alleles or homologous chromosomes, analyzed markers need to be informative, and that probability increases with the proportion of the population that is heterozygous for specific markers. The analysis is carried out in a PCR-reaction using one specific fluorescent

primer so that multiple single strand copies of the repeat are created. The resulting copies are then separated on an automatic sequencer together with a size standard reference so that the lengths of the fragments can be determined. The reaction is semi-quantitative, meaning that the relative amounts of different alleles can be determined.

QF-PCR is a specific application of microsatellite marker analysis used in clinical settings to identify aneuploidies in chromosomes 13, 18, 21, X and Y. Compared to chromosome analysis this PCR-based method is fast (days compared to weeks) but gives no structural information.

Analysis of microsatellite markers and QF-PCR was used to determine the causative mechanism behind the different cell populations detected in the patient reported in study I and also to determine the parental origin of and mechanism behind the formation of the marker chromosome found in the patient described in study II.

3.6 MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION (MLPA)

Multiplex Ligation-dependent Probe Amplification is a multiplex PCR-based method used for detection of copy number changes in specific chromosomal locations. It can also be used for additional applications, for example methylation detection (MS-MLPA). Shortly, MLPA starts with hybridization of two adjacent target-specific oligonucleotide probes to patient DNA. Successful exponential amplification of the probes is dependent on specific binding of both probes to their complementary DNA sequences, so that a subsequent ligation step results in the formation of one probe for each target sequence. The designed probes all carry identical forward and reverse primer-binding sequences and therefore can be exponentially amplified in a subsequent multiplex PCR reaction using one primer pair, on requirement that the ligation reaction is successful. PCR-amplified fragments are separated on capillary electrophoresis according to size and relative amount detected through fluorescent intensity. Intra- and intersample normalization using reference probes and reference samples, so that the copy number results measure relative amounts of the target sequences.

MLPA can be used for confirmation of copy number changes detected with other methods or, because of the ease of use, to screen large numbers of patients for specific copy number changes.

In our studies, MLPA was used for confirmation of small copy number variants detected by array CGH in study III.

3.7 CHROMOSOME ANALYSIS

Analysis of chromosome structure is accomplished by culturing of cells, chemical arrest of the cell cycle in metaphase, fixation of chromosomes on glass slides and chromatin staining followed by microscopy. With microscopy analysis of metaphase spreads, whole chromosome loss or gain (aneuploidy), extra structural abnormal chromosomes such as isochromosomes and ring chromosomes, as well as structural chromosomal balanced or

unbalanced rearrangements such as deletions, duplications, translocations, inversions and insertions can be detected at a resolution limited to around 4-5 Mb depending on the type of abnormality and location in the genome. Due to heteromorphisms, interchromosomal variations in amount and arrangement of repetitive DNA, chromosomal homologues can sometimes be distinguished in chromosome analysis. The method provides structural information, provides possibility to analyze single cells and identifies balanced rearrangements. The same information is not obtained from more modern methods, such as MLPA and array CGH analysis, and chromosomal analysis will likely continue to be an important clinical genetic method in the future. An important drawback with this method is that the requirement for dividing cells and cell culture (which may take up to 3 weeks) renders the method time consuming, and in addition the resolution is low.

Chromosome analysis was used in individual cases in study I-V to determine the arrangement of chromosomal aberrations and to distinguish different cell populations.

3.8 FLUORESCENT IN SITU HYBRIDIZATION (FISH)

Hybridization of fluorescently labelled DNA probes of target-complementary sequence to denatured meta- or interphase spreads allows the detection of loss or gain of DNA at a higher resolution than chromosome analysis. Different types of probes can be used for specific purposes such as single copy probes, dual fusion probes, break-apart probes, chromosome enumeration probes (CEP) and painting probes. Analysis can be performed on metaphase or interphase spreads, with different advantages and drawbacks. Metaphase FISH allows a more accurate structural analysis but requires living cells that undergo cell culture to reach the metaphase stage. Interphase FISH is convenient to overcome the requirement for cell culture and is suitable for analysis on tissues difficult to culture, and is also convenient for analysis of a high number of cells in analysis of mosaicism. The resolution of FISH analysis is higher for deletions than for duplications. The form of FISH analysis with the highest resolution is called fibre FISH and involves denaturation of the protein structures that keep chromatin wound together, resulting in a free chromatin chain where targets down to around 5 kb can be analyzed. The obvious drawback of FISH analysis is that it is a specific analysis where the target analyzed must be chosen beforehand and FISH can thus not be used as a screening method.

FISH analysis was used in studies I, II, IV and V to determine the origin of additional genetic material identified by chromosome analysis, to determine rearrangements and to confirm aberrations detected by array CGH.

3.9 ARRAY COMPARATIVE GENOMIC HYBRIDIZATION (ARRAY CGH)

3.9.1 Principle

Array CGH is a method that combines screening analysis with high resolution, and has since introduced in research and clinical practice become the first choice method in clinical genetic investigations of patients with DD/ID, autism spectrum disorders and multiple congenital

malformations (91). The method is based on an “array” of several hundred thousand up to millions of DNA probes attached to glass slides. The probes can be seen as the equivalent of FISH probes, and together represent the whole genome. By analysis with array CGH, DNA isolated from patients can be investigated for loss or gain of chromosome material, copy number variants, through comparison with DNA isolated from healthy individuals.

3.9.2 Different types of microarray platforms

3.9.2.1 Bacterial artificial chromosome (BAC) array

The first type of microarrays used BAC probes of 200-300 kb, and the resolution for detection of CNVs was approximately the same as the size of the probes. BAC array slides were produced by a method that involved printing spots of the different probes on to the glass slides. The protocols used for BAC arrays involved dye-swap experiments, in which two hybridization reactions were performed per patient sample. If the patient sample was labelled with Cy3 and the reference with Cy5 in the first reaction, the opposite was done in the second reaction to minimize false positive signals.

3.9.2.2 Oligonucleotide array

Microarrays using oligonucleotide probes, which are synthesized directly on the glass slide, have a considerably higher resolution than BAC arrays. Many oligonucleotide platforms can detect aberrations down to a few kb, while the platform with the highest resolution reported could detect aberrations down to 0.5 kb (92).

3.9.2.3 SNP array

SNP arrays combine detection of copy number variants with genotyping, enabling detection of loss of heterozygosity/uniparental disomy in the same experiment. In SNP array hybridization reactions, the patient sample is hybridized to the array. Fluorescence intensities from the experiment are then compared to a reference dataset to detect copy number variation in the patient, and intensities from the different alleles are also compared for each SNP, so that genotype can be determined. SNP arrays generally have a somewhat higher resolution than oligonucleotide arrays, although the distribution of SNPs is not evenly spread out in the genome. However, oligonucleotide probes can be added to SNP arrays to increase coverage. SNP arrays have the possibility to detect mosaicism at a lower level than oligonucleotide platforms.

3.9.2.4 Studies I-V

Array CGH was used in all studies, primarily for CNV screening in studies III, IV and V, but also to fine map aberration breakpoints detected by chromosome analysis in study I and II. Three different array platforms have been used for array analyses described in this thesis. The majority of patients in studies III-V were analyzed using a 180K oligonucleotide platform with 60-mer oligonucleotide probes and whole genome coverage, manufactured at Oxford Gene Technology (OGT).

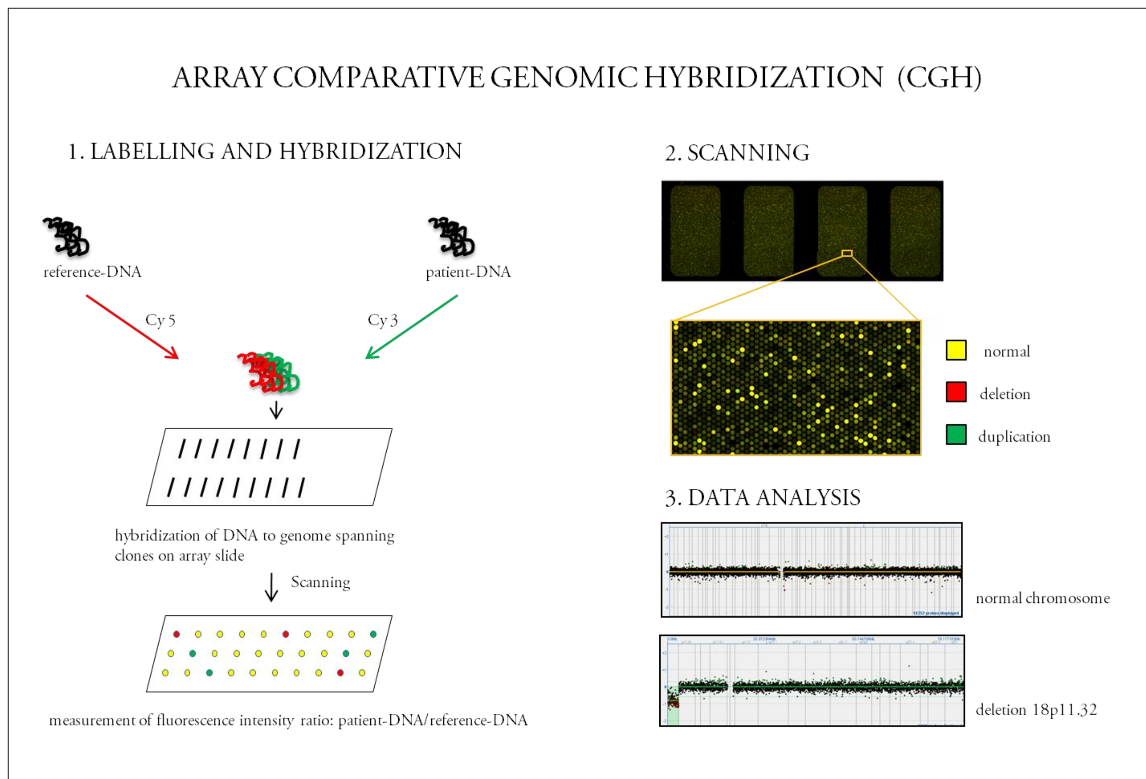

Figure 13. Array CGH analysis. Figure showing the procedure of array CGH analysis. The genome to be investigated (the patient DNA) is separately labelled with one of the two fluorophores Cy3 and Cy5. The reference DNA is labelled with the opposite fluorophore, using nick translation, so that the genomes can be distinguished and compared. In a screening analysis, the reference DNA to which the patient DNA is compared is usually a mixture of DNA pooled from a number of individuals without known disease.
In a hybridizations reaction, denatured and labelled patient and reference DNA samples are added to the array slide, and left to hybridize to the probes present on the glass slide at even temperature for around 24 hours. Unbound probe is removed in a washing step. Results are obtained by measuring fluorescence intensities from all probes using a fluorescence scanner. Computer analysis of raw data involves calculation of the log2-ratios of fluorescence intensity for patient and control samples for each probe, followed by data normalization. CNVs are detected by analysis using software that employs specific CNV detection algorithms.
During the hybridization, the patient and reference DNA binds to probes with complementary sequences in a competitive way, amd if no aberrations are present in a specific region, equal amounts of Cy3 and Cy5 are bound to the corresponding probes, resulting in detection of yellow fluorescence (a log2-ratio of 0 equals two copies in the patient). In contrast, a duplication of a region in the patient DNA results in an overweight of patient DNA bound to the corresponding probes, and detection of green fluorescence (a log2-ratio of $\geq$0.58 equals three or more copies in the patient). Conversely, a deletion results in deficit of the specific region and less patient DNA bound to probes, leading to emission of red fluorescence (a log 2-ratio of $\leq$-1 equals one or no copies (heterozygous or homozygous deletion) in the patient).
Results are visualized for each chromosome separately, with the chromosome position along the x-axis.and the corresponding log2-ratio values of the individual probes along the y-axis.

This platform is used at the Department of Clinical genetics at Karolinska University Hospital, Stockholm, Sweden, and has an in-house design with higher density of probes within genes and a background of evenly spaced probes over the entire genome. Data analysis is performed using the Feature Extraction software for raw data analysis and CytoSure Interpret software for CNV analysis. The practical resolution of the platform is 30-50 kb, varying with probe density in different regions.

The patients in studies I and II have been analyzed by a 244K oligonucleotide array from Agilent Technologies followed by raw data analysis in the Feature Extraction program and CNV detection in the DNA Analytics, both software programs from Agilent Technologies. The resolution of this platform is approximately 50 kb.

Three patients in study III were analyzed with a 38K BAC array containing 38 370 BAC clones and manufactured by the Swegene DNA Microarray Resource Center at Lund University. For analysis of array data, Bio Array Software Environment (BASE) was used for both raw data analysis and CNV detection (93).

3.9.3 Interpretation of CNVs

Interpretation of CNV pathogenicity is a well-known challenge due to the widespread presence of CNVs in the population and their reduced penetration and variable expressivity for associated phenotypes. CNVs can show different types of inheritance; dominant, recessive, X-linked or complex and can also unmask recessive alleles on the other homologous chromosome. Classically, size, gene content, type of CNV, inheritance and presence in databases of healthy or diseased individuals are factors that are assessed when considering pathogenicity of a detected CNV. Databases where variants detected in large studies of healthy individuals, such as the Database of Genomic Variants (DGV) (http://projects.tcag.ca/variation/), or individuals with developmental phenotypes, such as DECIPHER (https://decipher.sanger.ac.uk/) and ECARUCA (http://www.ecaruca.net), have been created to guide in classification.
It is also helpful to compare detected variants with variants from analyzed samples in an in-house database, since detection of some CNVs may be related to platform-specific problems, and also these samples may represent a more relevant genetic background. CNVs where pathogenicity cannot be determined are termed variants of unknown significance (VOUS) (66).

4 RESULTS AND DISCUSSION

4.1 GENETIC CAUSES OF CONGENITAL MALFORMATIONS DETECTED BY ARRAY CGH (STUDIES III, IV AND V)

Screening for pathogenic gene dose alterations in patients and fetal cases with congenital malformations using array CGH (study III-V) identified pathogenic aberrations in six patients (Table 3). These clearly pathogenic aberrations were detected in patients with syndromic phenotypes only, and no pathogenic aberrations were identified in patients with isolated malformations. The sizes of the pathogenic aberrations varied from 0.04 Mb to 51 Mb, the smallest representing a single gene deletion and the largest a chromosomal aberration. Four of the pathogenic variants were large chromosomal aberrations, and three of these were identified in males and were the result of unbalanced translocations. The two largest imbalances were inherited from fathers that were carriers of the balanced forms of the translocations (patients FC14(III) and P15(IV), respectively), while the smaller one had occurred *de novo* in the proband (P4(IV)). In the fourth patient, P24(V), a mosaic rearrangement on chromosome 15 was found to have occurred *de novo* in the patient. Of the four detected aberrations, all but the 18;22 translocation should are readily detected by chromosome analysis.

Large chromosomal imbalances are known causes of congenital malformations and since we did not perform cytogenetic pre-screening before array CGH analysis, the identification of several large structural aberrations is expected. Unbalanced chromosomal translocations are usually unique within families and do not represent causes that can be found in many patients, although some common features may be identified between patients with overlapping aberrations. Candidate disease gene identification from chromosomal aberrations becomes rather speculative since the abnormalities usually affect large numbers of genes and both duplications and deletions are present in the same patient.

Two pathogenic findings were not detectable by chromosome analysis, a whole gene *FANCB* deletion in a fetal case (FC10 (III)) and a deletion in the distal 22q11 deletion region in a patient with syndromic anal atresia (P12(V)) (Table 3). The *FANCB* deletion shows X-linked recessive inheritance and was transmitted from a healthy mother to a male fetus while the 22q11.22 deletion was inherited from a healthy father. In this case it remains unclear which role this aberration has in the patient phenotype since this deletion is associated with a mild, primarily cognitive, phenotype.

VOUS were detected in 32 patients, and variants in 10 of these patients were chosen based on size, copy number type and gene content, and further characterized by investigation of inheritance status (Table 4). In one case the CNV turned out to be *de novo* in the patient (patient P24(V)). In two cases the familial samples available were not enough to determine inheritance. In seven cases, the aberrations were inherited from apparently healthy parents.

Table 3. Clinical features and pathogenic aberrations detected in patients with congenital malformations by screening with array CGH analysis (studies III, IV and V).

	Patient/ fetal case	Sex	Malformation phenotype	Additional features	Detected aberration	Inheritance
Chromosomal aberrations	FC14(III)	M	complex heart malformation, esophageal atresia, unilateral mild hydronephrosis and hydroureter, dilatation of the 4th cerebral ventricle	-	46,XY,der(9)t(9;18)(p24.1;q12.3) del(9)(p24.3p24.1) 7.3 Mb dup(18)(q12.3q23) 34.3 Mb	father carrier of balanced translocation t(9;18)(p24.1;q12.3)
	P4(IV)	M	multiple VSDs, right rotation of the heart, right ventricular outflow tract obstruction, craniosynostosis, clubfoot	strabismus, mild dysmorphic features	46,XY,der(18)t(18;22) (p11.32;q13.31) del (18)(p11.32p11.32) 2.6 Mb dup(22)(q13.31q13.33) 4.9 Mb	*de novo*
	P15(IV)	M	VSD, pulmonary valve stenosis, small penis, cryptorchidism, ankyloglossia, absent thymus, hemangioma, narrow auditory canals	dysmorphic facial features, seizures at birth, stridor, hearing impairment	46,XY,der(5)t(5; 8)(p15.33;q22.1) del(5)(p15.33p15.33) 4.3 Mb dup(8)(q22.1q24.3) 51 Mb	father carrier of balanced translocation t(5; 8)(p15.33;q22.1)
	P24(V)	F	EA, VSDs	DD, epilepsy, severe respiratory tract infections	46,XX,add(15)(p10)[12]/47,XX,15ps-,+mar[8]/46,XX,15ps-[2] dup(15)(q11.2q13.1) 7.7 Mb	*de novo*
CNVs	FC10(III)	M	anal atresia, complex Fallot-like heart malformation, horseshoe kidney, bilateral rudimentary thumbs, malformed right ear with atresia of the external auditory canal	-	del(X)(p22.2p22.2) 0.01-0.04 Mb	X-linked inheritance from healthy mother
	P12(V)	M	anal atresia, vertebral defects, small VSD	-	del(22)(q11.22q11.22) 0.67 Mb	inherited from healthy father

VSD ventricular septal defect, EA esophageal atresia.

Table 4. Patients with variants of unclear clinical significance (VOUS) detected by screening analysis using array CGH (paper III, IV and V).				
Patient/FC	**Sex**	**Phenotype**	**Detected aberration**	**Inheritance**
V19(III)	F	EA, right-sided aortic arch	dup(19)(q12q12) 0.39 Mb	inherited from healthy mother
FC10(III)	M	AA, complex Fallot-like malformation, absent ductus arteriosus, right-sided aortic arch, horseshoe kidney, rudimentary thumbs bilaterally, small malformed left ear without external auditory canal, low-set right ear	dup(16)(p13.11p12.3) 2.9 Mb	not inherited from mother, sample from father unavailable
P5(IV)	M	TGA, hypospadias	del(7)(p21.3p21.3) 0.40 Mb	inherited from healthy mother (mosaic)
P16(IV)	M	TOF	dup(17)(q23.2q23.2) 0.73 Mb	inherited from healthy mother
P13(IV)	F	CoA, VSD, BAV	del(6)(q23.2q23.2) 0.58 Mb	inherited from healthy mother
P1(IV)	F	TGA, bilateral pes cavus, overlapping 3rd toes, laryngomalacia, rigid hips, dysmorphic facial features, DD, GR	del(6)(q26q26) 0.28 Mb	inherited from healthy mother
P25(V)	F	EA, duodenal atresia, pelvic kidney and aortic arch anomalies	dup(15)(q26.3q26.3) 0.4 Mb	parental samples not available, duplication detected in healthy monozygous twin
P24(V)	F	EA, VSD, DD, seizures	del(15)(q26.3q26.3) 0.7 Mb	*de novo*
P1(V)	M	EA, vertebral defects, three small VSDs	del(13)(q21.1q21.1) 1.0 Mb	inherited from healthy father
P6(V)	F	EA, ASD, VSD, persisting left v. cava sup.	del(5)(q14.2q14.2) 0.18 Mb	inherited from healthy father

EA esophageal atresia, AA anal atresia, TGA transposition of the great arteries, TOF tetralogy of Fallot, CoA coarctation of the aorta, VSD ventricular septal defect, BAV bicuspid aortic valve, DD developmental delay, GR growth retardation, ASD atrial septal defect, v. cava sup. vena cava superior.

Altogether, screening for gene dose alterations in 86 patients with isolated or syndromic congenital malformations revealed pathogenic aberrations in 6/87 patients (7%) (Table 5) and since pathogenic variants were detected only in syndromic cases, the proportion of pathogenic findings was higher when patients with isolated malformations were excluded (10% and 20%, respectively, including and excluding the VACTERL group). Pathogenic CNVs that would not be detectable by array CGH were found in 2/86 patients (2% and 4% including and excluding VACTERL, respectively). VOUS were detected in 32/86 patients (37%). A similar study using microarray analysis to investigate copy number variants in 95 fetal cases with isolated and syndromic congenital malformations reported CNVs >100 kb in size in 21% of patients (94). Similarly, in study III-V, 26/86 patients and fetal cases (30%) had at least one CNV > 100 kb.

Considering the phenotypes of patients and fetal cases screened in study IV-V, 58% had isolated malformations, which is likely higher than in most clinical settings. It is noteworthy that none of the clearly pathogenic findings were identified in patients with isolated malformations. Comparisons between patients with isolated and syndromic phenotypes did not reach statistical significance, due to small sample sizes and low number of genetic findings, although the results may indicate that copy number variants > 50 kb are rare causes of isolated malformations. Different proportions of causative findings have been reported using array CGH in patients with isolated heart malformations, and the utility of array CGH screening in this patient group has been discussed previously (95). The heterogeneity of the group "congenital heart malformations" makes conclusions about utility difficult in this specific group.

Nine patients (9/86; 10%) had two or more CNVs > 50 kb and four patients had two or more CNVs >100 kb (4/86; 5%). A two-hit model has been proposed for intellectual disability, in which one CNV can confer risk for neuropsychiatric disease and the combination with a second CNV aggravates the phenotype (96), and it is plausible that a similar mechanism could be involved in congenital malformation phenotypes, although this would be difficult to prove in individual cases. In two of the four patients with more than one CNV > 100 kb (P12(V) and P24(V)), one of the identified CNVs was pathogenic but could not completely explain the phenotypes seen in the patients. In the other two cases, it was not possible to decide whether the detected variants were associated with the patient phenotypes. It is conceivable that the combinations of CNVs are pathogenic, but presently their etiologic role cannot be determined in these cases.

The low proportion of causative findings may partly be explained by the high proportion of patients with isolated malformations and the conclusion by us and others that CNVs are rarely seen in patients with VACTERL association. The choice of platform naturally influences detection rates depending on resolution and robustness. The resolution of the 180K

Table 5. Copy number variations detected by array CGH analysis in patients with congenital malformations (study III, IV, V).								
Study	Number of patients	Patients with isolated malformations	Patients with detected pathogenic CNVs	Patients with detected VOUS >50 kb	Patients with syndromic phenotypes and detected pathogenic CNVs	Patients with isolated malformations and detected pathogenic CNVs	Patients with syndromic phenotypes and detected VOUS	Patients with isolated malformations and detected VOUS
III	39	-	2/39 (5%)	12/39 (31%)	2/39 (5%)	-	12/39 (31%)	12/39 (31%)
IV	23	16/23 (70%)	2/23 (9%)	9/23 (39%)	2/7 (29%)	0 (0%)	2/7 (29%)	2/7 (29%)
V	25	12/25 (48%)	2/25 (8%)	11/25 (44%)	2/13 (15%)	0 (0%)	7/13 (54%)	7/13 (54%)
Total	**86[1]**	**28/48[2] (58%)**	**6/86[1] (7%)**	**32/86[1] (37%)**	**6/58[1] (10%) 4/20[2] (20%)**	**0/28 (0%)**	**21/58[1] (36%) 9/20[2] (45%)**	**21/58[1] (36%) 9/20[2] (45%)**

VOUS variant of unclear significance.

[1]Correcting total number of patients for one patient that was included in study III and IV.

[2]Excluding patients with VACTERL association or a VACTERL-like phenotype in study III.

array used in our study is 30-50 kb, and thus we have not detected smaller CNVs which are known to be frequent in the genome. Also, patient selection in our studies differs from other studies in that we have only included surgically treated patients, so that mosaicism could be investigated. This has limited the number of patients included and potentially introduced a bias, although the proportion of patients with identified aberrations in our studies is similar to those reported in other studies. In addition we did not perform chromosome analysis before array CGH analysis, reflecting the finding of several large chromosome aberrations in our studies. Importantly, the patient material in our studies is heterogeneous, both within and between studies, and this fact likely reduces the possibility to detect aberrations present in specific malformations.

4.1.1 Copy number variants in patients with congenital heart malformations

Pathogenic gene dose alterations were detected in two patients with heart malformations (2/23; 9%) and VOUS were detected in nine patients (9/23; 39%). Based primarily on size and gene content, inheritance status was investigated in three patients with VOUS and shown to be inherited from unaffected parents and are thus continuously regarded as unclear variants that need further study. One of these regions, 6q23.2, may represent a susceptibility region for congenital heart malformations and is further discussed in section 4.1.5.1. It is noteworthy that inherited variants have been reported as causative in cases of heart malformations (97, 98).
Published studies of copy number variants in patients with heart malformations have reported similar or higher frequencies, although varying inclusion criteria, definition of copy number variants, diagnostic platforms and study design make comparisons difficult (98-102).

4.1.2 Copy number variants detected in patients with esophageal atresia

Screening for pathogenic aberrations in esophageal atresia revealed one clearly pathogenic aberration and VOUS in all five patients. We identified two novel regions in which gene dose alterations could be associated with increased risk for EA, discussed in section 4.1.5.2. Genetic causes can presently be identified in around 11-12% of patients with esophageal atresia and gene dose alterations are reported in individual cases, while more systematic studies have yet to be published (103). One study reporting results from array CGH in seven monozygotic twin pairs discordant for esophageal atresia did not report pathogenic variants (104). All patients in our material had syndromic EA, which is a possible explanation to a high proportion of pathogenic or VOUS identified in our study. The number of patients is too low to contribute to the determination of proportion of pathogenic copy number variants in patients with EA.

4.1.3 Copy number variants detected in patients with anal atresia

In our study of patients with anal atresia, a pathogenic CNV was detected in one patient (1/14; 7%) and VOUS were detected in five patients (5/14; 36%). For comparison, probable disease-causing CNVs were reported in 17% of patients with anorectal malformations in

combination with a CNS malformation (105). In our study, 50% of patients had isolated malformations, possibly decreasing the likelihood of detecting pathogenic CNVs.

4.1.4 Copy number variants in patients with hydronephrosis

In our study, we did not identify pathogenic CNVs in patients with congenital hydronephrosis. Two patients (2/6; 33%) had VOUS. The study group is very small and no patients had additional malformations, and thus we cannot draw any conclusions about the role of copy number variants in congenital hydronephrosis.

4.1.5 Interesting copy number variant findings in patients with congenital malformations

4.1.5.1 Deletion at 6q23 in a patient with CoA, VSD and bicuspid aortic valves (BAV)

A deletion at 6q23.2 was detected in patient P13(IV), with an isolated, combined heart malformation including CoA, VSD and BAV. The 0.58 Mb deletion was found to be inherited from a mother without a known heart malformation. The deletion overlaps a region reported to be associated with hypoplastic left heart syndrome (106), and four of the six deleted genes have been reported to be expressed in heart tissue and to be involved in heart or vasculature function or development. Homozygous mutations in the *ENPP1* gene, encoding an ectoenzyme involved in hydrolysis of extracellular nucleotides, cause generalized arterial calcification of infancy (OMIM:#208000). A second gene within the 6q23.2 region, *ENPP3*, encodes an isoenzyme of *ENPP1*, and has no known disease association. Interestingly, the mouse homologous gene, *Enpp3*, shows expression in the second heart field and may be regulated by *Nkx2.5* (107-109). Further, the *MED23* gene encodes a subunit of the large mediator complex (MED) involved in transcription activation in which mutations in genes encoding other subunits have been reported in congenital heart disease while the subunit encoded by *MED23* has so far not been associated with heart malformations (110). Nevertheless, a *Med23*$^{-/-}$ mouse model is embryonic lethal and displays disorganization of the vasculature (111). The *ARG1* gene is expressed in heart and may be linked to coronary heart disease (112). A search in the DECIPHER database reveals one patient (patient 251447) with an overlapping 7 Mb-deletion and congenital anomalies not including heart malformation.

It is known that cardiovascular genetics are complicated, with single gene mutations associated with heart malformations showing reduced penetrance and varying expressivity, and it has been suggested that the causality of CNVs should be evaluated on the basis of gene content rather than inheritance and size (113). Considering that four genes in the region seem to be involved in heart or vasculature development or function, two of them possibly in the same pathway, it seems likely that the variant affects development of the heart and confers risk for erroneous morphogenesis with influence of additional genetic or other factors.

4.1.5.2 Copy number variants at 15q26.3 in patients with esophageal atresia

Copy number variants at 15q26.3 were detected in three patients with esophageal atresia and other features. One was a *de novo* 0.7 Mb deletion affecting seven genes while the other two

were gains (a 0.5 Mb duplication and a 0.4 Mb triplication inpatients P4(V) and P25(V), respectively), affecting three genes, and one of them was detected in a healthy monozygous twin. The duplication region overlaps with CNVs in the DGV.

The deletion was detected in patient P24(V) who had partial tetrasomy 15. Esophageal atresia has not been described with the tetrasomy 15 genotype previously, drawing attention to the 15q26.3 region in relation to esophageal atresia. The causal relation of the deletion and EA could be one of three possible alternatives: 1) the deletion is unrelated to EA, 2) the deletion in combination with tetrasomy 15 can cause EA, or 3) the deletion in itself is causal.

Combining the results of these three patients and the fact that esophageal atresia is an unusual phenotype, the 15q26.3 region seems interesting to study further. Several of the genes affected by the detected CNVs are recessive disease genes not associated with esophageal atresia, and since the deletion and duplications are not overlapping, it seems likely that a putative developmental effect is exerted through disruption of gene expression regulation.

4.1.6 Summary

The combined results of study III-V emphasize importance of copy number variant analysis in patients and fetal cases with malformations, since variants with high penetrance conferring a high recurrence risk are detected. The importance of detection of similar pathogenic variants is very high to individual families, with the possibility to provide an explanation, to lower recurrence risk estimates and to offer prenatal diagnostics in future pregnancies. Our analysis is too small for proportion estimates of causative CNVs in individual malformations, but provides further evidence for the contribution of genetic factors in development of malformations.

In patients with isolated malformations, the likelihood of finding pathogenic causes by copy number analysis down to 50 kb may be low, although some studies that find no difference between *de novo* CNVs in syndromic and isolated phenotypes provide evidence to the contrary (98). Additionally, VOUS which may represent risk factors are detected in a substantial proportion of these patients and the impact of such variants will be important to investigate to further clarify the etiology behind malformations.

4.2 MOSAICISM AND CHIMERISM IN PATIENTS WITH CONGENITAL MALFORMATIONS (STUDIES I, II, IV AND V)

A specific aim of this thesis was to investigate the role of genetic mosaicism for copy number variants as a cause of congenital malformations. Altogether, 18 patients from study IV and V (including ten patients with heart, three patients with EA, three patients with anal atresia and two patients with congenital hydronephrosis) were analyzed for pathogenic aberrations or VOUS > 50 kb in tissue from the malformation and also in blood samples. We did not detect discordance of VOUS patterns between tissue from the malformations and blood in the same individuals, and thus our studies do not show that mosaic copy number variation is a common underlying cause of congenital malformations. A number of inherent challenges in studying

tissue-specific mosaicism as well as specific limitations to our studies, including choosing the appropriate tissue for analysis, difficulty of obtaining tissue samples from malformations and sensitivity of the analysis method complicate interpretation of the study results.

So far, no systematic studies comparing copy number variations detected by array CGH between tissue from malformations and blood in the same individuals have been published. Recently, Bednarczyk *et al.* reported the presence of five different copy number variants detected in esophageal tissue but not detectable in blood in a patient fulfilling criteria for VACTERL association (114). Studies of heart malformations investigating genetic mosaicism for specific genetic causes using other methods such as DNA sequencing, MLPA and FISH have been published (115-130). A number of studies by Reamon-Buettner *et al.* reported somatic mutations in cardiac disease genes identified in formalin-fixed tissue from hearts affected by malformations, that other groups were not able to replicate analyzing fresh frozen heart tissue. Apart from that, one study where fresh frozen heart tissue was used for analysis showed somatic mutations in a subset of patients with sporadic tetralogy of Fallot (130). Single studies screening for mosaic forms of X-chromosome aberrations or 22q11-deletions in cardiac tissue with fluorescent *in situ* hybridization (FISH) and short tandem repeat markers have not reported positive findings (115-117).
For comparison, mosaicism for CNVs detected in DNA from blood samples in children with developmental delay and/or multiple congenital anomalies has been reported with prevalences of 0.5-3.74% (86).

In our studies, an obvious disadvantage is the low number of patients overall and in separate diagnostic groups, a result of the difficulty to collect tissue samples from surgical procedures. Ethical restrictions to our study allowed collection of discarded material only, and this has limited the number of surgical procedures where collection was permitted and also the possibility to choose a common source of material. When studying mosaicism, there is a great uncertainty as to which tissue to study and conclusions must be drawn accordingly, so that in our studies, we could only hope to detect mosaicism present in the whole malformation or organ, but not restricted to specific regions. This presumably could exclude aberrations that have occurred later during development in more differentiated cells that would affect a specific region. In heart development, the difficulty of choosing the appropriate tissue is complicated by the hemodynamic effect on heart development so that an initial primary abnormally developing structure that affects blood flow can result in secondary malformations that may be more prominent (131).
The resolution of the microarray platform, both regarding size of aberrations and level of mosaicism, is important and the 30-50 kb resolution of the 180K oligonucleotide platform used in our studies precludes detection of small CNVs. In general oligonucleotide microarrays, with minimal mosaicism detection rates around 10-20%, are less sensitive compared to SNP arrays and possibly also BAC arrays in detecting low-level mosaicism (132, 133). The application of mosaicism filters in the CNV algorithm that we used corresponds to a mosaicism level of 15% combined with a size of 5 Mb, although visual inspection for smaller aberrations with log2-ratios indicating mosaicism was performed. It is

however possible that small aberrations present in <15% of cells have not been detected. Somatic CNVs present in more than one tissue (heart, kidney and skin) have been reported in a deceased individual without malformations (79), and since CNVs in tissue from congenital malformations have not systematically been studied before, we chose to include patients with syndromic CHD, and it is possible that this has lowered our chances of finding pathogenic mosaicism.

4.2.1 Aspects of mosaicism

Different aspects of mosaicism can be highlighted with examples from our studies, for example parental mosaicism and inheritance, mechanism of formation, phenotypes and prognosis associated with the genotype.

4.2.1.1 Genotype-phenotype correlations in mosaic disorders

A known difficulty of genetic counselling in mosaic disorders is the varying phenotype associated with mosaic forms of chromosomal aberrations, and this is exemplified by the patient described in study I. The underlying explanation is likely the variation of timing for the occurrence of the mutation, the cell type affected and the phenotypic effects of the aberration on the cell type that will influence distribution and effects of the aberration. In the patient of study I, analysis of several different tissues as well as different techniques was used to determine the proportion of the trisomic cell population. The patient had a very mild phenotype compared to other patients with trisomy 14 mosaicism, despite a detected proportion of trisomic cells of up to 30%. Also, it seems that the clinical phenotype in trisomy 14 mosaicism is unrelated to the level of mosaicism detected in blood (134). The explanation may be a low proportion of trisomic cells in the CNS, or in cell populations important during the development of CNS, but this explanation is difficult to prove.

4.2.1.2 Update of prognosis for non-mosaic chromosomal aberrations

Several conditions are described in the literature as non-compatible with survival in non-mosic form, for example trisomy 8, 9, 14, 17 and 22 (135), while constitutional trisomies for chromosomes 13, 18, 21 and X can be seen in live-born babies, but with drastically reduced survival for trisomy 13 and 18 (90% are reported to die within the first year of life from cardiac, renal or neurological complications, or in trisomy 18 from infections). The tendency to surgical treatment in this patient category has been restrictive since it has been reported that survival is not improved by surgery (136, 137).
In the patient described in study II, partial tetrasomy 14 was detected. The patient is severely affected with multiple malformations, hearing and visual impairment, global developmental delay and severe reactions to respiratory tract infections which have required intensive care treatment and tracheostomy. Since all previously reported patients with non-mosaic forms of the condition have been lethal during infancy, the parents were counselled accordingly. Analysis using array CGH, chromosome and FISH analysis did not give evidence of mosaicism in this patient. In chromosome analysis, 25 metaphases each from cultured peripheral blood lymphocytes and fibroblasts were analyzed, which can exclude mosaicism at

the 12% level with 95% confidence in both tissues separately (138). In addition, 100 interphases from peripheral blood lymphocytes were analyzed without evidence of mosaicism. It is still possible that a low-level mosaicism is present, or that a tissue-specific pattern for the marker chromosome with lower levels in specific tissue not analyzed by us is present. Still, this is the first patient with apparent non-mosaic tetrasomy 14 and long-time survival, likely reflecting advances in pediatric intensive care treatment during the last decades.

In light of these advances, it is important to reevaluate survival of patients severely affected by chromosomal aberrations, since prognosis is likely based on data accumulated during times when surgical techniques and intensive care were less developed. A recent study of the effects of surgical treatment of congenital heart malformations in patients with trisomy 13 and 18 show increased survival in these patients compared to patients who received expectant management. The study has limitations such as short follow-up time and a possible bias for patients with less severe phenotypes undergoing treatment, but if a subset of patients can benefit from surgical repair, it is still important to take into account in determining treatment strategies for these patients (139).

4.2.1.3 Formation of mosaic partial tetrasomy 15

In study V, an unusual structural variant of the inv dup(15) syndrome was identified in patient P24(V), equalling partial tetrasomy 15 in 82% of peripheral blood cells. The cell populations harbouring the tetrasomy 15 showed two different alternative rearrangements with the extra material located to 15p or a marker chromosome, respectively. In a subset of cells, the extra material was lost. From the existing information it is still possible to suggest that the most frequent disomic cell population with extra material on 15p was the original one and that instability of the arrangement has resulted in formation of a ring chromosome and subsequent loss of the ring in a subset of cells. Formation of the more common inv dup(15) is described as the result of a U-type exchange mechanism in maternal meiosis (140). The U-type exchange mechanism cannot alone explain the rearrangement in the patient, and it is possible that an initially formed isodicentric marker underwent a subsequent rearrangement including non-homologous end-joining of the isodicentric marker and 15p to result in the observed karyotype. A similar mechanism has been described to explain the occurrence of translocations of 15q, believed to be the reciprocal products of the inv dup(15) chromosome (141). Further determination of the mechanism in patient P24(V) is dependent on microsatellite marker analysis in parents and proband.

4.2.1.4 Parental mosaicism

A 0.40 Mb deletion on 7p21.3 detected in patient P5(IV), a boy who was born with transposition of the great arteries and severe hypospadias, was found to be inherited from a mother who was reportedly healthy. Parental analysis was performed in an array CGH analysis using DNA from buccal cells, and hybridization of maternal against paternal DNA. Visual examination of the array result for the parents revealed a mean log2-ratio of -0.16 for

the region deleted in the patient, indicative of mosaicism in the mother. Analysis of meta- and interphase chromosomes (15 and 332 cells analyzed, respectively) from peripheral blood of the patient's mother showed the deletion to be present in a mosaic state in 16% of lymphocytes while analysis in a healthy brother of the patient revealed a normal pattern. The inheritance pattern indicates that the CNV could be pathogenic, although further studies need to clarify this. The recurrence risk cannot be estimated, both due to the unclear significance of the deletion and also due to the limitations to investigating the proportion of oocytes carrying the deletion. The family was counselled that they are likely to have an increased risk of having an affected child in future pregnancies, and that intensified ultrasound investigations are indicated, but presently we may not offer prenatal testing due to the unclear nature of the CNV.

Recently, low-level somatic mosaicism for genetic aberrations in parents was reported to be the cause of constitutional genetic syndromes in their offspring (87, 98). This is an important discovery since the recurrence rate would be estimated to be considerably higher than if the aberrations had occurred *de novo*, possibly as high as 50% (7). Genetic counselling will be difficult if pathogenicity can not be determined for the variant, since a high recurrence risk cannot be excluded, but prenatal diagnosis cannot be offered to the family.

In this case, the deletion interrupts two genes, *THSD7A* and *PHF14*, removing the last 19 exons of *THSD7A* and the last exon of one splice variant of *PHF14*. *THSD7A* encodes a protein that has been shown to be important in endothelial cell migrations in zebrafish angiogenesis (142) while *PHF14* is likely involved in chromatin-mediated transcriptional regulation (143). Further studies of this region are needed to determine a potential role in development of congenital malformations.

4.2.2 Chimerism

To learn more about genetic mechanisms behind malformations, we wanted to specifically investigate the mechanism behind the unusual gene dose alterations found in clinical investigations of individuals with congenital malformations. In a boy with disorder of sex development, mild skin manifestations and normal psychomotor development, chromosome analysis revealed a 46,XX/47,XY,+14 karyotype. The parental contributions to genetically distinct cell populations in an individual can be determined using microsatellite marker analysis and results in this patient revealed chimerism as an underlying cause. Although chimerism is not necessarily a cause of phenotypic abnormalities, and is sometimes discovered by chance when results of blood group analysis are ambiguous, some types of chimerism may result in abnormalities. When the sex chromosome content of two cell populations is discordant, varying degrees of disorder of sex development may result, although phenotypes can be normal. A second disease-causing mechanism is when one of the cell populations shows uniparental disomy (144), leading to imprinting disorders in a varying degree in the case of a maternal isodisomic clone, and placental mesenchymal dysplasia in case of a paternal clone. The identification of chimerism also provides a unique possibility to discover mechanisms involved in human fertilization and the first part of embryogenesis.

4.2.3 Summary

Our studies do not provide evidence for tissue-specific mosaicism as a common cause in congenital malformations, however, the sample-size and heterogeneous diagnoses would only allow detection if mosaicism was a frequent underlying cause. It is still possible that tissue-specific mosaicism for CNVs or other types of genetic aberrations are causative, and will be detected in future studies using other techniques. The finding of parental mosaicism for a putative pathogenic CNV also points to the importance of detecting genetic mosaicism in parents, to improve recurrence risk estimations and genetic counselling.

4.3 CAUSES OF VACTERL ASSOCIATION (STUDY III)

4.3.1 Copy number variants

The proportion of pathogenic CNVs identified in patients with VACTERL association was 5%, which is considerably lower than in syndromic patients from studies IV and V (20%), although comparisons are not statistically significant. The finding of a lower proportion of pathogenic findings in patients with VACTERL association (study III) compared to patients with malformations that were not included based on a VACTERL diagnosis (although some patients fulfil VACTERL criteria in studies IV and V), may still support the finding in study III and other reports that copy number variants are not common causes behind VACTERL association, even compared to patients with multiple malformations. The proportion of causative gene dose alterations is similar to the result reported in a study of the same size by Hilger *et al.* who reported 6% causative *de novo* CNVs (145). With a number of published microarray studies, it seems that CNVs are causative in a small subset of patients with VACTERL association; however, these singular cases are still important for implications of genes and pathways involved in pathogenesis (146-148). It is still possible that small CNVs, not detected by current microarray platforms, are important. Notably, analysis in the 19 fetal cases included was performed on DNA isolated from tissue samples (lung, spleen, heart or liver). It is possible, but not likely, that pathogenic copy number variants could be restricted to blood cells or other tissues not sampled.

It seems that copy number variants can explain a small proportion of cases in the patient group, but points to other possible causes of VACTERL that have not been identified yet.

4.3.2 Single gene defects

Mutation screening in *CHD7*, *PCSK5* and *HOXD13* revealed a pathogenic mutation in *CHD7* in one patient. The *de novo CHD7* mutation was identified in a patient with a VACTERL-like phenotype, establishing a diagnosis of CHARGE syndrome. The phenotypic overlap between VACTERL and CHARGE include esophageal atresia, cardiac defects, anal atresia, renal and limb anomalies and our result emphasizes the importance of differential diagnosis in patients with these phenotypes.
The search for a common genetic cause in VACTERL association is often compared to the search for and identification of a causative gene in patients with CHARGE syndrome, since

CHARGE also occurs sporadically and was shown to be caused by *de novo* mutations. An important difference between the VACTERL and CHARGE phenotypes is that patients with CHARGE syndrome have a cognitive impairment which likely results in a low reproduction rate in these patients. In contrast neurodevelopmental phenotypes are not seen in patients with VACTERL association, and a historically low rate of reproduction is likely due to the severity of malformations, especially ARM. With improvement of surgical treatment, it is possible that patients with VACTERL association will be more reproductively active and that a different recurrence risk could be observed in the future.

Screening for mutations in *PCSK5* revealed three rare missense variants, with genotype frequencies in dbSNP of 0.03-2.8%, respectively. The variant with the lowest frequency affected a highly conserved Cys residue, resulting in a Cys→Tyr amino acid change, possibly disrupting a disulphide bond. The frequency reported is too high to represent a penetrant mutation, but could represent a variant with reduced penetrance. The way in which *PCSK5* mutations are believed to cause a VACTERL phenotype is through reduced cleavage of one of its' target proteins, GDF11, which has been shown to participate in regulation of *HOX* genes, long regarded as candidate genes for VACTERL association (74).

So far, it seems that single gene defects represent causative factors in a subset of VACTERL patients. Massive parallel sequencing has not provided a common cause, as in many other congenital disorders, although new recessive mutations in *TRAP1* were recently reported to be present in 2/300 patients with VACTERL association including kidney abnormalities (149, 150), and so far no VACTERL exome sequencing studies have reported a high proportion of *de novo* mutations.

4.3.3 Patient selection

Since VACTERL association is a rare disorder with relatively few genetic findings, we chose to include both patients (12) and fetal cases (19) fulfilling diagnostic criteria and patients with esophageal atresia and at least one additional malformation in the VACTERL spectrum, not fulfilling diagnostic criteria (8), to increase the study group and increase the likelihood of finding CNVs with potential candidate genes that could point to specific pathways important in development of the malformations seen in VACTERL.

It is interesting that both pathogenic gene dose alterations and all variants detected in *PCSK5* were identified in fetal cases. All of the included fetal cases, except one, were aborted after detection of fetal abnormalities at ultrasound screening, and thus were not spontaneous abortions. It may be the case that the fetal phenotypes that are diagnosed with ultrasound are more severe compared to the patient phenotypes in our study, and therefore have a higher degree of genetic findings, but due to the over-all low frequency of findings this remains unclear.
Inclusion of fetal cases where the neurological phenotype cannot fully be evaluated may result in inclusion of cases that, in a postnatal setting, would not have been diagnosed with VACTERL because of developmental delay, and this may decrease the likelihood of finding

true VACTERL causes in the patients in our study. It is also possible that the patients not fulfilling VACTERL criteria in our study would be less likely to have genetic aberrations, and that inclusion of these patients lower the expected diagnostic proportion. In our case, the inclusion of fetal cases and patients not fulfilling VACTERL criteria makes the results less suitable to generalize in terms of frequency, and can be regarded more as a search for candidate genes. Given the high number of included fetal cases fulfilling VACTERL criteria and pathogenic findings in 11% of these (2/19), one conclusion that may be drawn is that genetic analysis in fetal cases are important.

4.3.4 Summary

Systematic studies, including our own, have so far not revealed causative single gene defects, copy number variants or chromosomal aberrations in a large number of patients. Some pathways involved in development of mesodermal structures and mesodermal processes are implicated by several studies (14). Because of the intricate signalling network involved in the embryologic pattern formation and organogenesis, it is possible that *de novo* mutations in different genes involved in the same network may be responsible, or that oligogenic influence is of importance. A theory that should be systematically tested is the role of epigenetic factors in VACTERL association. It is also possible that whole genome sequencing will reveal mutations.

5 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

In 2008, when we set out to study copy number variants and mosaicism, array CGH had relatively recently been introduced as a method used in routine clinical genetic diagnostics, and most cases of mosaicism had been diagnosed in the form of aneuploidy by chromosome analysis, although examples of mosaic single gene defects were known. During the course of these years the technical development has allowed introduction of exome, genome and transcriptome sequencing, which can identify the majority of single nucleotide variants and structural rearrangements found in a human genome in one single reaction, and in addition has the ability to detect mosaicism down to the level of 1% (151). During the next few years, massive parallel sequencing will successively replace array CGH as a screening method in patients with developmental disease and congenital malformations in research and clinical practice.

We have, along with other research groups, studied the occurrence of pathogenic copy number variants in patients with different types of malformations, and the results indicate that copy number variants likely represent important and fully penetrant causes in a subset of patients, but do not explain the great proportion of isolated or syndromic cases. The characterization of inherited copy number variants and determination of their role in genetic disease is ongoing and will likely continue. Similarly for single gene defects, the contribution of *de novo* and inherited mutations in congenital malformations will be further investigated using massive parallel sequencing. Because of the great number of genes involved in human development, the number of possible target genes for mutations is very large and it is possible that *de novo* single gene or copy number mutations, alone or in combination with other genetic variants, in a multitude of different target genes are causative. It has been shown that aneuploidy is very common in cleavage stage embryos, possibly due to a lower degree of cell cycle control, and perhaps single gene mutations are also introduced at this stage. In developmental disease, the importance of regulation of gene expression has come more into focus and in heart malformations, chromatin regulation has been proposed to provide a new important “pathway“ where genetic events can result in disease (152). Since signalling networks active during embryonic development are often involved in development of many organs, it is likely that they may provide explanations also in other diseases.

We have investigated if mosaicism for genetic aberrations detected in the malformed tissue could be an overlooked cause of malformations. Based on the results of our and other studies, mosaicism for different types of aberrations seems more widespread than previously known, and is likely to have important roles in many types of disease, although we could not show that mosaic copy number changes in malformed organs are common. It is likely that underlying mosaicism and chimerism will be encountered more often with new methods that can detect very low levels of distinct cell populations and methods that allow genome-wide analysis of single cells will contribute to increased knowledge about mechanisms or chain of events that lead to mosaic genotypes and phenotypes.

In VACTERL association, it seems that a common cause will not be found. It is likely that the phenotype can be the result of several different mechanisms that affect mainly mesodermal structures. Since large scale sequencing has so far not provided the solution, it is possible that the answer is not in the coding sequence. Deregulation of developmentally important genes could be involved, either through mutations or rearrangements of regulatory units, mutations in non-coding RNAs or through changes in chromatin regulation by genetic or environmental influences. Studying altered gene expression in fetuses with diabetic mothers in animal models could provide insight into deregulated pathways. Also, the finding of a single umbilical artery in a subset of patients with VACTERL association may be a somewhat overlooked feature that could provide a clue as to the pathogenic processes that are involved (68, 153).

The knowledge about genomic structure and genetic variation increases continuously, and is becoming increasingly complex. It seems that the genetic background in an individual, all the different nucleotide variants, microsatellites and copy number variants provide an important explanation, but not a solution, to the varying phenotypes seen in the majority of genetic disorders. The interplay between this background of variants with varying penetrance in different phenotypes, is likely an important cause of sporadic disease, whether it be congenital malformations or later-onset phenotypes.

6 SAMMANFATTNING PÅ SVENSKA

Medfödda missbildningar är en viktig orsak till dödlighet och sjuklighet under tiden kring födelsen och identifieras hos ca 4% av alla barn som föds i Sverige. Förbättrade behandlingsmöjligheter inom barnkirurgi och intensivvård har lett till ökad överlevnad hos många barn, men trots detta kan missbildningar medföra livslånga besvär och behov av specialiserad sjukvård för drabbade individer. För familjer där barn har fötts med missbildningar är möjligheten att få information om bakomliggande orsak, prognos samt återupprepningsrisk vid ny graviditet centrala frågor. Ungefär hälften av alla familjer med missbildning i kombination med utvecklingsförsening kan få en förklaring till varför deras barn drabbats, och för patienter med enbart missbildning är siffran betydligt lägre. Syftet med våra studier har varit att öka kunskapen om bakomliggande orsaker till medfödda missbildningar för att kunna erbjuda förbättrad klinisk genetisk diagnostik och rådgivning, samt att identifiera gener involverade i normal och avvikande organutveckling.

Viktiga begrepp i våra studier är *copy number variation* (CNV) och *mosaicism.* Man har på senare år upptäckt att tillskott (duplikation) eller förlust (deletion) av arvsmassa, så kallad *copy number variation*, förekommer hos friska individer som en del av den normala variationen mellan människor men att vissa duplikationer och deletioner också kan påverka fosterutveckling och orsaka missbildningar eller utvecklingsstörning. Tidigare har sådana avvikelser studerats med mikroskopi av celler i delningsfas, men detta har medfört att man endast kunna identifiera stora kromosomavvikelser. Med en nyare teknik, array comparative genomic hybridization (array CGH), kan mindre sjukdomsorsakande avvikelser (duplikationer och deletioner) påvisas. Tekniken används i klinisk genetisk diagnostik för att undersöka individer med utvecklingsförsening och/eller medfödda missbildningar. Begreppet *mosaicism* syftar på att man hos vissa individer upptäckt att den genetiska uppsättningen i deras celler skiljer sig åt mellan två eller flera grupper av celler, och sådana skillnader uppstått efter befruktningen. Man vet numer att mosaicism sannolikt förekommer hos många människor utan att det har någon tydlig effekt, men ibland innehåller en av cellgrupperna en genetisk avvikelse som leder till sjukdom hos individen, vilket innebär att man vid en genetisk analys kan upptäcka tillståndet. Då mosaicism definitionsmässigt innebär att en genetisk avvikelse förekommer i en andel av det totala antalet celler i kroppen medför det en oftast en ”utspädning” av sjukdomseffekten. En annan mycket ovanlig bakomliggande orsak till missbildningar och genetiskt åtskiljbara cellgrupper hos en människa är chimerism vilket innebär att ett embryo uppstått genom sammansmältning av två separat befruktade ägg.

Vi har studerat enskilda fall samt grupper av patienter med olika typer av medfödda missbildningar med microarray-teknik samt ett flertal andra genetiska metoder för att undersöka vilken betydelse copy number variation och mosaicism har för uppkomsten av olika typer av medfödda missbildningar, samt hur vissa deletioner och duplikationer uppstått. Vi har utfört analys på arvsmassa isolerad från celler från missbildad vävnad och jämfört med celler från blod hos individer med missbildningar, för att undersöka om genetiska avvikelser begränsade till missbildad vävnad, alltså mosaicism, kan vara en orsak till missbildningar.

Sammanfattning av resultat från våra studier:

- med array CGH analys har vi hittat sjukdomsorsakande duplikationer och deletioner hos 5-8% av patienter med medfödda missbildningar
- sjukdomsorsakande duplikationer och deletioner förekommer hos en liten andel av patienter med VACTERL association
- analys för duplikationer och deletioner i vävnad från missbildningar har inte påvisat mosaicism för copy number variants
- resultat från molekylära analyser indikerar att partenogenetisk aktivering av ett ägg är den bakomliggande orsaken till chimerism hos en patient med könsutvecklingsstörning
- CHARGE syndrom har påvisats hos en patient med symptom liknande de som ses vid tillståndet VACTERL association
- hos en patient med mosaicism för trisomi 14 (förekomst av en extra kromosom 14) har analys av ett flertal olika vävnader visat att den avvikande cellpopulationen förekommer i proportioner varierande mellan 0 och 30%.
- vi har rapporterat överlevnad till skolåldern hos en patient med tillståndet tetrasomi 14 (förekomst av två extra kopior av en del av kromosom 14) i icke-mosaisk form, ett tillstånd som tidigare beskrivits som letalt

Utifrån resultaten av våra studier kan följande slutsatser dras:

- prognoser för ovanliga genetiska tillstånd med missbildningar kan behöva omvärderas då barnkirurgi och barnintensivvård genomgått stora framsteg
- CHARGE syndrom är ett tillstånd som symptommässigt liknar VACTERL association och som är viktigt att särskilja då det har en känd genetisk orsak
- vilka symptom en patient med en genetisk avvikelse i mosaisk form får är svårt att förutsäga, då det sannolikt beror på hur fördelningen av celler med den genetiska avvikelsen sett ut under fosterutvecklingen
- undersökning av duplikationer och deletioner med känslig metodik är viktigt hos patienter med medfödda missbildningar, då upptäckt av sjukdomsorsakande avvikelser har mycket stor betydelse för enskilda familjer i form av bättre information om orsak, prognos och återupprepningsrisk
- mosaicism för sjukdomsorsakande deletioner och duplikationer förefaller inte vara en vanlig orsak till medfödda missbildningar, dock är antalet undersökta fall så litet att man inte kan utesluta detta som en sjukdomsorsak

7 ACKNOWLEDGEMENTS

I thoroughly appreciate the cooperation of patients and families who have chosen to participate in the studies and invested their time so that future patients and families may have the opportunity of improved health care.
During the course of these studies I am grateful for having had the opportunity to meet and work with great people.

Agneta Nordenskjöld, main supervisor. Your devotion to your work is admirable, and so is your energy and strength. I still remember our first meeting together with Ann and Peter when I was still at the Geriatrics department at Danderyd's Hospital, and I still feel extremely grateful that I was given the opportunity to work on this project, it has meant so much to me. I am also grateful for the way you have let me take the detours in life that have appeared during the course of these years.

Ann Nordgren, co-supervisor. You are truly inspirational in science and in life, being ever enthusiastic and infusing energy in all projects and all people that you meet, and somehow, you always find interesting problems that need solving. With your generosity and caring, you create a nice working environment. For the future, I am happy that there are still many unsolved issues, like polar body fertilization and the underlying causes of "the lump".

Peter Gustavsson, co-supervisor. I am really glad that I've had the opportunity to work closely with you during these years and come to know you as a colleague and friend. With great pedagogic skills, patience and calm you always make time for questions, big and small. You always share interesting clinical issues and I enjoy our discussions. You have participated so much to this work.

***Magnus Nordenskjöld**,* group leader. You are a wizard in always creating possibilities for research. Thank you for letting me join your group.

The perinatal pathology unit, Karolinska Huddinge
Nikos Papadogiannakis Each visit to the Perinatal pathology unit has left me with the feeling that I want to come back and learn more. You have generously let me visit both for sample collection as well as for personal interest, taking time from your busy schedule. For me this has been a most appreciated collaboration. Also many thanks to ***Anette Niklasson*** and ***Annika Vestlund*** for being so friendly and helpful with data and fetal samples, always making me feel welcome.
The centres for pediatric cardiac surgery in Gothenburg and Lund
Håkan Berggren, Torsten Malm, Sune Johansson, Boris Nilsson, Sofi Andersson-Stavridis, and ***Jens Johansson Ramgren*** Thank you for believing in our project and for sharing your valuable time. I have greatly appreciated our collaboration.
The Department of Pediatric Surgery, Astrid Lindgren's Children's Hospital ***Tomas Wester, Magdalena Fossum and colleagues*** Thank you for remembering our project during your busy days. Your cooperation has made this project feasible.

Pat and present CMM colleagues
Sigrid Sahlén you have been invaluable, thank you for teaching me everything about work in the genetic laboratory, ***Nina Jäntti*** I am so thankful for all the help in the lab while I was at Huddinge, for always letting me "steal" your computer and for nice conversations, ***Christina Nyström*** so many thanks for being ever helpful, especially with always looking out for my samples and helping with DNA isolation, ***Anna-Lena Kastman*** you brighten everyday at CMM, also thanks for generously sharing your gel wells, ***Britt-Marie Anderlid*** I so much enjoy sharing thoughts about research and clinical work with you, ***Ann-Britt Wikström*** many thanks for being ever helpful and for finding out all the obscure details of the great but complicated concept of research internship.

Marie Meeths My best research friend and a great friend in life. These years would not have been the same without our laughs, research discussions, training sessions and too long coffee breaks. ***Ingegerd Öfverholm*** A fantastic friend on this research and clinic journey, thank you for sharing the horrors of AT and the struggle of getting a driver's license.

Josephine Wincent for valuable research, clinical, life and now also book discussions, ***Ellika Sahlin*** master of genetics and great training sessions, ***Malin Kvarnung*** it's always nice to share thoughts about work and life with you, ***Tobias Laurell*** for always bringing your good mood and sharing interest for mosaicism, ***Anna Brehmer*** for sharing array problems and many conversations in good and bad, ***Vasilios Zachariadis*** for being so generous with your knowledge, ***Alexandra Löfstedt*** for being such a fun person and for bringing discussion skills to the writing room, ***Annika Sääf*** you are still missed in the writing room, ***Anna Svenningsson*** for great swan experiences, ***Susanna von Holst*** for sharing your experiences of dissertation, day care and many other things, ***Amelie Norling*** for bringing some surgery flair to the genetics corridor and for your passionate knowledge of prams, ***Anna Skarin Nordenvall*** for always looking out for me during internship at Huddinge, ***Jia Cao*** for being such a friendly and bubbly person, it's impossible to be in a bad mood after talking to you!

Mårten Winge, Bianca Tesi, Samina Asad, Fulya Taylan, Giedre Grigelioniene, Anna Lindstrand, Wolfgang Hofmeister, Maria Pettersson, Anders Kämpe, Miriam Armenio, Izabella Körberg, Miriam Entesarian, Yougen Chen, Xufeng Zhang, Shulu Zu, Johanna Rantala, Ellen Markljung and Anna Löf Granström for nice discussions and good times.

The Department of Clinical genetics
Agne Liedén, Johanna Lundin and Anh Nhi Tranh Thank you for helping me in my attempts at interpreting array CGH results and for making time for questions at all times. Discussing genetic matters with you always elevates the scientific level.
Kristina Lagerstedt Robinson Many thanks for your help with microsatellite marker analysis.
Irene White, Margareta Lagerberg, Helen Née, Isabel Neira, Ingela Forsberg, Lisa Stolpe and everyone at the Blood section Without your help and experience, this work could not have been done.

Others
Mats A A Persson My first encounter with the science environment in your lab during PRAO was exciting and left a great impression. Thank you for introducing me to research!
Göran Annerén Thank you for sharing your knowledge and your VACTERL association connections.
Erik Sundström and friends at The Research school in molecular medicine for clinicians at Karolinska Institutet (Team Molly 2008), especially ***Veronica Jackson, Kristin Samuelsson, Sam Ghazi*** and ***Emilia Titelman***. Thank you for broadening my molecular knowledge and for great times.
Everyone at Centrum för Sällsynta diagnoser: Eva Ekblom, Katja Ekholm, Kristina Gustavsson Bonnier, Rula Zain, Annika Larsson and ***Alva Ekblom***. Thank you for a nice time writing this thesis, your help and company. The nice atmosphere has been much enjoyed, even if I've mostly been glued to my computer.

Friends and family
All my friends Each with your own special history, always supportive. You are all so cool.
Gunneli and Roland Möllby You have become very important in my life, especially during the last year. Warmest thanks for always helping out.
My grandmother Ragnvi You are an impressive person, and you have been and still are, very important to me, both during childhood and as a grown up. I am grateful for everything you have taught me.
Thomas, ***Thyra*** and ***Edith*** Being part of your family is so valuable.
My sister Josephine and my brother Carl I feel that it was a long time since I was your *older* sister. You are both so cool and inspiring, and I learn from and am inspired by you.
My parents Annica and Per I am so lucky to have you as my parents. Your love, help and support now and growing up has meant everything. The route is not always straight and you're always there for me. Thank you for being who you are.
Rasmus My centre of the universe, thank you for providing so much stability, for knowing me better than anyone and for challenging me on what's important in life. This thesis could never have been accomplished without *your* hard work at home.
Dagmar Our little tiger, the best person I know.

8 REFERENSER

1. Tobias CaF-S. Essential medical genetics. 6th ed. Chichester: Wiley-Blackwell; 2011.

2. Sanchez F, Smitz J. Molecular control of oogenesis. Biochim Biophys Acta. 2012 Dec;1822(12):1896-912.

3. Stukenborg JB, Kjartansdottir KR, Reda A, Colon E, Albersmeier JP, Soder O. Male germ cell development in humans. Horm Res Paediatr. 2014;81(1):2-12.

4. Sadler TW. Langman's Medical Embryology. Baltimore: Lippincott Williams & Wilkins; 2010.

5. Ikawa M, Inoue N, Benham AM, Okabe M. Fertilization: a sperm's journey to and interaction with the oocyte. J Clin Invest. 2010 Apr;120(4):984-94.

6. Langley AR, Smith JC, Stemple DL, Harvey SA. New insights into the maternal to zygotic transition. Development. 2014 Oct;141(20):3834-41.

7. Veltman JA, Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. 2012 Aug;13(8):565-75.

8. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, et al. Rate of de novo mutations and the importance of father's age to disease risk. Nature. 2012 Aug 23;488(7412):471-5.

9. de Ligt J, Veltman JA, Vissers LE. Point mutations as a source of de novo genetic disease. Curr Opin Genet Dev. 2013 Jun;23(3):257-63.

10. Baggot LM. Human Reproduction. Cambridge: Cambridge University Press; 1997.

11. Schleich JM, Abdulla T, Summers R, Houyel L. An overview of cardiac morphogenesis. Arch Cardiovasc Dis. 2013 Nov;106(11):612-23.

12. Lindsey SE, Butcher JT, Yalcin HC. Mechanical regulation of cardiac development. Front Physiol. 2014;5:318.

13. Khoshnood B, Lelong N, Houyel L, Thieulin AC, Jouannic JM, Magnier S, et al. Prevalence, timing of diagnosis and mortality of newborns with congenital heart defects: a population-based study. Heart. 2012 Nov;98(22):1667-73.

14. Stevenson RE, Hunter AG. Considering the Embryopathogenesis of VACTERL Association. Mol Syndromol. 2013 Feb;4(1-2):7-15.

15. Shaw-Smith C. Oesophageal atresia, tracheo-oesophageal fistula, and the VACTERL association: review of genetics and epidemiology. J Med Genet. 2006 Jul;43(7):545-54.

16. Jacobs IJ, Que J. Genetic and cellular mechanisms of the formation of esophageal atresia and tracheoesophageal fistula. Dis Esophagus. 2013 May-Jun;26(4):356-8.

17. Fausett SR, Klingensmith J. Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus. Wiley Interdiscip Rev Dev Biol. 2012 Mar-Apr;1(2):184-202.

18. Levitt MA, Pena A. Anorectal malformations. Orphanet J Rare Dis. 2007;2:33.

19. Wijers CH, van Rooij IA, Marcelis CL, Brunner HG, de Blaauw I, Roeleveld N. Genetic and nongenetic etiology of nonsyndromic anorectal malformations: a systematic review. Birth Defects Res C Embryo Today. 2014 Dec;102(4):382-400.

20. Garcia-Barcelo MM, Wong KK, Lui VC, Yuan ZW, So MT, Ngan ES, et al. Identification of a HOXD13 mutation in a VACTERL patient. Am J Med Genet A. 2008 Dec 15;146A(24):3181-5.

21. Song R, Yosypiv IV. Genetics of congenital anomalies of the kidney and urinary tract. Pediatr Nephrol. 2011 Mar;26(3):353-64.

22. Arlen AM, Smith EA. Disorders of the bladder and cloacal anomaly. Clin Perinatol. 2014 Sep;41(3):695-707.

23. Hutson JM, Li R, Southwell BR, Newgreen D, Cousinery M. Regulation of testicular descent. Pediatr Surg Int. 2015 Apr;31(4):317-25.

24. Grimbizis GF, Camus M, Tarlatzis BC, Bontis JN, Devroey P. Clinical implications of uterine malformations and hysteroscopic treatment results. Hum Reprod Update. 2001 Mar-Apr;7(2):161-74.

25. Abrao MS, Muzii L, Marana R. Anatomical causes of female infertility and their management. Int J Gynaecol Obstet. 2013 Dec;123 Suppl 2:S18-24.

26. Nordenvall AS, Frisen L, Nordenstrom A, Lichtenstein P, Nordenskjold A. Population based nationwide study of hypospadias in Sweden, 1973 to 2009: incidence and risk factors. J Urol. 2014 Mar;191(3):783-9.

27. Stein R. Hypospadias. European Urology. 2012;11(2):33-45.

28. Hughes IA, Houk C, Ahmed SF, Lee PA. Consensus statement on management of intersex disorders. Arch Dis Child. 2006 Jul;91(7):554-63.

29. Hughes IA. Disorders of sex development: a new definition and classification. Best Pract Res Clin Endocrinol Metab. 2008 Feb;22(1):119-34.

30. Ostrer H. Disorders of sex development (DSDs): an update. J Clin Endocrinol Metab. 2014 May;99(5):1503-9.

31. Giampietro PF, Raggio CL, Blank RD, McCarty C, Broeckel U, Pickart MA. Clinical, genetic and environmental factors associated with congenital vertebral malformations. Mol Syndromol. 2013 Feb;4(1-2):94-105.

32. Furniss D, Kan SH, Taylor IB, Johnson D, Critchley PS, Giele HP, et al. Genetic screening of 202 individuals with congenital limb malformations and requiring reconstructive surgery. J Med Genet. 2009 Nov;46(11):730-5.

33. Ekblom AG, Laurell T, Arner M. Epidemiology of congenital upper limb anomalies in 562 children born in 1997 to 2007: a total population study from stockholm, sweden. J Hand Surg Am. 2010 Nov;35(11):1742-54.

34. Zuniga A, Zeller R, Probst S. The molecular basis of human congenital limb malformations. Wiley Interdiscip Rev Dev Biol. 2012 Nov-Dec;1(6):803-22.

35. Wigglesworth. Perinatal Pathology. J B, editor. Philadelphia: W.B. Saunders Company; 1984.

36. Rutledge JC. Developmental toxicity induced during early stages of mammalian embryogenesis. Mutat Res. 1997 Dec 12;396(1-2):113-27.

37. Cragan JD, Friedman JM, Holmes LB, Uhl K, Green NS, Riley L. Ensuring the safe and effective use of medications during pregnancy: planning and prevention through preconception care. Matern Child Health J. 2006 Sep;10(5 Suppl):S129-35.

38. Nordenskjold M, editor. Genetiska sjukdomar. 1st ed. Stockholm: Liber AB; 2011.

39. Socialstyrelsen. Fosterskador och kromosomavvikelser 2012: Sveriges officiella statistik, Hälso- och sjukvård2013.

40. Lee C, Scherer SW. The clinical context of copy number variation in the human genome. Expert Rev Mol Med. 2010;12:e8.

41. Forsgren M, Böttiger B, Johansen K, Karltorp E, Teär-Fahnehjelm K, Broliden K. Rubella (röda hund). Kunskapscentrum för infektioner under graviditet - INFPREG; 1998 [cited 2015]; Available from: http://www.medscinet.se/infpreg/specinfo/specinfo.asp.

42. Lewensohn-Fuchs I, Forsgren M, Petersson K, Engman M-L, Fischler B, Karltorp E, et al. Cytomegalovirus (CMV). Kunskapscentrum för infektioner under graviditet - INFPREG; 1998 [cited 2015]; Available from: http://www.medscinet.se/infpreg/specinfo/specinfo.asp.

43. Petersson K, Barrigan A, Engman M-L, Teär-Fahnehjelm K. Toxoplasma-infektion. Kunskapscentrum för infektioner under graviditet - INFPREG; 1998 [cited 2015]; Available from: http://www.medscinet.se/infpreg/specinfo/specinfo.asp.

44. Salbaum JM, Kappen C. Diabetic embryopathy: a role for the epigenome? Birth Defects Res A Clin Mol Teratol. 2011 Aug;91(8):770-80.

45. Vereczkey A, Gerencser B, Czeizel AE, Szabo I. Association of certain chronic maternal diseases with the risk of specific congenital heart defects: a population-based study. Eur J Obstet Gynecol Reprod Biol. 2014 Nov;182:1-6.

46. van Gelder MM, van Rooij IA, Miller RK, Zielhuis GA, de Jong-van den Berg LT, Roeleveld N. Teratogenic mechanisms of medical drugs. Hum Reprod Update. 2010 Jul-Aug;16(4):378-94.

47. Harden CL, Meador KJ, Pennell PB, Hauser WA, Gronseth GS, French JA, et al. Management issues for women with epilepsy-Focus on pregnancy (an evidence-based review): II. Teratogenesis and perinatal outcomes: Report of the Quality Standards Subcommittee and Therapeutics and Technology Subcommittee of the American Academy of Neurology and the American Epilepsy Society. Epilepsia. 2009 May;50(5):1237-46.

48. Diav-Citrin O, Shechtman S, Tahover E, Finkel-Pekarsky V, Arnon J, Kennedy D, et al. Pregnancy outcome following in utero exposure to lithium: a prospective, comparative, observational study. Am J Psychiatry. 2014 Jul;171(7):785-94.

49. Sandberg T, Ladfors L. Antibiotika under graviditet. Kunskapscentrum för infektioner under graviditet - INFPREG; 1998 [cited 2015]; Available from: http://www.medscinet.se/infpreg/specinfo/specinfo.asp.

50. Hackshaw A, Rodeck C, Boniface S. Maternal smoking in pregnancy and birth defects: a systematic review based on 173 687 malformed cases and 11.7 million controls. Hum Reprod Update. 2011 Sep-Oct;17(5):589-604.

51. Simpson JL. Birth defects and assisted reproductive technologies. Semin Fetal Neonatal Med. 2014 Jun;19(3):177-82.

52. McCollough CH, Schueler BA, Atwell TD, Braun NN, Regner DM, Brown DL, et al. Radiation exposure and pregnancy: when should we be concerned? Radiographics. 2007 Jul-Aug;27(4):909-17; discussion 17-8.

53. Nora JJ. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. Circulation. 1968 Sep;38(3):604-17.

54. Smith DW, Aase JM. Polygenic inheritance of certain common malformations. Evidence and empiric recurrence risk data. J Pediatr. 1970 May;76(5):652-9.

55. Weischenfeldt J, Symmons O, Spitz F, Korbel JO. Phenotypic impact of genomic structural variation: insights from and for human disease. Nat Rev Genet. 2013 Feb;14(2):125-38.

56. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nat Rev Genet. 2006 Feb;7(2):85-97.

57. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat Genet. 2004 Sep;36(9):949-51.

58. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong association of de novo copy number mutations with autism. Science. 2007 Apr 20;316(5823):445-9.

59. McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet. 2008 Oct;40(10):1166-74.

60. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature. 2006 Nov 23;444(7118):444-54.

61. Henrichsen CN, Vinckenbosch N, Zollner S, Chaignat E, Pradervand S, Schutz F, et al. Segmental copy number variation shapes tissue transcriptomes. Nat Genet. 2009 Apr;41(4):424-9.

62. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 2009 Apr;84(4):524-33.

63. Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet. 2009;10:451-81.

64. Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet. 2005 Dec;1(6):e49.

65. Lu XY, Phung MT, Shaw CA, Pham K, Neil SE, Patel A, et al. Genomic imbalances in neonates with birth defects: high detection rates by using chromosomal microarray analysis. Pediatrics. 2008 Dec;122(6):1310-8.

66. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. Genet Med. 2011 Jul;13(7):680-5.

67. Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, et al. De novo rates and selection of large copy number variation. Genome Res. 2010 Nov;20(11):1469-81.

68. Solomon BD. VACTERL/VATER Association. Orphanet J Rare Dis. 2011;6:56.

69. Quan L, Smith DW. The VATER association. Vertebral defects, Anal atresia, T-E fistula with esophageal atresia, Radial and Renal dysplasia: a spectrum of associated defects. J Pediatr. 1973 Jan;82(1):104-7.

70. Reutter H, Ludwig M. VATER/VACTERL Association: Evidence for the Role of Genetic Factors. Mol Syndromol. 2013 Feb;4(1-2):16-9.

71. Brosens E, Eussen H, van Bever Y, van der Helm RM, Ijsselstijn H, Zaveri HP, et al. VACTERL Association Etiology: The Impact of de novo and Rare Copy Number Variations. Mol Syndromol. 2013 Feb;4(1-2):20-6.

72. Solomon BD, Pineda-Alvarez DE, Raam MS, Cummings DA. Evidence for inheritance in patients with VACTERL association. Hum Genet. 2010 Jun;127(6):731-3.

73. Kim JH, Kim PCW, Hui CC. The VACTERL association: lessons from the Sonic hedgehog pathway. Clin Genet. [Article]. 2001 May;59(5):306-15.

74. Szumska D, Pieles G, Essalmani R, Bilski M, Mesnard D, Kaur K, et al. VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase Pcsk5. Genes Dev. 2008 Jun 1;22(11):1465-77.

75. Reardon W, Zhou XP, Eng C. A novel germline mutation of the PTEN gene in a patient with macrocephaly, ventricular dilatation, and features of VATER association. Journal of Medical Genetics. 2001 Dec;38(12):820-3.

76. Alter BP, Rosenberg PS. VACTERL-H Association and Fanconi Anemia. Mol Syndromol. 2013 Feb;4(1-2):87-93.

77. Castori M. Diabetic embryopathy: a developmental perspective from fertilization to adulthood. Mol Syndromol. 2013 Feb;4(1-2):74-86.

78. Herion NJ, Salbaum JM, Kappen C. Traffic jam in the primitive streak: the role of defective mesoderm migration in birth defects. Birth Defects Res A Clin Mol Teratol. 2014 Aug;100(8):608-22.

79. Piotrowski A, Bruder CE, Andersson R, de Stahl TD, Menzel U, Sandgren J, et al. Somatic mosaicism for copy number variation in differentiated human tissues. Hum Mutat. 2008 Sep;29(9):1118-24.

80. Vanneste E, Voet T, Le Caignec C, Ampe M, Konings P, Melotte C, et al. Chromosome instability is common in human cleavage-stage embryos. Nat Med. 2009 May;15(5):577-83.

81. Mertzanidou A, Wilton L, Cheng J, Spits C, Vanneste E, Moreau Y, et al. Microarray analysis reveals abnormal chromosomal complements in over 70%

of 14 normally developing human embryos. Hum Reprod. 2013 Jan;28(1):256-64.

82. Dekel-Naftali M, Aviram-Goldring A, Litmanovitch T, Shamash J, Yonath H, Hourvitz A, et al. Chromosomal integrity of human preimplantation embryos at different days post fertilization. J Assist Reprod Genet. 2013 Jun;30(5):633-48.

83. Golubovsky MD. Postzygotic diploidization of triploids as a source of unusual cases of mosaicism, chimerism and twinning. Hum Reprod. 2003 Feb;18(2):236-42.

84. Kuliev A, Verlinsky Y. Meiotic and mitotic nondisjunction: lessons from preimplantation genetic diagnosis. Hum Reprod Update. 2004 Sep-Oct;10(5):401-7.

85. Erickson RP. Somatic gene mutation and human disease other than cancer: An update. Mutat Res. 2010 Oct;705(2):96-106.

86. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. Nat Rev Genet. 2013 May;14(5):307-20.

87. Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, et al. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. Am J Hum Genet. 2014 Aug 7;95(2):173-82.

88. Gammon A, Jasperson K, Pilarski R, Prior T, Kuwada S. PTEN mosaicism with features of Cowden syndrome. Clin Genet. 2013 Dec;84(6):593-5.

89. Freed D, Stevens EL, Pevsner J. Somatic mosaicism in the human genome. Genes (Basel). 2014;5(4):1064-94.

90. Buschiazzo E, Gemmell NJ. Conservation of human microsatellites across 450 million years of evolution. Genome Biol Evol. 2010;2:153-65.

91. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. American Journal of Human Genetics. [Article]. 2010 May;86(5):749-64.

92. Vandeweyer G, Kooy RF. Detection and interpretation of genomic structural variation in health and disease. Expert Rev Mol Diagn. 2013 Jan;13(1):61-82.

93. Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg A, Peterson C. BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. Genome Biol. 2002 Jul 15;3(8):SOFTWARE0003.

94. Serra-Juhe C, Rodriguez-Santiago B, Cusco I, Vendrell T, Camats N, Toran N, et al. Contribution of rare copy number variants to isolated human malformations. PLoS One. 2012;7(10):e45530.

95. White PS, Xie HM, Werner P, Glessner J, Latney B, Hakonarson H, et al. Analysis of chromosomal structural variation in patients with congenital left-sided cardiac lesions. Birth Defects Res A Clin Mol Teratol. 2014 Dec;100(12):951-64.

96. Girirajan S, Campbell CD, Eichler EE. Human copy number variation and complex genetic disease. Annu Rev Genet. 2011;45:203-26.

97. Yuan S, Zaidi S, Brueckner M. Congenital heart disease: emerging themes linking genetics and development. Curr Opin Genet Dev. 2013 Jun;23(3):352-9.

98. Warburton D, Ronemus M, Kline J, Jobanputra V, Williams I, Anyane-Yeboa K, et al. The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. Hum Genet. 2014 Jan;133(1):11-27.

99. Breckpot J, Thienpont B, Peeters H, de Ravel T, Singer A, Rayyan M, et al. Array comparative genomic hybridization as a diagnostic tool for syndromic heart defects. J Pediatr. 2010 May;156(5):810-7, 7 e1-7 e4.

100. Silversides CK, Lionel AC, Costain G, Merico D, Migita O, Liu B, et al. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. PLoS Genet. 2012;8(8):e1002843.

101. Lalani SR, Shaw C, Wang X, Patel A, Patterson LW, Kolodziejska K, et al. Rare DNA copy number variants in cardiovascular malformations with extracardiac abnormalities. Eur J Hum Genet. 2013 Feb;21(2):173-81.

102. Yan Y, Wu Q, Zhang L, Wang X, Dan S, Deng D, et al. Detection of submicroscopic chromosomal aberrations by array-based comparative genomic hybridization in fetuses with congenital heart disease. Ultrasound Obstet Gynecol. 2014 Apr;43(4):404-12.

103. Brosens E, Ploeg M, van Bever Y, Koopmans AE, H IJ, Rottier RJ, et al. Clinical and etiological heterogeneity in patients with tracheo-esophageal malformations and associated anomalies. Eur J Med Genet. 2014 Aug;57(8):440-52.

104. Veenma D, Brosens E, de Jong E, van de Ven C, Meeussen C, Cohen-Overbeek T, et al. Copy number detection in discordant monozygotic twins of Congenital Diaphragmatic Hernia (CDH) and Esophageal Atresia (EA) cohorts. Eur J Hum Genet. 2012 Mar;20(3):298-304.

105. Dworschak GC, Draaken M, Hilger AC, Schramm C, Bartels E, Schmiedeke E, et al. Genome-wide mapping of copy number variations in patients with both anorectal malformations and central nervous system abnormalities. Birth Defects Res A Clin Mol Teratol. 2014 Sep 24.

106. Hinton RB, Martin LJ, Rame-Gowda S, Tabangin ME, Cripe LH, Benson DW. Hypoplastic left heart syndrome links to chromosomes 10q and 6q and is genetically related to bicuspid aortic valve. J Am Coll Cardiol. 2009 Mar 24;53(12):1065-71.

107. Barth JL, Clark CD, Fresco VM, Knoll EP, Lee B, Argraves WS, et al. Jarid2 is among a set of genes differentially regulated by Nkx2.5 during outflow tract morphogenesis. Dev Dyn. 2010 Jul;239(7):2024-33.

108. Rutsch F, Ruf N, Vaingankar S, Toliat MR, Suk A, Hohne W, et al. Mutations in ENPP1 are associated with 'idiopathic' infantile arterial calcification. Nat Genet. 2003 Aug;34(4):379-81.

109. Bollen M, Gijsbers R, Ceulemans H, Stalmans W, Stefan C. Nucleotide pyrophosphatases/phosphodiesterases on the move. Crit Rev Biochem Mol Biol. 2000;35(6):393-432.

110. Schiano C, Casamassimi A, Vietri MT, Rienzo M, Napoli C. The roles of mediator complex in cardiovascular diseases. Biochim Biophys Acta. 2014 Jun;1839(6):444-51.

111. Balamotis MA, Pennella MA, Stevens JL, Wasylyk B, Belmont AS, Berk AJ. Complexity in transcription control at the activation domain-mediator interface. Sci Signal. 2009;2(69):ra20.

112. Jenkinson CP, Grody WW, Cederbaum SD. Comparative properties of arginases. Comp Biochem Physiol B Biochem Mol Biol. 1996 May;114(1):107-32.

113. Breckpot J, Thienpont B, Arens Y, Tranchevent LC, Vermeesch JR, Moreau Y, et al. Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. Cytogenet Genome Res. 2011;135(3-4):251-9.

114. Bednarczyk D, Makowska I, Sasiadek MM, Smigiel R. Somatic mosaicism in esophageal atresia. Am J Gastroenterol. 2014 Dec;109(12):1954-6.

115. Bernardini L, Giuffrida MG, Francalanci P, Capalbo A, Novelli A, Callea F, et al. X chromosome monosomy restricted to the left ventricle is not a major cause of isolated hypoplastic left heart. Am J Med Genet A. 2010 Aug;152A(8):1967-72.

116. Jianrong L, Yinglong L, Xiaodong L, Cuntao Y, Bin C, Bo W. 22q11.2 deletion mosaicism in patients with conotruncal heart defects. Birth Defects Res A Clin Mol Teratol. 2006 Apr;76(4):262-5.

117. Rauch A, Hofbeck M, Cesnjevar R, Koch A, Rauch R, Buheitel G, et al. Search for somatic 22q11.2 deletions in patients with conotruncal heart defects. Am J Med Genet A. 2004 Jan 15;124A(2):165-9.

118. Reamon-Buettner SM, Borlak J. Somatic mutations in cardiac malformations. J Med Genet. 2006 Aug;43(8):e45.

119. Reamon-Buettner SM, Borlak J. HEY2 mutations in malformed hearts. Hum Mutat. 2006 Jan;27(1):118.

120. Reamon-Buettner SM, Borlak J. Somatic NKX2-5 mutations as a novel mechanism of disease in complex congenital heart disease. J Med Genet. 2004 Sep;41(9):684-90.

121. Reamon-Buettner SM, Borlak J. GATA4 zinc finger mutations as a molecular rationale for septation defects of the human heart. J Med Genet. 2005 May;42(5):e32.

122. Reamon-Buettner SM, Borlak J. TBX5 mutations in non-Holt-Oram syndrome (HOS) malformed hearts. Hum Mutat. 2004 Jul;24(1):104.

123. Reamon-Buettner SM, Cho SH, Borlak J. Mutations in the 3'-untranslated region of GATA4 as molecular hotspots for congenital heart disease (CHD). BMC Med Genet. 2007;8:38.

124. Reamon-Buettner SM, Ciribilli Y, Inga A, Borlak J. A loss-of-function mutation in the binding domain of HAND1 predicts hypoplasia of the human hearts. Hum Mol Genet. 2008 May 15;17(10):1397-405.

125. Reamon-Buettner SM, Ciribilli Y, Traverso I, Kuhls B, Inga A, Borlak J. A functional genetic study identifies HAND1 mutations in septation defects of the human heart. Hum Mol Genet. 2009 Oct 1;18(19):3567-78.

126. Reamon-Buettner SM, Hecker H, Spanel-Borowski K, Craatz S, Kuenzel E, Borlak J. Novel NKX2-5 mutations in diseased heart tissues of patients with cardiac malformations. Am J Pathol. 2004 Jun;164(6):2117-25.

127. Esposito G, Butler TL, Blue GM, Cole AD, Sholler GF, Kirk EP, et al. Somatic mutations in NKX2-5, GATA4, and HAND1 are not a common cause of tetralogy of Fallot or hypoplastic left heart. Am J Med Genet A. 2011 Oct;155A(10):2416-21.

128. Salazar M, Consoli F, Villegas V, Caicedo V, Maddaloni V, Daniele P, et al. Search of somatic GATA4 and NKX2.5 gene mutations in sporadic septal heart defects. Eur J Med Genet. 2011 May-Jun;54(3):306-9.

129. Draus JM, Jr., Hauck MA, Goetsch M, Austin EH, 3rd, Tomita-Mitchell A, Mitchell ME. Investigation of somatic NKX2-5 mutations in congenital heart disease. J Med Genet. 2009 Feb;46(2):115-22.

130. Huang RT, Xue S, Xu YJ, Yang YQ. Somatic mutations in the GATA6 gene underlie sporadic tetralogy of Fallot. Int J Mol Med. 2013 Jan;31(1):51-8.

131. McBride KL, Zender GA, Fitzgerald-Butt SM, Koehler D, Menesses-Diaz A, Fernbach S, et al. Linkage analysis of left ventricular outflow tract malformations (aortic valve stenosis, coarctation of the aorta, and hypoplastic left heart syndrome). Eur J Hum Genet. 2009 Jun;17(6):811-9.

132. Pham J, Shaw C, Pursley A, Hixson P, Sampath S, Roney E, et al. Somatic mosaicism detected by exon-targeted, high-resolution aCGH in 10,362 consecutive cases. Eur J Hum Genet. 2014 Aug;22(8):969-78.

133. Cohen AS, Wilson SL, Trinh J, Ye XC. Detecting somatic mosaicism: considerations and clinical implications. Clin Genet. 2014 Sep 15.

134. Shinawi M, Shao L, Jeng LJ, Shaw CA, Patel A, Bacino C, et al. Low-level mosaicism of trisomy 14: phenotypic and molecular characterization. Am J Med Genet A. 2008 Jun 1;146A(11):1395-405.

135. Spinner NB, Conlin LK. Mosaicism and clinical genetics. Am J Med Genet C Semin Med Genet. 2014 Dec;166C(4):397-405.

136. Verloes A. Trisomy 13. 2008 [cited 2015]; Available from: http://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=EN&data_id=337&Disease_Disease_Search_diseaseGroup=patau&Disease_Disease_Search_diseaseType=Pat&Disease(s)/group of diseases=Trisomy-13--Patau-syndrome-&title=Trisomy-13--Patau-syndrome-&search=Disease_Search_Simple.

137. Verloes A. Trisomy 18. 2008 [cited 2015]; Available from: http://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=EN&data_id=339&Disease_Disease_Search_diseaseGroup=trisomy-18&Disease_Disease_Search_diseaseType=Pat&Disease(s)/group of diseases=Trisomy-18&title=Trisomy-18&search=Disease_Search_Simple.

138. Hook EB. Exclusion of chromosomal mosaicism: tables of 90%, 95% and 99% confidence limits and comments on use. Am J Hum Genet. 1977 Jan;29(1):94-7.

139. Costello JP, Weiderhold A, Louis C, Shaughnessy C, Peer SM, Zurakowski D, et al. A Contemporary, Single-Institutional Experience of Surgical Versus Expectant Management of Congenital Heart Disease in Trisomy 13 and 18 Patients. Pediatr Cardiol. 2015 Jan 23.

140. Schreck RR, Breg WR, Erlanger BF, Miller OJ. Preferential derivation of abnormal human G-group-like chromosomes from chromosome 15. Hum Genet. 1977 Apr 7;36(1):1-12.

141. Rossi E, Giorda R, Bonaglia MC, Candia SD, Grechi E, Franzese A, et al. De novo unbalanced translocations in Prader-Willi and Angelman syndrome might be the reciprocal product of inv dup(15)s. PLoS One. 2012;7(6):e39180.

142. Wang CH, Chen IH, Kuo MW, Su PT, Lai ZY, Huang WC, et al. Zebrafish Thsd7a is a neural protein required for angiogenic patterning during development. Dev Dyn. 2011 Jun;240(6):1412-21.

143. Huang Q, Zhang L, Wang Y, Zhang C, Zhou S, Yang G, et al. Depletion of PHF14, a novel histone-binding protein gene, causes neonatal lethality in mice due to respiratory failure. Acta Biochim Biophys Sin (Shanghai). 2013 Aug;45(8):622-33.

144. Strain L, Warner JP, Johnston T, Bonthron DT. A human parthenogenetic chimaera. Nat Genet. 1995 Oct;11(2):164-9.

145. Hilger A, Schramm C, Pennimpede T, Wittler L, Dworschak GC, Bartels E, et al. De novo microduplications at 1q41, 2q37.3, and 8q24.3 in patients with VATER/VACTERL association. Eur J Hum Genet. 2013 Apr 3.

146. Dworschak GC, Draaken M, Marcelis C, de Blaauw I, Pfundt R, van Rooij IA, et al. De novo 13q deletions in two patients with mild anorectal malformations as part of VATER/VACTERL and VATER/VACTERL-like association and analysis of EFNB2 in patients with anorectal malformations. Am J Med Genet A. 2013 Dec;161A(12):3035-41.

147. Schramm C, Draaken M, Bartels E, Boemers TM, Aretz S, Brockschmidt FF, et al. De novo microduplication at 22q11.21 in a patient with VACTERL association. Eur J Med Genet. 2010 Sep 16.

148. Hernandez-Garcia A, Brosens E, Zaveri HP, de Jong EM, Yu Z, Namwanje M, et al. Contribution of LPP copy number and sequence changes to esophageal atresia, tracheoesophageal fistula, and VACTERL association. Am J Med Genet A. 2012 Jul;158A(7):1785-7.

149. Solomon BD, Pineda-Alvarez DE, Hadley DW, Hansen NF, Kamat A, Donovan FX, et al. Exome Sequencing and High-Density Microarray Testing in Monozygotic Twin Pairs Discordant for Features of VACTERL Association. Mol Syndromol. 2013 Feb;4(1-2):27-31.

150. Saisawat P, Kohl S, Hilger AC, Hwang DY, Yung Gee H, Dworschak GC, et al. Whole-exome resequencing reveals recessive mutations in TRAP1 in individuals with CAKUT and VACTERL association. Kidney Int. 2014 Jun;85(6):1310-7.

151. Pritchard CC, Smith C, Marushchak T, Koehler K, Holmes H, Raskind W, et al. A mosaic PTEN mutation causing Cowden syndrome identified by deep sequencing. Genet Med. 2013 Dec;15(12):1004-7.

152. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, et al. De novo mutations in histone-modifying genes in congenital heart disease. Nature. 2013 Jun 13;498(7453):220-3.

153. de Jong EM, Felix JF, Deurloo JA, van Dooren MF, Aronson DC, Torfs CP, et al. Non-VACTERL-type anomalies are frequent in patients with esophageal atresia/tracheo-esophageal fistula and full or partial VACTERL association. Birth Defects Res A Clin Mol Teratol. 2008 Feb;82(2):92-7.