Genetic, genomic and epigenetic alterations in congenital malformations: implications in genetic counseling

Clara Serra Juhé

DOCTORAL THESIS UPF / 2012

THESIS SUPERVISOR

Prof. Luis A. Pérez Jurado

Departament de Ciències Experimentals i de la Salut

UNIVERSITAT POMPEU FABRA

Als meus pares,
a l'Alba, en Pau i l'Aniol,
a en Carles

ACKNOWLEDGEMENTS

Al llarg d'aquests anys de doctorat, he tingut la sort de conèixer i d'aprendre de moltíssimes persones i m'agradaria no deixar-me'n cap en aquests agraïments.

En primer lloc m'agradaria agrair a en Luis, no només la oportunitat que em va donar de fer el doctorat, sinó també la oportunitat de conèixer, acabant la carrera, el món de la genètica, que em va fascinar des del primer moment. Per tot el que he pogut aprendre durant aquests anys i per la seva dedicació.

A tot el grup de genètica de la UPF (Mariví, Ivon, Cris, Tina, Marta, Aïda, Raquel, Débora, Maria, Gaby, Andreu, Fátima, Verena, Olaya, Benja, Anna...) per haver col·laborat de manera MOLT important a què aquests anys de tesi hagin estat fantàstics entre sopars, make-up parties, partits de vòlei i tantes altres coses que hem compartit!

A en Benja, per ser el meu primer mestre al laboratori. Per ensenyar-me a fer MLPA i (intentar) ensenyar-me a fer anar l'R però, sobretot, per estar sempre disposat a donar un cop de mà, fins i tot el dia abans de marxar de vacances!

A la Ivon, per tenir tanta paciència com per a contestar totes les preguntes i dubtes de tot aquest temps i per ajudar-me molt els meus primers mesos al laboratori, intentant esbrinar què se suposava que havia de fer amb tants arrays...

A la Raquel, per tota l'ajuda durant aquests últims mesos i per tot el que m'ha ajudat a descobrir a part de la ciència (com trobarem a faltar l'Esprit!).

A l'Olaya, per tots els bons consells sobre la tesi, per animar-me i ajudar-me.

A la Maria, per tocar sempre de peus a terra i fer-nos veure tot sovint que tot s'ha de relativitzar… I pels seus ànims i ajuda (ai, les quantis…) durant aquests últims dies.

A la Gaby, per haver-me fet entendre la importància dels "brownie points" i per les llargues xerrades a la terrassa i al passadís. Per totes les correccions de l'anglès i per intentar aprendre plegades l'embriologia del cor...

A l'Aïda, per totes les hores barallant-nos amb els MLPAs de metilació i amb el Genome Studio, per ser tan detallista, per la visita a Nijmegen.

A Cris, por ser la mejor compañera de mesa que se puede tener! Y por enseñarme a recitar una secuencia mientras leía otra y escribía la complementaria…

A la Irene, per haver estat (i encara ser!) la millor companya que em podria haver buscat en el llarg camí de la genètica, des que a segon vam sentir a parlar de l'assessorament genètic i ens va cridar l'atenció. Per tots els treballs, ABPS i classes que hem compartit però, sobretot, per la gran amistat que n'hem tret.

A tot l'equip de genètica de Vall d'Hebron (a la Teresa, en Miguel, l'Alberto, l'Asun, l'Anna, la Mercè, la Beth, la Mati, la Leo, la Carmen, la Neus, la Montse, la Beni…) per haver-nos acollit tan bé en les primers pràctiques que hi vam fer amb la Irene i per totes les que van seguir. I especialment a la Teresa per tota l'ajuda durant la tesi.

Als Q-genomics (en Manel, la Cristina, la Sonia, en Lluís, en Xavi, la Ma Jesús), perquè tot i estar al "lado oscuro" i tenir pipetes més guais que les nostres són uns grans companys de laboratori.

A la junta de SEAGen, especialment a l'Estela, perquè sense les nostres reunions aquests últims temps no haguessin estat el mateix!

I also would like to thank all the people in the Genomic Disorders in Nijmegen (Michael, Peer, Ilse, Christian, Konny, Irene, Lisenka, Marloes, Petra, Thessa, Sabine...). It was really nice to work with all of you and I learned a lot (about exome sequencing and about The Netherlands, after travelling a lot!). I would like to thank Joris for the opportunity to go there and for his comments and suggestions writing the thesis. And specially, I would like to thank Alex for his help before going there, for the six months in Nijmegen and also for his contribution and support writing the thesis. It was gezellig!

I a en Joan Marc, la Laia, en Joan, l'Eva, la Carme, la Marta, l'Anabel, l'Anna, la Mireia, l'Ivan... per totes les estones divertides al llarg d'aquests anys. A l'Isa, per tots els anys compartint pis, per les trucades al tren i per les tardes de rebaixes. I també a en Jesús pels vespre a Laietana.

A la Rosa Ma i en Pepe, per acollir-nos tot sovint a Barcelona, i també a Canet, per cuidar-nos tan bé i per fer-me croquetes!

Als meus pares, per haver-nos donat la oportunitat d'aprendre i formar-nos i per haver-nos ensenyat a valorar-ho, i pel seu recolzament en tots els reptes que emprenem. A l'Alba, en Bosco, en Pau, la Laura i l'Aniol per totes les estones junts, els caps de setmana a l'Armentera, els dinars a Salt, Badalona, el Carmel o Caldes (on sigui per trobar-nos!) i pels tips de riure que ens fem quan estem plegats. A la Berta, per ser la nebodeta més maca i riallera que es pot tenir, i a l'Aleix que està a punt d'arribar. I als tius Pitus i a l'àvia.

I finalment a en Carles, pel seu suport incondicional durant aquests quatre anys (que no sempre ha estat fàcil!). Per no dubtar en marxar cap a Nijmegen tot i el fred, els frikandels i la falta de muntanyes... Per la seva paciència durant aquests últims quinze dies i per tantes altres coses.

Moltes gràcies a tots!

ABSTRACT

Mechanisms underlying congenital malformations are largely unknown despite its high incidence, affecting 2-3% of liveborn infants. A broader knowledge about the causes of birth defects would provide valuable information regarding the outcome and prognosis of the anomaly, the development and establishment of diagnostic protocols, the design of therapeutic strategies and genetic counseling to the family. Different approaches have been used in the present thesis regarding technologies and model diseases to elucidate the contribution of genetic and epigenetic alterations in the etiopathogenesis of congenital malformations. Copy number variations, methylation patterns, as well as point mutations have been explored. Moreover, a study to analyze genetic counseling in relation to one of the new molecular techniques used has been performed. Obtained data reveal a relevant role of genetic and epigenetic alterations in congenital malformations, in some cases as a unique cause to explain the disease and in others as part of an oligogenic or multifactorial model.

RESUM

Els mecanismes causants de les malformacions congènites són poc coneguts malgrat l'elevada incidència d'aquestes patologies, que afecten el 2-3% de recent nascuts. Un coneixement més ampli de les causes de les anomalies congènites proporcionaria informació rellevant pel que fa a al pronòstic de l'anomalia, el desenvolupament i establiment de protocols diagnòstics, el disseny d'estratègies terapèutiques, així com l'assessorament genètic a la família. En la tesi que es presenta s'han utilitzat diferents estratègies, pel que fa a tecnologies i models de malalties, amb l'objectiu d'esbrinar la contribució d'alteracions genètiques i epigenètiques en l'etiopatogènia de les malformacions congènites. S'han analitzat variacions en número de còpia, patrons de metilació, així com mutacions puntuals. D'altra banda, també s'ha realitzat un estudi per aprofundir en l'assessorament genètic en relació a una de les noves tècniques moleculars utilitzades. Els resultats obtinguts indiquen que les altercacions genètiques i epigenètiques tenen una contribució molt rellevant en l'etiologia de les malformacions congènites, en alguns casos com a causa única de la malaltia i en altres com a component d'un model oligogènic o multifactorial.

PROLOGUE

The eclosion in recent years of new molecular techniques in genetic and epigenetic research has provided an amazingly large amount of data regarding the genome and its alterations leading to disease. The application of those techniques in several disorders has supplied valuable information that contributes to the understanding of genetic and epigenetic mechanisms involved in the etiopathogenesis of several diseases.

This thesis presents the results of the application of new molecular techniques, such as chromosomal microarray analysis, exome sequencing and methylation arrays in the study of congenital malformations, a group of severe disorders affecting 2-3% of liveborn infants. The clinical application of one of those techniques and its implications in genetic counseling was also explored due to the relevance of translating research into clinical practice.

This thesis is divided in several chapters following the classical structure.

In the **introduction** a general overview regarding congenital malformations, possible causes for the disorder and model diseases used in this thesis are included, as well as a brief description of those techniques.

In the main body of the thesis, the **articles** describing the different studies and approaches used are included, as well as a short explanation of the reasons to design each study.

A general **discussion** contrasting all the results obtained and possible interpretations is provided. Following the discussion and as a final chapter, the main **conclusions** of the thesis are summarized.

CONTENTS

LIST OF FIGURES

Introduction

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

LIST OF TABLES

Introduction

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

INTRODUCTION

1 CONGENITAL MALFORMATIONS

The term anomaly or defect refers to structural abnormalities that, when present at birth, are called congenital anomalies or birth defects. Anomalies and defects may be classified into two main categories:

-Major defect: A structural abnormality that requires medical and/or surgical treatment or creates significant problems for the patient. Neural tube defects or renal dysplasia are examples of major defects.

-Minor defect: A structural abnormality that does not require medical or surgical treatment but might be an indicator of some problem during development. Minor defects can also be defined as the features that vary from those that are most commonly seen in the population but do not increase morbidity. Hypotelorism and single palmar crease are minor defects.

When a major defect is detected, it is relevant to search for associated minor defects in order to define whether the major defect is part of a more complex syndrome affecting more organs or systems.

Four categories may be considered for major birth defects (1, 2):

-Malformation: fetal growth and development do not proceed normally due to underlying genetic, epigenetic or environmental factors which alter a particular structure. Holoprosencephaly and transposition of great arteries are examples of malformations in the central nervous system and heart, respectively.

-Deformation: an abnormal external force during in utero development results in abnormal growth or formation of a fetal structure. For example, fetuses that grow in a uterine environment where not enough amniotic fluid is present may have a flattened face due to compression of the face against the uterine wall.

-Disruption: structural defect caused by the destruction of some previously well-develop organ or tissue due to a disrupting event. The most common disruptions are those caused by amniotic bands where a digit or extremity growing normally is disrupted due to the development of amniotic band at the end of that extremity. This may result in missing fingers, toes, hands or feet depending on the localization of the amniotic band.

-Dysplasia: abnormal tissue structure because of an alteration in size, shape and/or organization of the cells. Many skeletal syndromes and some renal diseases such as multicystic kidney are due to dysplasia in the developing bone or renal tissue.

It is important to properly classify major defects into the four mentioned categories because the causes underlying them are substantially different. Deformations and disruptions do not have a genetic origin, although they can be the consequence of a genetic disease. As mentioned before, a low quantity of amniotic fluid may cause some deformations in fetuses; oligohydramnios could be caused by a genetic disease like multicystic kidney. Otherwise, dysplasies and malformations might have a genetic cause (1, 2). From now on, the term congenital malformation will be used referred to major birth defects with a potentially genetic basis included in two of the categories described above, malformation as well as dysplasies.

Congenital malformations often occur in the setting of multiple congenital anomalies, including dysmorphic facial features, or in association with different organ malformations, developmental abnormalities, or growth abnormalities (3, 4). When a combination of multiple birth defects occurs together it is known as a polymalformation syndrome. On the other hand, in an important proportion of cases, only one malformation is detected without the presentation of other minor or major defects.

1.1 Incidence

A potentially lethal or disabling major defect occurs in 2–3% of liveborn infants (5). Congenital malformations have become the main cause of infant mortality during the first years of life (6) and are associated with long term morbidity (7, 8).

In the next table, most common major malformations and their incidence in Spain between 1980 and 1995 are shown (9).

The most common congenital malformations are heart defects, with a global incidence of 8/1.000 livebirths, and neural tube defects, with an incidence of 1/1.000.

	RACAV	RDCV	RDCA	REDCB	EMCEMC	EUROCAT
Anencephaly	4.6	4.4	5.6	4.1	2.6	3.6
Spina bifida	4.8	3.5	7	4.7	4.1	5.1
Encephalocele	1.2	0.6	0.5	0.4	1	1.3
Hydrocephaly	5	1.5	4.4	9	4.5	5
Microtia	0.4	-	0.5	0.4	1.5	0.9
An / microphthalmia	1.2	0.6	2.5	1.4	2.1	1.6
Congenital cataract	1.6	0.9	2.1	0.4	-	1
Left heart hypoplasia	2.4	1.2	1.9	1.7	-	2.4
Transposition of great arteries	4.4	1.2	3.2	5.7	-	3.7
Truncus	0.3	0.9	0.9	1	-	0.9
Cleft lip (and / or palate)	5.6	5.6	8.6	4.3	5.7	8.9
Cleft palate (not cleft lip)	5.1	2.6	6	4.1	4.9	6.4
Esophageal atresia	2.5	0.9	4.9	2.1	2.1	2.7
Small Intestinal Atresia and Stenosis	2	1.2	2.8	1.4	-	2.3
Atresia / stenosis of anus and / or rectum	3.2	2.3	3.9	2.1	2.3	3.3
Hypospadias	8.4	15.2	14.8	9	18.1	9.8
Renal agenesia/hypoplasia	11.3	5	6.3	3.7	-	5
Renal displasia (except for hydronephrosis)	5.3	4.4	6.3	2.9	-	4.7
Limb reduction (not dysplasia)	5.1	6.7	9.5	5.7	6.9	6.1
Diaphragmatic hernia	3.2	1.2	4.9	3.7	2.4	3.5
Omphalocele	3.3	1.5	1.6	2.1	1.4	2.5
Gastroschisis	0.5	0.3	0.7	0.6	0.4	1.3
Down syndrome	20.3	22	14.6	17.9	13.7	16.8

Table 1. Incidence per 10,000 liveborns of the most common malformations reported in Spain between 1980 and 1995 by the registers for congenital malformations existing in the conuntry.

1.2 The relevance of identifying the cause of congenital defects

An increase of knowledge of the etiology of congenital malformations would provide valuable insight about the pathophysiology of these developmental anomalies. Establishing genotype-phenotype correlations may provide information about the outcome and prognosis of the anomaly and could also be relevant in order to design therapeutic strategies. In addition, the identification of disease causing alterations might be essential for the development and establishment of diagnostic procedures and strategies in similar patients to ensure efficient diagnostic protocols. As an example, some birth defects can be detected during pregnancy and, when it occurs, identifying the cause of those alterations is relevant to monitor the gestation, to determine the fetal viability and / or to anticipate medical problems at birth or later.

Moreover, the identification of the basis of genetic diseases is essential to provide an accurate definition of disease risk, a critical element to ensure proper genetic counseling and disease prevention. Genetic counseling has become more relevant in this area considering that more individuals with congenital malformations are living into adulthood due to advances in medical and surgical care and may have the opportunity to reproduce (3).

1.3 Causes

Congenital disorders vary widely in causation. A congenital disorder may be the result of a non-appropriate intrauterine environment, medication or drug exposures or maternal infections during pregnancy, or the result of a genetic abnormality.

It is hypothesized that in a significant percentage of congenital malformations, especially in cases where the malformation is isolated and none dimorphic features are identified, the etiology might be multifactorial with variable contribution of genetic and environmental factors that alter developmental pathways.

1.3.1 Teratogenic agents: medication and drug exposures

Some medications and drugs may cause congenital malformations by disturbing the development of the embryo or fetus.

Alcohol intake during pregnancy has been strongly related to a well-established disease called Fetal Alcohol Spectrum Disorder. The main features of this disorder include affectation of the central nervous system -causing intellectual disability and behavioral problems-, distinctive facial features and growth retardation. Some malformations, like congenital heart defects (10) or genitourinary malformations also have an increased incidence in children with Fetal Alcohol Spectrum Disorder (11).

Other drug exposures, such as cocaine, may also be related to congenital anomalies. Cocaine intake during gestation seems to be involved in several complications during pregnancy. The mentioned drug might act having disruptive effects mediated by a vasoconstriction which, depending on the stage of the pregnancy, might cause spontaneous miscarriages, prematurity, intrauterine growth restriction or congenital malformations (12, 13). Several birth defects have been related to maternal cocaine intake, including urogenital anomalies (14), distal limb deformities (12), intestinal atresia (12), cardiac defects (15, 16) and central nervous system malformations (17).

Not only drugs but also medications have been linked to birth defects. Nearly 50% of pregnant women are exposed to at least one medication during gestation (18), while 52% of 200 individuals referred for genetic counseling had been exposed to more than one potential teratogen (19).

Some antiepileptic medications, such as valproic acid and carbamazepine, have been implicated in congenital malformations, mainly in neural tube defects. The use of valproic acid monotherapy in the first trimester has been associated with significantly increased risk for other congenital malformations such as atrial septal defect, cleft palate, hypospadias, polydactyly and craniosynostosis (20). Systemic retinoids, used for several diseases such as skin conditions or some types of cancer, are related to an increased risk of spontaneous miscarriage, dysmorphic features, congenital heart defects, limb defects, hydrocephaly and microcephaly.

1.3.2 Maternal diseases

Some infections and other diseases suffered by the mother during pregnancy might cause birth defects. Diabetes (21, 22), hyperthyroidism (23-25), hypothyroidism (24, 25), phenylketonuria (26, 27) and epilepsy (28, 29) are some of the diseases that may affect pregnancy outcome. Several maternal infections such as rubella, cytomegalovirus, varicella-zoster, toxoplasmosis, syphilis or parvovirus B19 may contribute to alterations of fetal development (30-33). The severity of birth defects depends on the developmental stage when the infection occurs; usually the outcome is worse if the infection takes place at earlier developmental stage. Infection for cytomegalovirus is one of the most frequent infections during pregnancy that may cause deafness, mental retardation or microcephaly to the fetus.

1.3.3 Genetic causes

Genetic factors are clear contributors to congenital malformations. Several genetic and genomic alterations have been reported in patients with multiple congenital anomalies, with or without developmental delay or growth abnormalities. Deletions and duplications, corresponding to recurrent genomic disorders and other genomic regions, can be detected in 15-24% (34, 35) of patients with multiple congenital anomalies and/or intellectual disability. Point mutations in genes such as *TBX1* (DiGeorge syndrome) (36) or *ASXL1* (Bohring-Opitz syndrome) (37) have also been described in patients with multiple congenital anomalies and intellectual disability (MCA/ID).

The role of genetic factors in isolated congenital malformations is not as clear as in MCA/ID. Epidemiological studies have shown that siblings and offspring of individuals with congenital malformations have an increased risk of this type of anomalies (38-40) pointing out the high heritability of this type of diseases. Nevertheless, when considering isolated congenital malformations it is hypothesized that, in an important proportion of cases, genetic factors are relevant included within a multifactorial model (41-43). As an example, for some malformations like left heart hypoplasia an oligogenic model has been proposed (44). In other cases, genetic alterations related to the disease are identified in patients but inherited from a healthy parent, meaning that other factors –genetic or environmental ones- are also related to the disease.

1.3.3.1 Genetic rearrangements

Birth defects are often caused by chromosomal imbalances, especially when associated with dysmorphic features or developmental delay. Nevertheless, there are many syndromes and diseases related to developmental abnormalities as well as major malformations without a known molecular cause. It is likely that some of these syndromes are due to genomic alterations involving small chromosomal regions not identified so far. Several techniques, such as standard karyotyping, MLPA (Multiplex ligation-dependent probe amplification), aCGH (array Comparative Genomic Hybridization) or SNP (Single Nucleotide Polymorphism) array are used in order to detect genetic rearrangements.

Standard karyotyping might detect numerical and structural anomalies larger than 5–10 Mb and other techniques, such as FISH (45) or MLPA (46, 47), allow the identification of submicroscopic chromosomal imbalances. In the last decade, the development of molecular karyotyping by CMA (Chromosomal microarray analysis, including aCGH as well as SNP arrays) has led to a detection rate of as much as 15–24% of causative segmental aneusomies in patients with multiple congenital anomalies and/or mental retardation (MCA/MR) (34, 35). Retrospective studies in fetuses with multiple malformations have obtained a detection rate of causative chromosomal imbalances from 8 to 15% by using CMA (48-50). Several recurrent genomic disorders are frequently associated with congenital heart defect (CHD). Specifically, microdeletions at 22q11 (51, 52) and microduplications at 1q21.1 (53, 54) are a common cause of conotruncal heart defects .

Concerning isolated congenital malformations, few data is available. In those cases, the detection of genomic rearrangements could represent the identification of the genetic basis of the disease, perhaps as part of a more complex syndrome without other recognizable manifestations at this stage of development, or the finding of genetic susceptibility factors contributing to mutational load in a multifactorial model. Therefore, submicroscopic deletions and duplications may play a significant role in the etiology of this condition, either as direct cause or as possible genetic risk factor for isolated congenital anomaly (55).

1.3.3.2 Methylation alterations

Epigenetic mechanisms seem to contribute to many physiological processes as development or aging. Therefore, its alterations could have an important role in several diseases. Epigenetics includes those changes in the regulation of gene expression not caused by modification of the nucleotide sequence. DNA methylation and histone modification are both epigenetic processes. DNA methylation refers to the addition of a methyl group in 5' carbon of citosines, which alters the secondary interactions of the DNA molecule modifying the gene expression pattern. Aberrant DNA methylation may result in changes in transcription and subsequently in gene expression. As an example, hypermethylation of CpG islands located in promoter regions has been described as a mechanism to prevent the transcription of some genes leading to an abnormal gene expression profile related to tumor development, for example.

Genetic diseases with a well-established cause such as Prader-Willi or Angelman syndrome are related to a misregulation of epigenetic mechanisms (56). Regarding methylation, some studies have shown its relevance in diseases with a genetic component such as schizophrenia or bipolar disorders. Furthermore, some data is already available regarding the role of methylation in congenital anomalies, including correlation of genetic variants in genes regulating methylation, such as the folate pathway, and the risk for heart malformations in Down syndrome patients (57). Folate-pathway has also been related in some articles with isolated heart defects without Down syndrome, as in a meta-analysis with polymorphisms in the *MTHFR* gene (58).

1.3.3.3 Point mutations

Point mutations have been described as the basis of many single gene disorders usually with Mendelian inheritance. Next generation sequencing

technologies developed in the last few years have increased the number of known genetic diseases caused by point mutations.

Point mutations in some genes have also been reported as the cause of congenital malformations. In MCA/MR mutations in several genes, such as *DHCR7* (59), *FOXF1* (60) and *NIPBL* (61) have been reported. In isolated congenital heart malformations mutations in genes with a crucial role in heart development have been detected, such as mutations in *GATA4* (62) or *NKX2-5* (63). Also in other types of malformations, as holoprosencephaly or microcephaly, point mutations are strongly related to the phenotype.

Otherwise, in neural tube defects point mutations detected in genes related to planar cell polarity have been described in affected patients, although an apparently healthy progenitor was also a carrier of the same variant. In those cases, point mutations might be a mutation with incomplete penetrance or a susceptibility factor that requires other genetic or environmental factors to cause the malformation.

1.4 The relevance to study fetal tissues

The availability of fetal samples to study congenital malformations may provide additional information compared to other types of studies. Usually, terminations of pregnancy due to congenital malformations are related with severe anomalies, with a higher chance to have a genetic background. In addition, the possibility to perform the experiments in the tissue affected, DNA as well as RNA analysis, would provide more accurate information. Thus, it would be possible to obtain additional data, such as the detection of somatic mosaicism or the expression level of candidate genes in the tissue affected. Apparently, the information regarding expression profiles should be more interesting if the affected tissue is considered. Otherwise, the study of expression levels in blood if the malformation is located in the heart could not be the most appropriate strategy in order to obtain information about the consequences of a genetic alteration.

2 MODEL DISEASES

2.1 Multiple malformations

Congenital malformations often occur in the setting of multiple congenital anomalies, including dysmorphic facial features, or in association with different organ malformations, developmental defects, or growth abnormalities (3, 4). In these cases in which there is a more complex syndrome, chromosomal aberrations are a frequent cause of disease, although point mutations in specific genes have also been described. As an example, Wolf-Hirschhorn (64) syndrome is a polymalformation syndrome caused by a recurrent microdeletion in 4p16.3. The most common clinical features in Wolf-Hirschhorn syndrome are growth retardation, microcephaly, skeletal anomalies, hypotonia, severe developmental delay and congenital heart defects. Other polymalformation syndromes are DiGeorge (65) syndrome or 1p36 deletions. In other diseases polymalformation syndromes are caused by a point mutation as Kabuki syndrome (66) or Rubinstein-Taybi (67).

Nevertheless quite often when multiple congenital anomalies are identified is not possible to recognize a known syndrome. In those cases, screening methodologies such as MLPA for subtelomeric rearrangements or recurrent genomic disorders or CMA are useful to diagnose the cause. In those cases, CMA leads to a detection rate of 15-24% of causative segmental aneusomies in patients with MCA/MR (34, 35).

2.2 Congenital Heart Defects

In particular, CHD represent a high percentage of clinically significant birth defects. The incidence of CHD is approximately 8 per 1000 live births, making CHD the most common malformation (68, 69). In a definition proposed by Mitchell et al, congenital heart defect is "a gross structural abnormality of the heart or intrathoracic great vessels that is actually or potentially of functional significance." Diseases such as hypertrophic or dilated cardiomyopathy, long QT and Wolff-Parkinson-White syndromes are usually not included in CHD because, even though the genetic alteration which causes the disorder is present at birth, the cardiomyopathy presents later in childhood or adolescence. In table 2 and 3 the incidence of most common congenital heart defects is described (70, 38).

Lesion	Number of Studies	Mean	SD	Lower Quartile	Median	Upper Quartile	NERICP 1975–1977
VSD	43	3,570	2,878	1,757	2,829	4,482	345
PDA	40	799	1,399	324	567	782	135
ASD	43	941	1,043	372	564	1,059	65
AVSD	40	348	165	242	340	396	110
PS	39	729	731	355	532	836	73
AS	37	401	543	161	256	388	41
Coarc	39	409	246	289	356	492	165
Tetralogy	41	421	188	291	356	577	196
d-TGA	41	315	115	231	303	388	218
HRH	32	222	199	105	160	224	—
Tricuspid atresia	11	79	52	24	92	118	56
Ebstein's anomaly	5	114	138	38	40	161	12
Pul Atresia	11	132	123	76	83	147	69
HLH	36	266	216	154	226	279	163
Truncus	30	107	71	61	94	136	30
DORV	16	157	103	82	127	245	32
SV	23	106	70	54	85	136	54
TAPVC	25	94	46	60	91	120	58
All cyanotic	37	1,391	590	1,078	1,270	1,533	888
All CHD*	43	9,596	7,484	6,020	7,669	10,567	2,033
BAV	10	13,556	13,049	5,336	9,244	13,817	—

*Excluding bicuspid nonstenotic aortic valves, isolated partial anomalous pulmonary venous connection and silent ductus arteriosus.

BAV = bicuspid aortic valve; CHD = congenital heart disease; Coarc = coarctation of the aorta; NERICP = New England Regional Infant Cardiac Program. Other abbreviations as in legend to Figure 5.

Table 2. Incidence per million live births of the most common congenital heart malformations reported in England between.

Index Case	First Recurrence (n)	Second Recurrence (n)	Third Recurrence (n)	Total Recurrence Cases (n)	Group Concordance	Exact Concordance	Parent/Child Pairs:Sibling Pairs
HLH (n = 12)	HLH: 4 CoA: 2 VSD: 5 TAPVD: 1	HLH: 1 CoA: 2	HLH: 1	16	Left heart defects: 6/12	HLH: 4/12	0:12
CoA (n = 15)	CoA: 2 HLH: 3 VSD: 6 AVSD: 1 TA: 1 Complex: 1 TGA: 1			15	Left heart defects: 5/15	CoA: 2/15	2:13
AS (n = 8)	AS: 3 HLH: 2 PS: 1 TA: 1 Complex: 1			8	Left heart defects: 5/8	AS: 3/8	3:5
Total (n = 35)	35	3	1	39	16/35 (46%)	9/35 (26%)	5:28

*Concordance is calculated using first recurrence data.

AS = aortic stenosis; AVSD = atrioventricular septal defect; CoA = coarctation of the aorta; Complex = complex heart defect; HLH = hypoplastic left heart; PS = pulmonary stenosis; TA = tricuspid atresia; TAPVD = total anomalous pulmonary venous drainage; TGA = transposition of the great arteries; VSD = ventricular septal defect.

Table 3. Recurrence risk of heart malformations and type of CHD in relatives of patients with left heart defects.

About half of congenital cardiovascular malformations are severe and require one or more surgical procedures in the neonatal period or during childhood (71). Because of its severity, congenital heart disease is a leading cause of morbidity and mortality during infancy; about 8% of all infant deaths is due to congenital heart defects (72).

There are several classification systems for CHD. In 2000 the International Congenital Heart Surgery Nomenclature was developed to provide a generic classification system.

2.2.1 Hypoplasia of the heart

Hypoplasia of the heart refers to an underdevelopment of one side of the heart and is the most severe form of CHD. It may compromise the right ventricle or the left one resulting in an ineffective blood pumping. If the hypoplasia affects the right ventricle, the blood circulation to the lungs is not functioning properly; hypoplasia of the left ventricle prevents the correct blood pumping to the body. Hypoplasia of the heart has an approximate incidence of 3/10.000 pregnancies. Regarding left heart hypoplasia, often associated with obstruction to left ventricle outflow when is severe, surgery is required for long-term survival due to the inability of the left ventricle to support the systemic circulation. There are two possible surgical treatments: neonatal cardiac transplantation or a sequence of complex open-heart operations in infancy in order to achieve a univentricular circulation in which the right ventricle supports the systemic circulation and pulmonary blood flow is passive. When hypoplasia is mild, the outflow obstruction repair may prevent further complications since the left ventricle may be capable of supporting the systemic circulation (73).

It is though that heart hypoplasia has an important genetic background, taking into consideration the proven high heritability. Nevertheless, the mechanisms causing this type of malformations are poorly known.

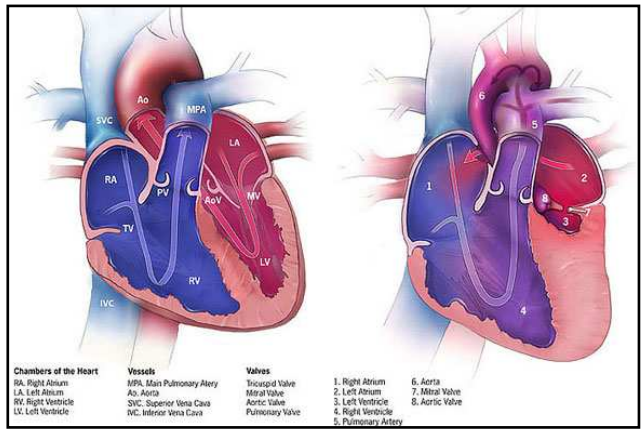


Figure 1. Picture of a normal developed heart (left) and a heart with left heart hypoplasia (right). Note the underdevelopment of the left part of the heart.

2.2.2 Obstruction defects

Obstruction defects refer to malformations of heart valves, arteries or veins, which difficult blood circulation because of its narrowing or blocking. When it occurs, the heart has to pump harder to achieve a proper circulation, causing heart hypertrophy.

Common obstructive defects include pulmonic stenosis, aortic stenosis, and coarctation of the aorta. Bicuspid aortic valve stenosis and subaortic stenosis are also obstructive defects with a lower incidence. Some studies have shown a high heritability of some of these conditions, hypothesizing an important genetic component in the etiology of this type of heart defects (74).

2.2.3 Septal defects

The septum is the partition wall that separates the left heart from the right heart. A defect in the septum allows blood to circulate from the left side of the heart to the right, reducing heart efficiency. Depending on where the defect is placed, it is called atrial septal defect or ventricular septal defect. Ventricular septal defects (VSD) are the most common type of CHD, affecting approximately 0.44-0.48 per 1.000 elementary school children (75). VSD is an abnormal opening in the interventricular septum which allows communication between the right and left ventricles. VSD might be part of a more complex heart malformation or an isolated malformation. Severity of VSD may depend on the size, location and conjunction with other heart malformations.

Point mutations in *GATA4*, *GATA6* and *JAG1* have been identified in patients with VSD, as well as deletions in 8p23.11, which have been related to VSD and auricular septal defects (ASD).

2.2.4 Cyanotic defects

Cyanotic defects include those heart malformations causing cyanosis because of a lack of oxygen in the body. The most frequent cyanotic defect is Tetralogy of Fallot, a complex heart malformation which consists of pulmonary infundibular stenosis, overriding aorta, VSD and right ventricular hypertrophy. This conjunction of heart malformations implies a low oxygenation of blood because of the mixing of oxygenated and deoxygenated blood. Tetralogy of Fallot has a high incidence in patients with 22q11.2 deletion syndrome (22%). In addition, point mutations in some genes such as *FOG2* and *Nkx2-5* have been identified related to the same malformation (76).

2.3 Central Nervous System malformations

There are several types of central nervous system malformations. Some of the most frequent ones are described below.

.2.3.1 Holoprosencephaly

Holoprosencephaly is a complex brain malformation in which the prosencephalon fails to develop into two hemispheres, affecting the forebrain and the face. Prevalence is estimated to be 1/10,000 live and stillbirths and 1/250 of pregnancies. Mechanisms underlying holoprosencephaly are partially understood and it seems that *HOX* genes, which may guide the placement of embryonic structures, do not function properly along the midline of the head. Activation of *HOX* genes prevents the merge of the left and right side of prosencephalon by the division of that embryonic structure (77).

Depending on the severity, three forms have been described: lobar, semi-lobar and alobar. Another form, the less severe one, has also been described and it is called microform and is characterized by midline defects without the typical holoprosencephaly brain malformation. Often it is not easy to classify this type of malformations into those categories taking into account that holoprosencephaly is a continuous spectrum of abnormal separation of the hemispheres (77).

The cases with cyclopia, proboscis, premaxillary agenesis or median or bilateral cleft lip/palate with also a severe brain malformation leading to miscarriages or stillbirths in some cases are at the most severe end of the spectrum. In less severe cases, patients might present a normal or near-normal brain development and facial defects such as hypotelorism, solitary maxillary median incisor or even normal face. In affected patients, a wide range of associated clinical manifestations is reported: developmental delay, hydrocephalus, motor impairment, feeding difficulties, oromotor dysfunction, epilepsy or hypothalamic dysfunction. Prognosis depends on the severity and the associated complications.

The etiology of holoprosencephaly is partially known and it is very heterogeneous. Chromosomal abnormalities such as trisomy 13 have been described as being related to the malformation. Other known syndromes, for example, Smith-Lemli-Opitz syndrome or Charge syndrome have also been related to holoprosencephaly. In non-syndromic holoprosencephaly, several

genes have been implicated: *SHH*, *ZIC2, SIX3, GIF, PTCH1, GLI2, FOXH1, TDGF1 , DISP1 , NODAL, FGF8, GAS1, DLL1, and CDON* (78, 79). Environmental factors such as maternal diabetes or hypocholesterolemia during gestation might be also related to that brain malformation.

2.3.2 Ventriculomegaly

Ventriculomegaly is one of the most common abnormal brain findings on prenatal ultrasound, occurring in around 1–2 per 1000 pregnancies. Ventriculomegaly might be an isolated malformation or part of a more complex syndrome. Ventriculomegaly is a brain malformation which occurs when the lateral ventricles are abnormally dilated or enlarged. Ventriculomegaly is diagnosed when the width of the atrium of the lateral ventricle measures more than 10 mm. Depending on that measurement, ventriculomegaly is described as mild to moderate (between 10 and 15 mm) or severe (greater than 15mm). In many cases of mild ventriculomegaly, there is resolution during the pregnancy. Ventriculomegaly may affect also the third and fourth ventricle. Enlargement of the ventricles may occur for a number of reasons, such as loss of brain volume or impaired outflow or absorption of cerebrospinal fluid from the ventricles. The underlying mechanisms for the decreased of brain volume or the accumulation of cerebrospinal fluid may include genetic alterations, maternal infections or tumors located in the fetal brain. Some genetic syndromes, like Walker-Walburg syndrome, include ventriculomegaly as one of their main features. Mutations in *L1CAM* have been reported in families with X-linked aqueductal stenosis causing ventriculomegaly. Often, however, there is no identifiable cause (80).

2.3.3 Neural tube defects

Neural tube defects (NTDs) are one of the most common birth defects, occurring in approximately one in 1,000 live births. The neural tube is a narrow channel that folds and closes during the third and fourth weeks of pregnancy to form the brain and spinal cord. An NTD is an opening in the spinal cord or brain that occurs very early in human development as a result of the failure in neural tube closure.

There are several types of neural tube defects depending on the placement and the presence or lack of cover. Open neural tube defects, which are more common, occur when the brain and/or spinal cord are exposed at birth through a defect in the skull or vertebrae without being covered; examples of open NTDs are anencephaly, encephaloceles, hydranencephaly and spina

bifida. Closed NTDs occur when the spinal defect is covered by skin; common examples of closed NTDs are lipomeningocele and tethered cord (81-84).

2.3.3.1 Anencephaly

Anencephaly is a neural tube defect that occurs when the head end of the neural tube is not closed properly resulting in an absence of a major portion of the brain and skull, including the cranial vault and the covering skin. As a consequence of this malformation, the brain mass is reduced and infants with this condition are born without the main part of the forebrain. Infants are either stillborn or usually die within a few hours or days after birth. Its prevalence at birth ranges from 1 in 5000 to 1 in 2000 (81).

2.3.3.2 Encephalocele

Sac-like protrusions of the brain and the membranes that cover it through opening in the skull are known as encephalocele. Encephalocele might be located in different parts of the skull as in the middle of it, between the forehead and nose or on the back side of the skull. The severity of encephalocele varies depending on its location and depending on the quantity of brain tissue affected (82).

2.3.3.3 Spina bifida

Spina bifida is another type of neural tube defect in which the opening of the tube is located along the spinal cord and therefore one or more vertebral arches may be incomplete. The most common locations of those malformations are the lumbar and sacral areas. Spina bifida malformations fall into three categories: spina bifida occulta, spina bifida cystica with meningocele, and spina bifida cystica with myelomeningocele (83).

Spina bifida occulta is the mildest form of spina bifida consisting in an incomplete close of, the outer part of some vertebrae. The opening is small and the spinal cord does not protrude through it. Many people with this type of spina bifida do not even know they have it, as the condition is asymptomatic in most cases. The incidence of spina bifida occulta is approximately 12% of the population (84).

Meningocele is the protrusion of the meninges of the spinal cord through a defect in the spinal column forming a cyst filled with cerebrospinal fluid that does not contain neural tissue. As the nervous system remains undamaged, individuals with meningocele are unlikely to suffer long-term health problems. They are treated surgically.

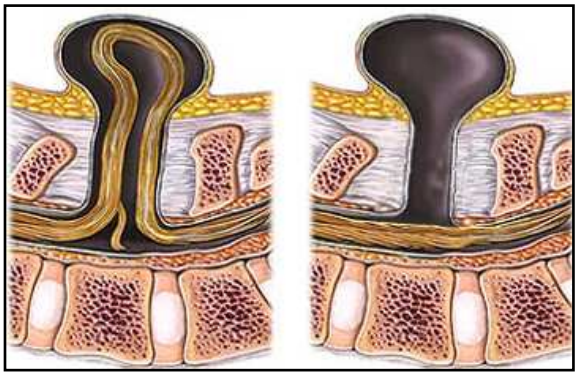


Figure 2. Picture of NTD: myelomeningocele and meningocele. Note the opening of the tube and the protrusion of the sack in both pictures. On the left side, the spinal cord protrudes as well as the meninges, so the neural tissue might be damaged; this image corresponds to a myelomeningocele. On the right side, a meningocele is shown: the meninges protrude trough the opening but not neural tissue is included within the

Myelomeningocele is the most significant form of spina bifida and it leads to disability in most affected individuals. In individuals with myelomeningocele, the unfused portion of the spinal column allows the spinal cord to protrude through an opening. The meningeal membranes that cover the spinal cord form a sac enclosing neural tissue. The neural tissue enclosed in the protrusion and the nerves that originate at that level of the cord are damaged or not properly developed. As a result, there is usually some degree of paralysis and loss of sensation below the affected area of the spine. Therefore, the higher the level of the myelomeningocele, the more severe the nerve dysfunction may be. Affected individuals may have loss of sensation, deformities of the hips, knees or feet, difficulty walking or inability to walk, loss of muscle tone, bladder and bowel incontinence, urinary tract infections and impaired renal function.

Many individuals with spina bifida will have an associated abnormality of the cerebellum, called the Arnold Chiari II malformation. In affected individuals, the cerebellar tonsils are displaced from the back of the skull down thorough

foramen magnum. In an important proportion of patients with Arnold-Chiari malformation, hydrocephalus will also occur because the displaced cerebellum interferes with the normal flow causing an accumulation of cerebrospinal fluid. Other central nervous system malformations such as corpus callosum abnormalities or cortex anomalies are also seen in patients with spina bifida.

The elected treatment for neural tube defect varies depending on the type of the malformation. For instance, some mild versions of spina bifida require minimal treatment. Otherwise, more severe forms require surgery to correct the problems and, despite the operation, the normal function to the affected part of the spinal cord might not be restored. Intrauterine surgery for spina bifida has also been performed and the safety and efficacy of this procedure is currently being investigated (85).

2.3.3.4 Prevention of neural tube defects

Regarding the cause of this malformation, it is hypothesized that neural tube defects present a multifactorial inheritance, the result of gene-environment interactions. Mutations in some genes related to planar cell polarity have been described (86-90) although in the majority of cases mutations were inherited from a healthy parent. Those results arise doubts about the role of these mutations in the etiology of the disease, probably being one of the contributing factors, but not the only one, to neural tube defects. Nonetheless, familial cases with a seemingly autosomal recessive mode of inheritance have been reported.

Since several years ago, the relation between folic acid intake before and during pregnancy and neural tube defect has been proven. The incidence of spina bifida may decrease by up to 70% when daily folic acid supplements are taken prior to conception and during the pregnancy. However, the mechanism underlying the role of folic acid in neural tube defects is poorly understood (91, 92).

2.4 Renal malformations

Congenital anomalies of the kidney and urinary tract occur in 1 in 500 births and are a major cause of morbidity and mortality in children. Some renal anomalies are part of a more complex syndrome, although most cases are isolated urinary tract malformations. Several genes have been reported in association to isolated renal malformations, such as *Ret, PAX2, SALL1 and BMP4* (93-97). Renal malformations might also be caused by mother's exposure to certain drugs or medication during pregnancy. Medications

which may cause renal malformations include blood pressure medicines such as angiotensin-converting enzyme inhibitors and angiotensin receptor blockers. Illegal drugs, such as cocaine, may also cause renal malformations in the fetus. Nevertheless, the majority of cases remain unsolved. Two of the most common renal malformations are renal hypoplasia/agenesis and renal dysplasia.

2.4.1 Renal agenesis

Renal agenesis or hypoplasia may have different levels of severity regarding the affectation of a single or both kidneys and also considering the grade of hypoplasia, being agenesis the most severe end of the spectrum. Bilateral renal agenesis is the failure of both fetus kidneys to develop during gestation. The absence of kidneys is a lethal condition that causes oligohydramnios, a deficiency of amniotic fluid in pregnancy that may lead to fetal malformations known as the Potter sequence. The etiology of bilateral renal agenesis is unclear, although some autosomal recessive cases have been described related to Fraser syndrome. It is also known that bilateral renal agenesis is more common when one of the progenitors has some type of kidney malformation, especially unilateral renal agenesis. Those results indicate a common genetic background related to unilateral and bilateral renal agenesis (98, 99).

2.4.2 Renal dysplasia

In kidney dysplasia, the internal structure of one or both kidneys is not developed properly and fluid-filled sacs called cysts replace normal kidney tissue. If kidney dysplasia only affects one kidney the prognosis is good, only requiring regular checkups to ensure the correct functioning of the remaining kidney. When kidney dysplasia affects both kidneys is generally a lethal condition, and those who do survive require dialysis and kidney transplant very early after birth. In some cases, a silent renal dysplasia in one progenitor is discovered after the birth of an affected child. When kidney dysplasia is discovered in an infant, an ultrasound examination may reveal the condition in one of the parents (99).

2.5 Down syndrome

Trisomy 21 or Down syndrome (DS) is a polymalformation syndrome characterized by the presence of a third copy of chromosome 21. In 95% of the cases, trisomy 21 is free, which means that the extra chromosome is not integrated to another chromosome and it is present in all analyzed cells. In those cases, the trisomy is due to a non-disjunction during meiosis. In 2-3%

of the cases the trisomy 21 is free and mosaic, namely, not found in all studied cells but only in a proportion of them. In the remaining 2-3% of the cases, the supernumerary chromosome 21 is integrated to another chromosome as part of a translocation. The diagnosis can be made by karyotyping, although other techniques may also detect the trisomy. The recurrence risk for the parents of an affected child is only slightly higher if the trisomy is free. In cases of DS caused by translocation, the risk is increased only if one of the parents has a balanced rearrangement. For a person with Down syndrome, the risk of transmitting the disease to the descendants is 1/3 (100).

The most severe features related to DS are congenital heart defects -mainly atrio-ventricular canal, digestive malformations as duodenal atresia, congenital cataract, small size, Hirschsprung disease, seizures, leukaemia, sleep apnea, sensory deficiencies, auto-immune and endocrine pathologies, earlier aging, Alzheimer disease and mental retardation. Patients with DS also present a recognizable dimorphic pattern with upslanting palpebral fissures, epicanthus, flat neck, round face, small nose or bilateral single palmar crease. Median life expectancy is now above the age of 50 years (101).

Although the cause of DS is well established, the factors influencing the appearance of different types of malformations are poorly understood. For example, it is well known that patients with trisomy 21 have a higher incidence of congenital heart malformations than other children, even though little is known about the factors related to that. Concretely, the incidence of CHD in these patients is between 43-58% (102, 103).

Considering that all patients with trisomy 21 carry the same genetic alteration, it is clear that other factors, genetic or epigenetic, are involved in the development of the CHD, as well as in the development of other malformations reported in DS.

3 TECHNIQUES TO DETECT THE CAUSE OF THE MALFORMATION

3.1 Techniques to study genetic rearrengements

There are multiple techniques that can be applied to the detection of possible etiology of congenital malformation, either driven by the clinical suspicion or as a screening approach. This review does not pretend to be exhaustive but only to introduce the genetic and genomic technologies that have been applied in the course of this project.

3.1.1 MLPA

The MLPA (Multiplex Ligation-dependent Probe Amplification) is a method for the detection of multiple chromosomal imbalances in a single reaction. This technique allows semiquantitative screening of specific genomic regions, based on the simultaneous hybridization and subsequent amplification of up to 40 different probes in a single reaction. It is a procedure easy to implement and requires a minimum of 50 ng of DNA. In the MLPA technique, the genomic DNA is hybridized in solution with a set of probes, each one divided into two fragments. One (L) consists of a sequence of approximately 30 nucleotides specific to a particular target in the DNA. This sequence is flanked by nucleotides that are complementary to a universal primer. The other fragment (R) also contains both the specific sequence for the chosen target DNA region and the hybridization sequence for the universal primer. Probe L, between the primer binding region and the target binding portion, includes a region variable in length called stuffer in order to generate size differences between probes included

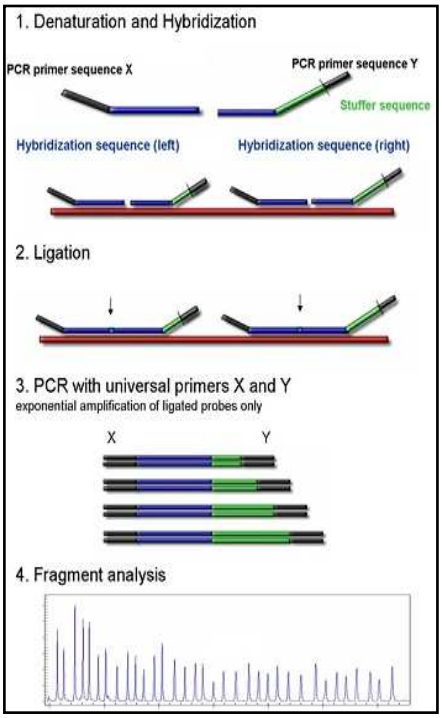


Figure 3. Overview of MLPA steps. In this figure the main steps of MLPA are shown (denaturation, hybridization, ligation, PCR and fragment analysis) as well as the structure of MPA probes.

in the same reaction. This strategy allows electrophoretic resolution of the amplified fragments (figure 3).

Both fragments of the MLPA probe complementary to the DNA sequence are designed in order to hybridize with the target sequence one adjacent to the other; then both are joined by ligase enzyme. The joining of both fragments generates a single probe for each target region that can be amplified by conventional PCR using universal primers; the fragments not attached to another one cannot be amplified. The amount of bound probes is proportional to the number of copies of the target sequence. After PCR amplification the relative height of the peaks may indicate deletion or duplication of the target sequence, namely the gene dose (104).

The MLPA technique is inexpensive, simple, fast and provides flexibility for the study of regions of interest. Moreover, MLPA allows the study of a large number of samples in the same experiment. Despite the advantages, the main limitation of this technique is the inability to detect balanced rearrangements, eg, translocations or inversions. MLPA might be useful to screen easily several samples and loci of interest (aneuploidy screening in prenatal diagnosis, regions associated to microdeletion and microduplication syndromes in patients with mental retardation) or to validate results of other techniques, such as the results obtained with CMA technologies.

3.1.2 ACGH

The most recently developed technique for the detection of duplications and deletions is based on comparative genomic hybridization, so called aCGH. This technique is based on a competitive hybridization between two DNAs, the sample of interest and a control sample. Each of these samples is labeled with a different fluorochrome and it is hybridized on a support containing an array of probes. Different types of probes might be used in aCGH such as BACs (Bacterial Artificial Chromosome) or oligonucleotides. The resolution of this technique depends on the size of the probes, the number of them and their distribution through the genome (Figure 4).

The competitive hybridization between the DNA sample and the control DNA allows the detection of deletions and duplications: an intense signal of the patient fluorochrome would indicate a duplication of the region and a decrease in the signal given by patient fluorochrome may correspond to a deleted region (105-107).

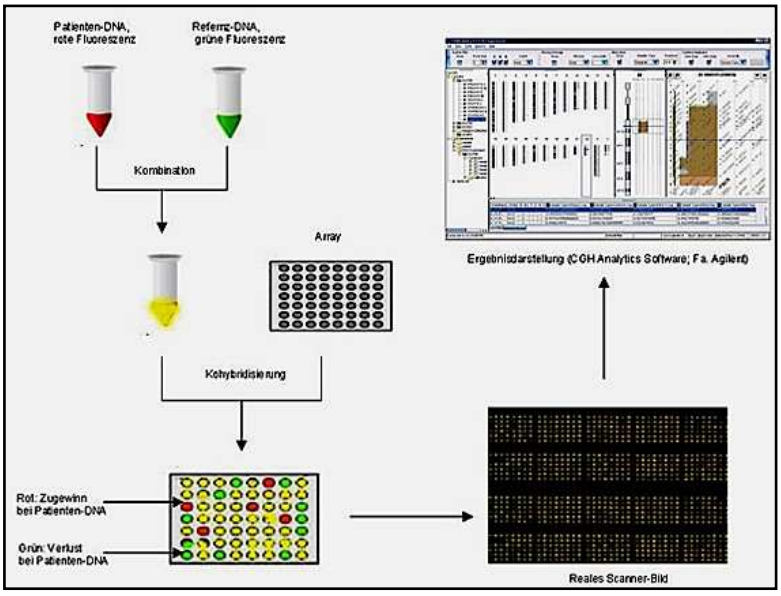


Figure 4. Schematic overview of aCGH. In the first step, genomic DNA samples from a test and a reference are differentially labeled. Then, both DNA samples are mixed in equal amounts and hybridized to the array containing the probes. After that, the array is scanned and computer image assesses the relative fluorescence for each probe in the array. In the spots where the same fluorescence is detected for both fluorochromes no alterations is present in the patients; when more fluorescence is detected for patients' fluorochrom a duplication is considered in patients samples; when the ratio is higher for reference fluorochrom, a deletion is present in the patient.

The aCGH permits the study of copy number variations in a large-scale manner by the screening of regions throughout the genome, as well as a very detailed analysis of some selected regions, depending on the design of the array probes. The aCGH has the same limitation than the MLPA, as it does not detect balanced alterations.

aCGH is useful for the screening of chromosomal imbalances for which there is no suspicion, as it may have coverage throughout the genome with high resolution. Using aCGH it is also possible to detect rearrangements in mosaicism. Recent published articles describe a detection of mosaicism as low as a 10% level.

3.1.3 SNP arrays

SNP arrays do not rely on the same system than comparative genomic hybridization described in the previous section, but the same term is used

when it is applied to study the copy number variants along the genome. SNP arrays allow the study of the alleles present in each individual in a specific number of SNPs, depending on the design of the experiment. This technique quantifies the intensity of each of the alleles of the SNPs analyzed. The information provided allows, in addition to deletions and duplications, the analysis of heterozygosity and uniparental disomy. There are two relevant parameters for the analysis of SNP array: the intensity of the fluorescent signaling for each position and the BAF (B allele frequency), which represents the proportion of signaling between both alleles, namely, the B allele signal divided by the sum of the A and B signals.

The data for the detection of deletions and duplications is provided by the intensities of the fluorescent signaling for each position and also by the relative intensity between the two alleles in each nucleotide tested. For example, in deleted regions the intensity is reduced and no heterozygous positions are detected. In duplicated regions the intensity is higher and in heterozygous positions one of the alleles is detected in a doubled intensity compared to the other. Homozygous regions might be recognized because normal intensities are detected but no heterozygous positions are identified inside that fragment, namely, BAF is 0 (not B allele detected) or 1 (only B allele is detected). Recently, some softwares have been developed to analyze the presence of mosaic rearrangements by using data from SNP arrays. As in the MLPA and aCGH, the main limitation of SNP arrays is the inability to detect balanced rearrangements (108).

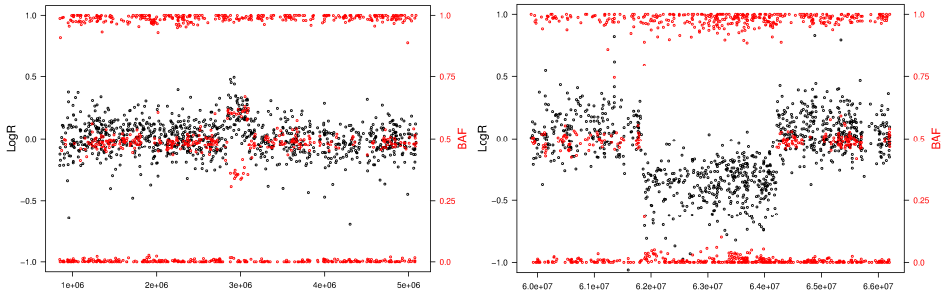


Figure 5. SNP array plots. A: Plot corresponding to a duplication. Within the duplicated region the intensity (LogR) is higher and BAF shows a proportion of the alleles of 2:1. B: Plot showing a deletion. LogR of that region is reduced and there are none heterozygous positions within the region, as there are no BAF values of 0.5 meaning a proportion of 1:1 of the alleles in this SNP.

3.2 Techniques to study methylation alterations

3.2.1 Methylation arrays

Methylation arrays have been developed recently in order to interrogate thousands of CpG sites per sample in the same experiment. Before the genome-wide methylation array, a bisulfite treatment of DNA is required. This procedure converts unmethylated citosine residues to uracile but does not modify 5-methylcitosine residues. Therefore, bisulfite treatment introduces modification in the DNA sequence depending on the methylation status of citosine nucleotides. The microarray contains two different bead-bound probes for each interrogated CpG. One of the bead-bound probe is complementary to the DNA sequence without the modification in citosine residue, while the other one is complementary to the DNA sequence with the conversion of the unmethylated citosine to uracile. The hybridization of DNA with bead-bound probe is followed by a single-base extension with a labeled nucleotide. The ratio of fluorescent intensity between the two bead-bound probes for each CpG provides information about the level of methylation in each CpG. Available softwares permit the comparison of methylation patterns between single samples and between groups of samples (109).

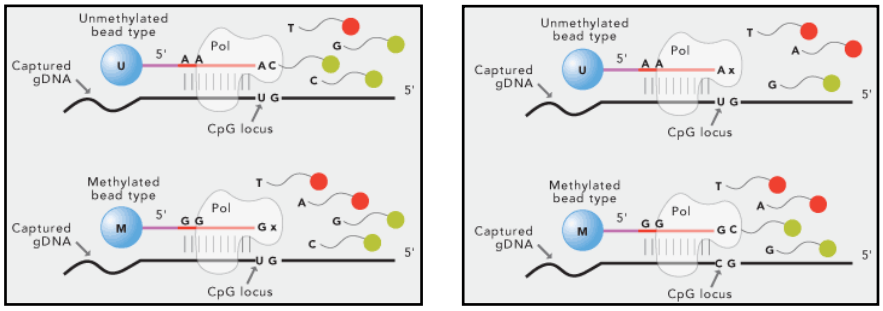


Figure 6. Principle of methylation array. A: as the CpG analyzed is unmethylated, the cytosine residue is converted to uracile in bisulfate treatment. Due to the modification, the gDNA sequence is complementary to the unmethyalted bead. Therefore, the extension only occurs in the unmethylated probe. B: the CpG tested is methylated. For that reason no modification occurs during bisulfate treatment and the gDNA sequence hybridizes with the methyalted probe and extension takes place in that bead.

3.2.2 Methylation specific MLPA

Methylation Specific MLPA (MS-MLPA) is a method to study methylation profiling based on the standard MLPA technique. The combination of the usual MLPA procedure with a methylation sensitive restriction enzyme provides information regarding methylation levels in target regions. As in standard MLPA, it is possible to study several regions in each reaction and by using MS-MLPA both information about CNV (copy number variation) and methylation levels is obtained.

MS-MLPA shares several steps with the standard MLPA procedure. First step, includes the denaturation of DNA, the hybridization with MLPA probes which are complementary to the regions of interest and the ligation of the two fragments in which each MLPA probe is divided. The MLPA probes must contain a restriction site for the methylation sensitive enzyme used in the procedure. After the ligation, an additional step is introduced in MS-MLPA compared to standard MLPA. At that point, the reaction should be split into two tubes in order to obtain information about CNV and methylation profiling. One of the tubes is processed as a standard MLPA reaction to obtain one undigested sample for copy number detection; the other tube is incubated with a methylation-sensitive restriction enzyme which cuts the fragments depending on the methylation status to obtain methylation data.

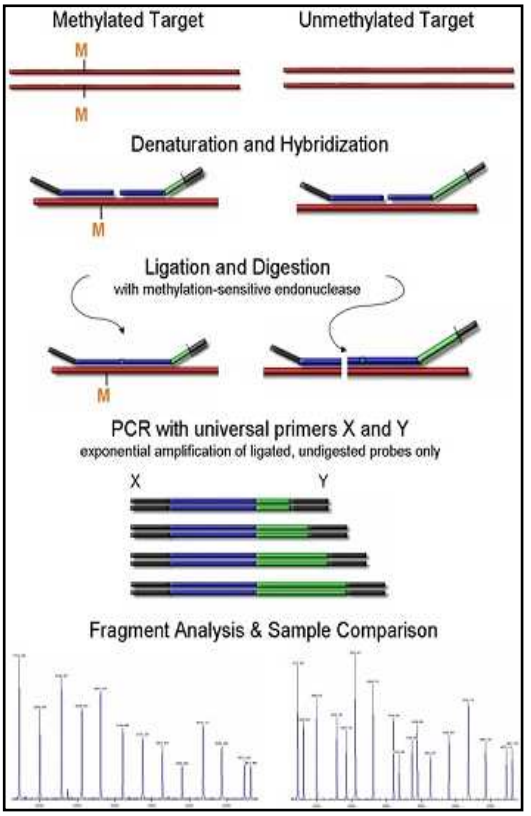


Figure 7. MS-MLPA overview. Steps in MS-MLPA are the same than in standard MLPA except for the digestion of probes by the methylation sensitive enzyme. This step allows the quantification of methylation level as the enzyme cuts depending on the methylation level.

After that, a PCR is done for both tubes separately; digested probes are not amplified during the PCR. Therefore, the comparison of produced signal during capillary electrophoresis between both tubes provides information about the proportion of digested probes and, consequently, the proportion of methylated/unmethylated DNA (110).

3.2.3 EpiTyper

EpiTyper is a methodology to study methylation ratios for individual CpGs within a target sequence. This procedure is based on mass spectrometry to quantify the different obtained fragments. Firstly, it is necessary to perform a bisulfite treatment of the DNA to convert unmethylated cytosine to uracil. After that, an RNA transcription is performed since in the next step an Rnase is used in order to cleave RNA specifically at cytosine and uracil nucleotides. The cleavage produces different patterns depending on the changes introduced to the DNA sequence during bisulfite treatment. Those fragments are quantitatively measured with MALDI-TOFF (Matrix-assisted Laser Desorption/Ionization-Time of Flight)technology (111).

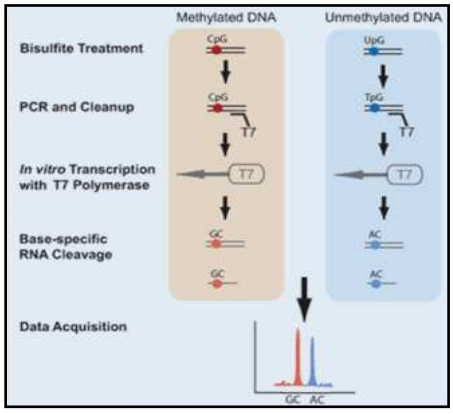


Figure 8. Scheme of the steps in EpiTYPER technology. On the left of the image, the procedure for a methylated CpG. As the CpG is methylated, no modification is done during the treatment with bisulfate. On the right side, as the CpG is not methylated, a modification is introduced in the sequence. Due to this difference, the pattern obtained after the base-specific RNA cleavage and, consequently the profile obtained in MALDI-TOF technology will be different.

3.3 Techniques to study point mutations

3.3.1 Sanger sequencing

Sanger sequencing was develop in 1975 by Sanger and Coulson in order to establish the nucleotide content for a target sequence. Before sequencing, a polymerase chain reaction is done with a pair of primers flanking the region of interest in order to obtain the target sequence amplified. Subsequently, cycles of template denaturation, primer annealing and primer extension are carried out with the four deoxynucleotide bases and a low concentration of

fluorescently labeled dideoxynucleotides, each dideoxynucleotide with a different color. The formers allow primer extension but the latters terminate the extension. At the end of the reaction different length fragments are obtained with a different labeled dideoxynucleotide at the end. Those fragments are size-separated by electrophoresis and the color is detected in order to compose the sequence. Nowadays, up to 1000 bases might be read using Sanger sequencing with a high accuracy. Nevertheless, this target sequence approach might be used to sequence concrete regions but it is limited by its throughput.

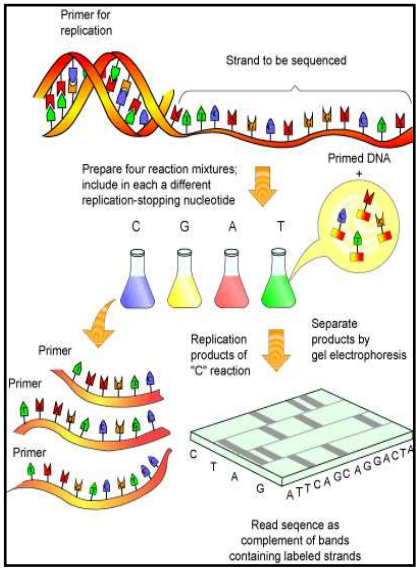


Figure 9. Schematic overview of Sanger sequencing procedure. The use of dideoxynucleotide might cause the end of the chain synthesis. As each dideoxynucleotide is labeled with a different colour, the separation of the products in an electrophoresis gel will allow the reading of the sequence.

3.3.2 Next generation sequencing

Regarding the relevant limitation of Sanger sequencing, a new technology called Next Generation Sequence (NGS) has been develop in order to sequence DNA using a large-scale approach. During the last five years, NGS or Massive parallel sequencing instruments have been develop and are capable to sequence millions of DNA fragments with a considerably lower cost. Several different technologies are now available. By using this technology, several genes related to Mendelian diseases have been identified (112, 113).

Considering the costs of the technology and the challenges in the interpretation of whole-genome sequencing data, DNA enrichment is a worthwhile option. DNA enrichment prior to sequence allows you to analyze target regions, such as the exome. The exome, which accounts for approximately 1% of the genome, contains the coding fragments of genes. Therefore, exome sequencing provides data from regions relevant in protein codification. Considering the costs of the technology and the challenges in the interpretation of whole-genome sequencing data, DNA enrichment is a

valuable option. Using DNA enrichment prior to sequence it is possible to analyze target regions, such as the exome.

The exome, which accounts for approximately 1% of the genome, contains the coding fragments of genes. Therefore, exome sequencing provides data from regions relevant in protein codification.

To perform exome enrichment before sequencing two different methods are mainly used: array-based hybridization and liquid-based hybridization. Both procedures rely on the hybridization of DNA of interest to target-specific probes. In the former process the probes are bond on a microarray surface and in the latter probes are in solution. Regarding the sequencing method itself several technologies are available, each one presenting different characteristics an features considering throughput, read length and accuracy.

Nonetheless, the main steps in the workflow are similar among different technologies and instruments. Firstly, a random fragmentation of DNA is performed followed by ligation of adapter sequences. Those adaptors allow us to amplify all fragments within the same PCR. After the amplification, an enrichment step might be considered depending on the aim of the experiment. If DNA enrichment is done, only target regions will be captured for the following step, sequencing. In order to sequence the fragments obtained, several cycles of cyclic reversible termination are performed with an image capture after each cyclic reversible termination. The reversible nature of the cycle allows the cleavage of the terminators and the repetition of the described cycle several times to sequence the fragment (114, 115).

Considering the amount of data generated by NGS, the parallel evolution of bioinformatics has been mandatory. The development of software has been completely imperative for data management and data analysis.

4 GENETIC COUNSELING

In 1975, the American Society of Human Genetics adopted the following definition of genetic counseling: "Genetic counseling is a communication process which deals with the human problems associated with the occurrence, or the risk of an occurrence, of a genetic disorder in the family. This process involves an attempt by one or more appropriately trained persons to help the individual or family to (1) comprehend the medical facts, including the diagnosis, probable course of the disorder, and the available management; (2) appreciate the way heredity contributes to the disorder, and the risk of recurrence in specified relatives; (3) understand the alternatives for dealing with the risk of occurrence; (4) choose the course of action which seems to them appropriate in view of their risk, their family goals, and their ethical and religious standards, to act in accordance with that decision; and (5) to make the best possible adjustment to the disorder in an affected family member and/or the risk of recurrence of that disorder."

Genetic counseling is an educational process that seeks to assist affected and/or at risk individuals to understand the nature of the genetic disorder, its transmission and the options open to them in management and family planning. The process requires the participation of one or more appropriately trained professionals to assist patients and / or their family to understand and cope with medical and non-medical aspects related to the disease trying to minimize the psychological stress that such information might cause. In genetic counseling, the way in which genetic counseling is provided is as important as the information given. Patients and / or their relatives may use the information provided in genetic counseling session to decide the most appropriate course of action considering the risk, family goals and the ethical and religious convictions. The ultimate goal of genetic counseling is enable patients and relatives to make the best possible adjustment to the disease and / or the risk of recurrence of it. In specific situations, such as common diseases with a relevant genetic component, the promotion of healthy habits in relation to the individual risk is also one of the objectives of genetic counseling.

In order to offer appropriate genetic counseling is necessary to establish an accurate diagnosis in the index case or, when it is not possible, try to determine the inheritance pattern of disease in the family. With this aim, the genetic counselor might use different tools and resources. The proper collection of genetic and clinical information is the first step in providing an adequate genetic counseling; otherwise, a wrong or incomplete collection of

data might lead to an inaccurate genetic counseling. Clinical data documentation of patients and their families is essential to have a complete picture of the case and correctly inform patients about the genetic disorder, associated risks and possible options available. Another crucial tool in genetic counseling is the correct drawing of the family tree, which helps to clearly visualize family structure, kinships and the most relevant clinical data. A family tree can also give information about the inheritance pattern of the genetic disease in the family, a crucial issue when determining genetic risks and identifying family members at risk.

In genetic counseling, apart from the information given, the way in which genetic counseling is provided is essential. Communication skills and empathy are relevant in order to achieve and effective communication process. Some studies have pointed out the importance of communication skills to explore the emotional state of patients or the level of comprehension in order to provide information in a comprehensive manner for patients and families. It is relevant to identify factors influencing comprehension, attitudes and decision making to improve genetic counseling practice and to ensure a proper attention to all patients and families. Previous publications have described a different risk perception for patients and / or families depending on the manner in which the advice was provided (118) or the vocabulary used (119, 120).

It is also important to be aware of the non-directiveness nature of genetic counseling as the genetic counselor provides information to family and patients that enable them to make a decision based on their beliefs and way of life. Genetic counseling should promote the autonomy and the decision-making process in the patient and / or relatives by providing the right information in a non-directiveness manner. The counselor should attempt to provide the information, tools and resources needed to enable the patient to make a real informed choice (121-123).

In the context of congenital malformations genetic counseling is essential. There are many situations were genetic counseling is indispensable, for instance when there is an ultrasound finding in an ongoing gestation or a termination of pregnancy due to birth defects, after a spontaneous miscarriage or when a newborn presents a congenital malformation. Several studies have proved the need for genetic counseling regarding prenatal screening and prenatal diagnostic tests to improve the information provided,

the understanding and also to reduce decision-making conflicts without increasing maternal anxiety (119).

Although the high incidence of congenital malformations in miscarriages, stillborns or liveborns the mechanism underlying those diseases are poorly understood. In the absence of well established cause, genetic counseling becomes quite difficult and only a strong family history of congenital anomalies may provide valuable information concerning risk determination. In most cases, no family history is reported with a unique case identified. If the malformation is part of a more complex syndrome the chances to find a diagnostic increase; otherwise, when an isolated malformation is described the likelihood of finding the cause is drastically reduced.

Since the genetic aspects of congenital malformations are largely unknown, the enforcement of new technologies to explore different genetic mechanism is valuable for the diagnosis and to provide an adequate genetic counseling to patients and families. Therefore, it is worthwhile to draw on to recently developed technologies in research as well as in clinical practice to elucidate the genetic mechanism underlying congenital malformations with the aim of improving patients' and family care. Nevertheless, it is important to take into consideration that the emerging molecular techniques involve a challenge in genetic counseling considering that they extend the diagnostic capability and the potential impact of variants of uncertain clinical significance. The increase of the difficulty related to new molecular techniques make genetic counseling even more necessary related to prenatal testing and congenital malformations.

HYPOTHESIS

1) The etiology of congenital malformations is heterogeneous with several factors involved, including genetics and environment.

2) Genetic rearrangements, point mutations, epigenetic alterations or a combination of them altering developmental pathways may contribute in a relevant manner in the etiology of congenital malformations.

3) Novel techniques developed in recent years for genetic, genomic and epigenomic analyses may provide valuable knowledge in the study of congenital malformations.

4) The identification of alterations underlying congenital malformations and the subsequent characterization would ensure proper genetic counseling to the family, with relevant implications for family planning.

5) The application of molecular techniques to improve patients care' and diagnosis should be accompanied by an appropriate genetic counseling and a high rate of acceptability by patients and families. A detailed and comprehensive explanation in the context of genetic counseling would lead to higher acceptability of the procedure. Systematic studies of patients' opinions about the provided service may contribute to the ultimate goal of improving genetic counseling with new technologies.

OBJECTIVES

1) Identify genomic rearrangements in fetuses with different types of congenital malformations by using BAC aCGH, oligonucleotides aCGH and SNP arrays and characterize those alterations in order to determine their possible implication in the etiology of congenital malformations. Explore the possible contribution of mosaicism, uniparental disomies and multiple hits in congenital malformations.

2) Explore the clinical utility of chromosomal microarray analysis in invasive prenatal diagnosis in comparison to other available techniques such as G-banding karyotype analysis, quantitative fluorescent PCR and Multiplex Ligation-dependent Probe Amplification in order to determine the detection rate of each technique as well as the identification of variants of uncertain clinical significance, costs and turn-around time.

3) Analyze the strengths and weaknesses of genetic counseling in relation to congenital malformations and the use of novel high throughput molecular techniques, such as chromosomal microarray analysis, in order to provide and appropriate genetic counseling to the families.

4) Investigate the contribution of epigenetic alterations to the etiology of isolated congenital malformations using isolated congenital heart defects as a disease model and DNA from the targeted tissue. Analyze how alterations in methylation affect gene expression to establish their implication in congenital heart defects. Define candidate genes and pathways involved in congenital heart disease

5) Elucidate the role of de novo and inherited point mutations in congenital malformations using exome sequencing data from two selected and well-phenotyped cohorts, one with isolated heart hyopoplasia and the other one with neural tube defect plus Arnold-Chiari malformation.

CHAPTER 1

Contribution of rare copy number variants to isolated human malformations

Clara Serra-Juhé, Benjamín Rodríguez-Santiago, Ivon Cuscó, Teresa Vendrell, Núria Camats, Núria Torán, Luis A. Pérez-Jurado

Accepted in PlosOne

The emergence during the last decade of molecular technologies based on CMA has implied a revolution in genetics field. The study of deletions as well as duplications with high resolution and in a genome-wide manner has provided valuable information regarding genome structure and has revealed a relevant contribution of genetic rearrangements in genetic diseases.

CMA, including array CGH and SNP microarrays, has allowed the identification of causative chromosomal imbalances in 15-24% of patients with multiple congenital anomalies and/or intellectual disability. Concerning the study of congenital malformations in fetuses, retrospective studies using CMA have reported a detection rate of 8-15% of causative segmental aneusomies. Those rearrangements included recurrent genomic disorders frequently associated with congenital malformations, as 22q11.2 microdeletion.

Despite those studies, few data is available regarding isolated congenital malformations. In order to elucidate the role of deletions as well as duplications in birth defects, a study based on CMA mainly in isolated congenital malformations was designed. Cryptic genomic rearrangements were analyzed in a cohort of fetuses with isolated congenital malformations and in a second cohort of fetuses with multiple birth defects, with the purpose of exploring the mechanisms underlying congenital anomalies.

Contribution of rare copy number variants to isolated human malformations

Clara Serra-Juhé, Benjamín Rodríguez-Santiago, Ivon Cuscó, Teresa Vendrell, Núria Camats, Núria Torán, Luis A. Pérez-Jurado

ABSTRACT

Background: Congenital malformations are present in approximately 2-3% of live born babies and 20% of stillborn fetuses. The mechanisms underlying the majority of sporadic and isolated congenital malformations are poorly understood, although it is hypothesized that the accumulation of rare genetic, genomic and epigenetic variants converge to deregulate developmental networks.

Methodology/Principal Findings: We selected samples from 95 fetuses with congenital malformations not ascribed to a specific syndrome (68 with isolated malformations, 27 with multiple malformations). Karyotyping and multiplex ligation-dependent probe amplification (MLPA) discarded recurrent genomic and cytogenetic rearrangements. DNA extracted from the affected tissue (46%) or from lung or liver (54%) was analyzed by molecular karyotyping. Validations and inheritance were obtained by MLPA. We identified 22 rare copy number variants (CNV) [>100kb, either absent (n=7) or very uncommon (n=15, <1/2,000) in the control population] in 20/95 fetuses with congenital malformations (21%), including 11 deletions and 11 duplications. One of the 9 tested rearrangements was *de novo* while the remaining were inherited from a healthy parent. The highest frequency was observed in fetuses with heart hypoplasia (8/17, 62.5%), with two events previously related with the phenotype. Double events hitting candidate genes were detected in two samples with brain malformations. Globally, the burden of deletions was significantly higher in fetuses with malformations compared to controls.

Conclusions/Significance: Our data reveal a significant contribution of rare deletion-type CNV, mostly inherited but also *de novo*, to human congenital malformations, especially heart hypoplasia, and reinforce the hypothesis of a multifactorial etiology in most cases.

INTRODUCTION

A potentially lethal or disabling major malformation occurs in 2–3% of live born infants and 20% of stillborn fetuses [1]. Congenital malformations have become the main cause of infant mortality during the first years of life [2] and are associated with long term morbidity [3,4]. In particular, congenital heart defects (CHD) represent a high percentage of clinically significant birth defects. The incidence of CHD is approximately 8 per 1,000 live births making CHD the most common malformation [5,6].

Congenital malformations often occur in the setting of multiple congenital anomalies, including dysmorphic facial features, developmental aberrations of different organs, or growth abnormalities [7,8]. In these cases with a more complex syndrome, chromosomal aberrations are a frequent cause of disease, although point mutations in developmental or metabolic genes have also been described in specific syndromes [9,10]. Standard karyotyping can detect numerical and structural anomalies larger than 5–10 Mb

and other techniques, such as fluorescent in situ hybridization (FISH) [11] or MLPA [12-14], allow the identification of submicroscopic chromosomal imbalances. In the last decade, the development of molecular karyotyping by array comparative genomic hybridization (aCGH) or single-nucleotide-polymorphism (SNP) microarrays, globally termed chromosomal microarray analysis (CMA), has allowed the detection of as much as 15–24% of causative segmental aneusomies in patients with multiple congenital anomalies and/or intellectual disability [15,16]. Retrospective studies in fetuses with multiple malformations have obtained a detection rate of causative chromosomal imbalances from 8 to 15% by using CMA [17-19], and the clinical utility of a targeted CMA has been demonstrated in standard invasive prenatal diagnosis [20,21]. CHD are among the malformations in which genomic rearrangements have been shown to play a major role. For instance, microdeletions at 22q11.2 [22,23] and microduplications at 1q21.1 [24,25] are a common cause of conotruncal heart defects .

In an important proportion of cases, only one malformation is detected without the presentation of other minor or major defects. Although some isolated congenital malformations can be caused by environmental risk factors, such as maternal diseases or exposure to teratogenic agents during pregnancy [4], there is strong evidence that genetics plays a major role, as epidemiological studies have shown an increased risk of this type of anomalies in siblings and offspring of individuals with sporadic congenital malformations, as well as increased paternal age and high concordance in monozygotic twins [26,27,28]. A small percentage can be attributed to point mutations in development related genes [29,30], although this type of genetic alterations have been insufficiently tested until recently.

Submicroscopic deletions and duplications may play a significant role in the etiology of this condition, either as direct cause or as possible genetic risk factor for isolated congenital anomaly [31]. Nevertheless, the mechanisms underlying the majority of non-chromosomal or sporadic congenital malformations are poorly understood.

Finding the cause of congenital malformations is necessary to better understand the pathophysiological basis of these developmental anomalies and define disease risks, both critical elements to ensure proper genetic counseling and disease prevention. Genetic counseling has become more relevant in this area considering not only the recurrence risk of healthy parents after having an index case, but also that more individuals with congenital malformations are living into adulthood due to advances in medical and surgical care and may have the opportunity to reproduce [7].

We have searched for cryptic genomic rearrangements in fetuses with isolated congenital malformations and fetuses with more than one congenital anomaly. Our data illustrate a significant contribution of rare deletion-type CNV, mostly inherited but also *de novo*, to human congenital malformations. These genomic rearrangements could represent the single genetic etiology of the disease, perhaps as part of a more complex syndrome without other recognizable manifestations at this stage of development, or genetic susceptibility factors contributing to the mutational load in multifactorial disorders.

METHODS

Ethics Statement

All studies were performed as part of an expanded diagnostic protocol approved by the Medical Ethical Committee of the Vall d'Hebron Hospital, after receiving written informed consent from the family.

Samples / Patients

Fetuses were selected from medically terminated pregnancies between 17 and 22 weeks of gestation owing to one or more malformations with bad prognosis detected during pregnancy. Samples were collected from frozen tissues stored in the Tissue Bank of Vall d'Hebron Hospital. A complete fetopathological examination had been performed and the samples were classified in two different groups: 1) 68 samples with an isolated congenital malformation, including 33 with isolated CHD, 26 with isolated central nervous system (CNS) malformation and 9 with isolated renal malformation; 2) 27 fetuses with more than a unique malformation. Prenatal GTG banding chromosome analysis was normal for all 95 fetuses. An overview of the clinical features of the fetuses included in the study is summarized in table 1 (detailed in Tables S1-S4).

Parental blood samples were collected in cases in which an alteration was identified.

MALFORMATION	SAMPLES
Congenital heart disease	
Conotruncal defect	13
Heart hypoplasia	17
Other	3
Central nervous system malformation	
Neural tube defect	16
Holoprosencephaly	3
Hydrocephalus	3
Ventriculomegaly	3
Agenesis of the corpus callosum	1
Renal malformations	
Agenesis	5
Dysplasia	3
Nephronophthisis	1
Multiple malformations	27

Table 1. Overview of malformations in the 95 analyzed fetuses.

DNA extraction from tissue and blood samples

In fetuses with an isolated congenital malformation, the affected tissue (heart, brain or kidney) was obtained when available (n=44); liver or lung tissue was used for the remaining samples with insufficient target tissue (n=24). For fetuses with multiple congenital anomalies (n=27), liver or lung tissue was used. Parental DNA was isolated from total blood. DNA was extracted using the Gentra Puregene Blood kit (Qiagen) according to manufacturer's instructions.

Multiple Ligation-Dependent Probe Amplification (MLPA)

Genomic rearrangements in subtelomeric regions (P036 and/or P070, MRC Holland) as well as recurrent microdeletion or microduplication syndromes (custom made, Table S5) were also discarded prior to selection by using two MLPA panels.

An MLPA assay was also designed to validate the genomic alterations detected by CMA and to study the inheritance in those cases with available parental samples. A total of 100 ng of genomic DNA from each sample was subject to MLPA using specific synthetic probes [Table S6] designed to target the specific CNV detected by different types of array. All MLPA reactions were analyzed on an ABI PRISM 3100 Genetic analyzer according to manufacturers' instructions. Each MLPA signal was normalized and compared to the corresponding peak height obtained in control samples [32,33].

Molecular karyotyping by CMA

The entire cohort was studied by using BAC (Bacterial Artificial Chromosome) aCGH. DNA samples (1ugr) were labeled by random priming with Cy3-dCTP and Cy5-dCTP and hybridized against a reference pool of the same gender. Samples were hybridized onto a BAC aCGH containing 5,600 clones with a backbone mean coverage of ~1Mb and increased density in hotspot regions for genomic rearrangements (subtelomeres, pericentromeres and regions flanked by segmental duplications). Analyses of BAC-aCGH data were performed as previously described [32]. A total of 25 samples were also studied by using an oligonucleotide Agilent H244K aCGH. Samples were processed and hybridized according to manufacturer's recommendations (Agilent Protocol v6.0, ref. G4410-90010). This technique allowed us to validate and better map the breakpoints of the alterations detected by BAC aCGH, as well as to increase the resolution of the study in samples in which no alteration had been detected using BAC aCGH. Only CNVs with genes, longer than 100kb and with a frequency in control samples lower than 1/2,000 were considered. The frequency of each CNV in the control population was determined using 1M Illumina SNP array data from a control database of 8,329 samples already reported [34], along with data from 1,991 Spanish adult samples from the Spanish Bladder Cancer/EPICURO study including 1034 patients with urothelial cell carcinoma of the bladder and 957 hospital-based generally healthy controls with a mean age of 63.7 years [35].

DNA from 70 samples was studied by using the 370K Illumina SNP array. This technique permitted us to increase the resolution in samples in which no alteration had been identified using BAC aCGH. Moreover, using SNP array uniparental disomy and regions with high level of homozigosity were studied. Copy number changes were identified using the PennCNV software with stringent filtering, as previously described [35]. Only CNVs with genes, longer than 100kb and with a frequency in control samples lower than 1/2,000 were considered. A search for possible mosaic copy number and copy neutral changes was also performed using the MAD algorithm [36].

Genetic counseling

Genetic counseling was offered to all couples when an alteration was identified in order to explain the findings and the need for further testing including parental samples. After the study of the parents' samples, follow-up counseling was provided along with a written report explaining the alteration, the putative relation with the phenotype and the implications to the family.

Bioinformatic and statistical analyses

The frequency of each CNV in the population was determined using 1M Illumina SNP array data from a control database of 8,329 samples already reported

[34], along with data from 1,991 Spanish adult samples studied in our laboratory with the same arrays [35].

In addition, already available data from a randomly selected cohort of 168 generally healthy Spanish adult control individuals (Spanish Bladder Cancer/EPICURO study genotyped with Illumina 1M SNP array [35]) was used in order to compare the different frequencies of rare rearrangements between controls and fetuses with congenital malformations (global CNV burden and CNV combinations). In order to avoid or minimize a possible bias due to the different detection yield of the array platforms used, we only considered alterations larger than 100kb that should be detected with any of the platform arrays. For the comparative analyses, only CNVs with genes, a minimum length of 100kb and a frequency lower than 1/2,000 were considered. Alterations totally overlapping with segmental duplications were also excluded to minimize biases due to the different probe coverage among microarray platforms.

Gene content and enrichment analyses

The gene content (genes included or disrupted) of the rare CNVs identified in the cohort of fetuses was analyzed using a computational resource, Consensus Path DB [37], to obtain an overview of the pathways which could be altered, with the aim of identifying overrepresented pathways. The p-value of each pathway was considered ($p>0.05$).

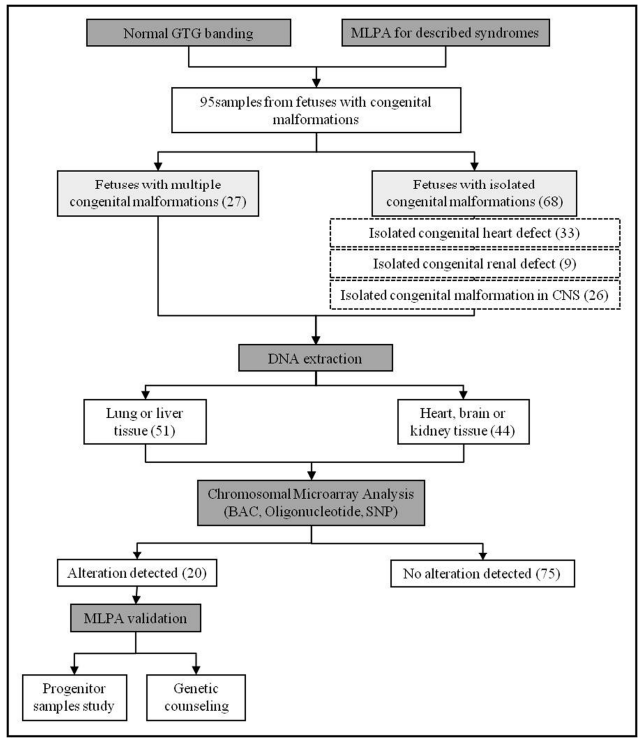


Figure 1. Strategy followed to study samples of fetuses with congenital malformations. MLPA: multiplex ligation-dependent probe amplification; CNS: central nervous system; BAC: bacterial artificial chromosome; SNP: single nucleotide polymorphism.

RESULTS

Prenatal GTG banding chromosome analysis was normal for all 95 fetuses. Known microdeletion/microduplication syndromes and subtelomeric genomic rearrangements were also discarded by MLPA in all cases. All samples were first studied using BAC aCGH and then by oligonucleotide or SNP array.

Globally, CMA detected 22 CNVs fulfilling the established criteria (>100kb, gene containing and present in <1/2000 controls) in 20 samples (21.05%), 11 deletions and 11 duplications (100.6-2,207kb in length), with 2 samples harboring two rearrangements. MLPA probes were designed to define the inherited or de novo nature of the CNVs in all 9 cases from whom parental samples were available. In 8 cases the alterations were inherited while the rearrangement was *de novo* in a single case. The detected alterations are listed in table 2, including information about the genomic coordinates, size, arrays used for detection and validation, inheritance and genes included in the region. Among the 22 alterations identified, 7 (4 duplications and 3 deletions) have never been found in the 10,320 adults used as controls. Two aberrations, both of them identified in fetuses with CHD, overlap with previously reported alterations associated with developmental anomalies and are likely the underlying genetic cause [38-40]: 1) A 363kb *de novo* deletion in 16q24.1, encompassing five genes (*FOXF1, FOXC2, MTHFSD, FLJ30679* and *FOXL1),* was detected in a fetus with left heart hypoplasia (case 1); 2) the recurrent 2Mb 15q13.3 deletion was identified in a fetus with right heart hypoplasia as well as in the healthy mother. The remaining 20 rearrangements have not been described in patients with disease.

Although not included in the listed 22 aberrations because its reported frequency in controls is 0.14% (>1/1,000), we also detected the recurrent 1.6Mb 16p13.11 duplication in two samples, one case of CNS malformation (neural tube defect and Arnold/Chiari malformation) and another with multiple malformations (anal imperforation, right heart hypoplasia and esophagus atresia). The reciprocal deletion of this region has been clearly associated with increased risk for congenital malformations and developmental difficulties but published data for the duplication are not clearly conclusive [41].

In order to define whether the global burden of rare CNVs in the fetuses with congenital malformations was or not significantly increased, we compared it with CNV data in a cohort of 168 control subjects analyzed with the Illumina 1M SNP array. For consistency, only CNVs larger than 100kb, containing genes, not totally overlapping with segmental duplications, and found at a frequency <1/2,000 were considered (listed in Table S7). Rare CNVs fulfilling criteria were identified in 17.86% of the control samples including 2 samples with 2 alterations. These rare CNVs in controls were predominantly duplications (78.12% vs 21.88% deletions). Thus, the global CNV burden in malformed fetuses was only slightly increased with respect to that in normal controls (21.05% vs 17.86%), although the number and proportion of deletions was significantly higher in malformed fetuses.

The proportion of samples with rearrangements was different between the different groups of malformations, being higher in fetuses with CHD (10/33 samples, 30.30%) and even higher if only heart hypoplasia was considered (8/17, 47.06%). The difference in aberrations frequency between groups was statistically significant comparing fetuses with heart hypoplasia and controls (p=0.009). The difference in the frequency of deletion-type CNV between controls and cases was also statistically significant (p=0.03) and more evident comparing only fetuses with heart hypoplasia and controls (p=0.001). These differences were due to the increased number of deletions, but not duplications, in cases with congenital malformations (Table 3).

Case #	Malformation	Gain /Loss	Region	Length (kb)	Start	End	Array used	Inheritance	Genes in the region	Control frequency (10,320)
6	CHD	Gain	5q35.2	297.3	175798945	176096236	SNP	-	*ARL10, CLTB, EIF4E1B, FAF2, GPRIN1, HIGD2A, NOP16, PCDH24, RNF44, SNCB, TSPAN17*	0
2	CHD	Loss	16q24.1	363.5	8638212	86745576	SNP	*De novo*	*FOXF1, FOXC2, MTHFSD, FLJ30679, FOXL1*	0
4	CHD	Loss	16q23.3	120.3	83869776	83990089	SNP	Paternal	*MLYCD, OSGIN1*	1
23	CHD	Loss	6p25.1	139.4	5249765	5389206	Oligonucleotide	-	*LYRM4, FARS2*	1
24	CHD	Gain	2p25.3	108.5	3579585	3688127	SNP	-	*COLEC11, RNASEH1, RPS7*	1
29	CHD	Gain	3p26.3	246.5	2869944	3116438	SNP	Paternal	*CNTN4, IL5RA*	1
1	CHD	Loss	15q13.3	2207	30755144	32962148	BAC and Oligo	Maternal	*FAN1, MTMR10, TRPM1, LOC283710, KLF13, OTUD7A, CHRNA7*	1
3	CHD	Loss	13q21.2	288.1	60410392	60698463	BAC and SNP	Maternal	*DIAPH3*	2
15	CHD	Gain	10q26.3	181.4	134572478	134753880	Oligonucleotide	-	*INPP5A, NKX6-2, TTC40*	2
13	CHD	Gain	9p21.1	369.2	28659143	29028380	SNP	-	*LINGO2*	4
48	CNS	Loss	17p12	1383.3	14090300	15473646	Oligonuclotide	-	*COX10, CDRT15, HS3ST3B1, PMP22, TEKT3, CDRT4, FAM18B2*	0
57	CNS	Loss	6p22.2	112.9	24401654	24514569	SNP	Maternal	*ALDH5A1, GPLD1, MRS2*	0
		Loss	12p12.3	311.6	18337494	18649057	SNP	Paternal	*PIK3C2G*	4
50	CNS	Gain	1p33	139.2	46814268	46953453	Oligonuclotide	-	*FAAH, DMBX1, KNCN*	0
		Gain	10q11.22	197.3	46951237	47148490	Oligonuclotide	-	*SYT15, GPRIN2, PPYR1*	4
59	CNS	Gain	5p13.2	150.3	37411054	37561355	BAC and Oligo	-	*WDR70*	4
53	CNS	Loss	18q22.1	2324.0	63733025	66057032	SNP	Maternal	*CDH19, DSEL., LOC643542*	4
65	Renal	Loss	4q12	883	52798624	53681594	SNP	Maternal	*DANCR, LRRC66, SGCB, SNORA26, SPATA18, USP46*	1
72	Multiple	Gain	10p14	128.4	11815455	11943885	SNP	-	*C10orf47, LOC219731*	0
82	Multiple	Gain	5q35.3	752.9	179833485	180586413	SNP	-	*BTNL3, BTNL8, BTNL9, CNOT6, FLT4, LOC729678, MGAT1, OR2V2, OR2Y1, SCGB3A1, ZFP62*	0
85	Multiple	Loss	7p14.1	130.1	40264889	40394987	SNP	-	*C7orf10*	1
84	Multiple	Gain	1p34.1	100.6	46252717	46353332	SNP	-	*MAST2*	5

Table 2. Summary of copy number variations detected in 95 samples of fetuses with congenital malformations. Control frequency refers to the frequency of the same type of rearrangement found in the fetus, deletion or duplication. Hg19 assembly. CHD, congenital heart defect; CNS: central nervous system; SNP: single nucleotide polymorphism; BAC: bacterial artificial chromosome.

The frequency of individuals with more than one CNV hit fulfilling the established criteria was not different between cases and controls, around 2% (Table 3).

Regarding the overrepresentation analysis, phosphatidylinositol phosphate metabolism was the only pathway significantly overrepresented in cases with respect to controls. Three genes directly involved in this pathway, *PIK3C2G*, *GPLD1* and *INPP5A*, are included in the CNVs identified in two fetuses. Interestingly, two of these genes are located in two deletions found in the same sample, a fetus with holoprosencephaly. One deletion encompassing three genes, *ALDH5A1*, *GPLD1* and *MRS2*, was inherited from the mother, while the other one including only one gene, *PIK3C2G*, was inherited from the father (Fig.2). An additional sample with two events was a fetus with hydrocephalus found to have two duplication CNVs, on chromosome bands 1p33 (including the genes *FAAH*, *DMBX1* and *KNCN*) and 10q11.22 (containing the genes *SYT15*, *GPRIN2* and *PPYR1*), but parental samples were not available in this case.

No large stretches of homozygosity suggestive of parental consanguinity or uniparental disomy (UPD) were identified in any sample (70/95 fetuses studied with SNP arrays). In addition, despite the use of DNA from the affected tissue in 46% of cases, no events of copy number or copy neutral changes suggestive of somatic mutations were detected.

GROUP	ALTERATIONS	DELETIONS	DUPLICATIONS	DOUBLE HIT	SAMPLES
Controls (168)	32	7 (4.2% / 21.88%)	25 (14.9% / 78.12%)	2 (1.19%)	30 (17.86%)
Fetuses (95)	22	11 (11.6% / 50%)	11 (11.6% / 50%)	2 (2.11%)	20 (21.05%)
CHD (33)	10	5 (15.2% / 50%)	5 (15.2% / 50%)	0 (0%)	10 (30.30%)
*Heart hypoplasia (17)	8	5 (29.4% / 62.5%)	3 (17.6% / 37.5%)	0 (0%)	8 (47.06%)
CNS malformations (26)	7	4 (15.4% / 57.14%)	3 (11.5% / 42.86%)	2 (7.69%)	5 (19.23%)
Renal malformations (9)	1	1 (11.1% / 100%)	0 (0% / 0%)	0 (0%)	1 (11.11%)
Multiple malformations (27)	4	1 (3.7% / 25%)	3 (11.1% / 75%)	0 (0%)	4 (14.81%)

Table 3. Comparisons of rare copy number changes >100kb detected in the fetuses with congenital malformations and controls. In brackets the proportion of samples with the CNV and the proportion of the specific type of rearrangement. *A subcategory of CHD only considering heart hypoplasia has been added to the table due to remark the different frequency of CNVs with respect to the other CHD.

CHD, congenital heart defect; CNS: central nervous system.

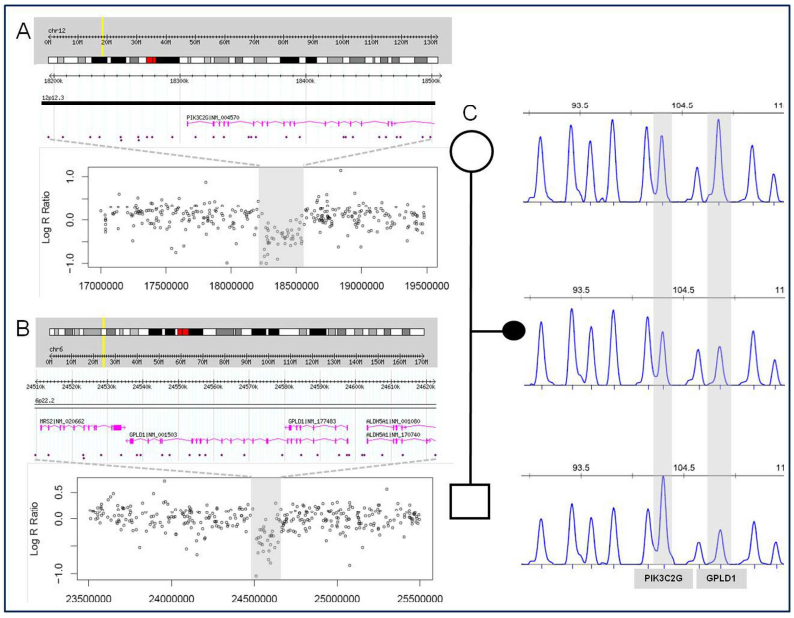


Figure 2. Detection, validation and inheritance of the two imbalances in case 57. A and B: Ideogram showing the regions included in the deletions and plot of Log R Ratio of SNP array; it shows the SNPs included in the deletions, as well as the ones included in the flanking regions. C: MLPA showing the inheritance of both deletions.

DISCUSSION

Chromosomal aberrations have been reported as a frequent cause of congenital malformations, especially when they are associated with growth or developmental delay, malformations affecting a second organ or dysmorphic features [6,16,18,42]. Many of the chromosomal unbalances associated with such syndromes are large and encompass multiple genes. A detection rate of 10% chromosomal abnormalities, including one marker chromosome and one rearrangement of 9Mb, has been reported studying by aCGH a population of 50 fetuses with more than three malformations [42]. A yield of 17%, considering aneuploidies and known syndromes, was found in a cohort of 638 newborns with birth defects. The role of submicroscopic deletions and duplications in isolated congenital malformations has been documented for CHD with the identification of 18 putatively pathogenic CNVs (17.1%) in 105 samples from infants with isolated CHD [31], including 4 recurrent rearrangements in 22q11.2 (responsible of DiGeorge syndrome), 17p11 (causative of Smith-Magenis syndrome), and 1q21.1, a large alteration of 14Mb, and an aberration with no genes.

In our series, chromosomal alterations detected by karyotyping and cryptic alterations in subtelomeric regions or known microdeletion / microduplication syndromes were previously excluded. Rare CNVs larger than 100kb were detected in 21% of fetuses with prenatally detected malformations, with a yield of 30.3% in fetuses with CHD. The CNV burden was slightly but significantly higher in malformed fetuses compared with controls (21.05% vs 17.86%). Deletions were also more prevalent in cases than controls (50% vs 21.88%), as well as larger in size. As

expected, large CNVs and mostly deletions are more likely to affect gene expression with relevant effect on developmental pathways. The difference in the detection rate in comparison with other studies might be explained by the different selection criteria and resolution of the array platforms used.

We detected abnormalities previously reported as causative of CHD in two cases. A 363kb *de novo* deletion in 16q24.1 encompassing the *FOX* gene cluster was detected in a fetus with left heart hypoplasia. Overlapping deletions have been previously reported in patients with alveolar capillary dysplasia, misalignment of pulmonary veins and distinct malformations including congenital heart defect, specifically hypoplastic left heart [38]. Deletion of *FOXF1* is thought to be responsible for alveolar capillary dysplasia while *FOXC2* is related to the lymphoedema-distichiasis syndrome. Larger deletions, as in our case, may cause a more complex syndrome which includes CHD likely due to additive effects of haploinsufficiency for contiguous genes [38].

We also identified the recurrent 2Mb 15q13.3 deletion in a fetus with right heart hypoplasia, inherited from the healthy mother. Interestingly, the brother of the mother also had a cardiac malformation on anamnesis but he rejected to be studied. Deletions and duplications at 15q13.3 have been related to different developmental anomalies, such as dysmorphic features, intellectual disability, seizures, schizophrenia, and in 17% of patients congenital heart defects [39]. Based on previous studies in animal models, *KLF13,* encoding the Kruppel-like factor 13, is the best candidate gene for the cardiac defects associated with the 15q13.3 deletion. *KLF13* knockdown in Xenopus embryos caused atrial septal defects and hypotrabeculation similar to those observed in humans or mice with hypomorphic *GATA4* alleles [40]. Rearrangements in this region show incomplete penetrance and variable expressivity, with various cases in which the deletion or duplication is inherited from a healthy progenitor, as in our case [39]. Given this incomplete penetrance of clinical manifestations and the relatively low proportion of patients affected by cardiac disease, it is assumed that factors other than the 15q13.3 deletion should also be involved in the appearance of the clinical traits. In this case, no additional genomic alterations were detected.

Among the additional rare rearrangements identified in fetuses with malformations, all tested were inherited from an apparently healthy progenitor, which is consistent with previous data [31,39]. The rarity and gene content of some of these rearrangements suggest their possible pathogenic implication in congenital malformations. Nevertheless, like in some recurrent microdeletion syndromes, the existence of healthy carriers among progenitors and the adult population indicates that the rearrangements are not the only cause of the disease. Considering the epidemiologic evidence for multifactorial etiology of major malformations, these rearrangements could represent just one of the several factors involved. In this regard, a case with holoprosencephaly showed two deletions, one inherited from the mother and the other from the father, both harboring genes of the same pathway (phosphatidylinositol metabolism). Two duplication-type CNV events were also found in a fetus with hydrocephalus, although parental samples were not available to determine their inheritance pattern. However, candidate genes for brain malformation were also located in both CNVs: *DMBX1* codes for a diencephalon-mesencephalon homeobox implicated in brain development and *GPRIN2* encodes a G-protein regulated inducer of neurite overgrowth involved in formation and extension of neurite-like processes [43,44]. Given the very low frequency of these alterations in controls, the functional relationship of altered genes and their inheritance from different progenitors at least in the first case, it is logical to propose that the double hits may have contribute to the fetal malformations by additive effect of the

CNVs on altering developmental regulation. A two-hit model with several recurrent and non-recurrent CNVs has been already reported for neurobehavioral and relatively severe phenotypes [45].

In addition, we also detected the 16p13.11 1.6Mb duplication in two cases with different phenotypes. This duplication has been found in 0.14% normal adult controls (12/8,329 controls) and in 0.27% patients with developmental delay and/or malformations (42/15,767) [34]. Given the higher frequency of this duplication in our series (2%) as well as in reported patients with developmental anomalies [41], the data highly suggest that this CNV is indeed a susceptibility variant for developmental disorders including congenital malformations. The different phenotypes related to the microduplication might also be related to the concurrence of this contiguous gene alteration with other undefined genetic or environmental second hit. Depending on the concurrence of other factors that may contribute to reach the gene dysfunction threshold in a specific tissue or developmental time, the phenotype would correspond to different diseases or malformations. Although additional CNVs were not found with increased frequency in cases with respect to controls in our cohort, including the two cases with 16p13.11, secondary events of other type, such as point mutations or epimutations cannot be ruled out.

On the other hand, UPD and shared homozygosity regions were discarded by SNP array and mosaic alterations were also not identified. Although the number of samples studied is low, UPD does not seem to be a highly frequent cause of isolated congenital defect. Since DNA from the affected tissue was analyzed in 46 samples, we can also conclude that mosaicism for large rearrangements in the abnormally developed tissue is not frequent in isolated congenital malformation.

In addition to the most common aneuploidies and genomic disorders also detected by karyotyping and targeted assays, CMA significantly increases the detection yield of cryptic segmental aneuploidies in fetuses with congenital malformations. The highest yield for rare CNVs was found in samples with hypoplasia of the left/right heart, doubling the frequency of any other group of malformations and suggesting a higher genetic component for this type of malformation, which is consistent with its higher heritability [46,47]. Recently published data [48] also suggests a higher frequency of rearrangements in patients with left heart hypoplasia comparing with controls even though the difference is only statistically significant for those aberrations smaller than 60kb. However and from a clinical perspective, CMA can detect the single causative alteration in a relatively low percentage of cases with isolated congenital malformations, about 2% once the most common aneuploidies and recurrent rearrangements are discarded. Therefore, although many rare CNVs detectable by CMA, like those reported here, presumably contribute to the disorder, they should be considered as variants of unknown significance until more information is available to better predict phenotype based on genotype.

Accumulation of multiple rare genomic and epigenetic variants converging to deregulate developmental genes leading to mutational loading of developmental networks may cause congenital malformations [49]. Rare copy number variants, point mutations and/or epigenetic variations, either inherited or *de novo*, can impact gene function or alter dosage and contribute to mutational load. Changes affecting multiple genes and networks related to development may induce developmental anomalies. This concept implies that if threshold levels of flux are exceeded, compensatory mechanisms may fail, leading to an inadequate development. This hypothesis has been tested in mouse model and some results suggest that the accumulation of alterations in regulatory development networks results in an

inadequate development [50]. Although it is reasonable to expect homologous genes to behave similarly in humans, more evidence supporting this hypothesis is needed. Further studies, including whole genome sequencing and epigenomic analyses as well as expression profiles of genes related to development should be done in order to improve the knowledge of the etiology and the diagnostic tools for isolated congenital malformations.

ACKNOWLEDGMENTS

We would like to thank Verena Terrado and Raquel Flores for technical assistance, as well as Gabriela Palacios and the anonymous reviewer of the manuscript for critical comments and suggestions. We also would like to thank all participant families for their consent and support.

REFERENCES

1 Kalter H, Warkany J (1983) Medical progress. Congenital malformations: etiologic factors and their role in prevention (first of two parts). N Engl J Med 308(8):424-31.

2 Kalter H, Warkany J (1983) Congenital malformations (second of two parts). N Engl J Med 308(9):491-7.

3 De Galan-Roosen AE, Kuijpers JC, Meershoek AP, van Velzen D (1998) Contribution of congenital malformations to perinatal mortality. A 10 years prospective regional study in The Netherlands. Eur J Obstet Gynecol Reprod Biol 80(1):55-61.

4 Botto LD, Correa A (2003) Decreasing the burden of congenital heart disease: an epidemiologic evaluation of risk factors and survival. Prog Pediatr Cardiol 18:111–21.

5 Mitchell SC, Sellmann AH, Westphal MC, Park J (1971) Etiologic correlates in a study of congenital heart disease in 56,109 births. Am J Cardiol 28(6):653-7.

6 Richards AA, Santos LJ, Nichols HA, Crider BP, Elder FF, et al. (2008) Cryptic chromosomal abnormalities identified in children with congenital heart disease. Pediatr Res 64(4):358-63.

7 Pierpont ME, Basson CT, Benson DW, Jr., Gelb BD, et al. (2007) Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation 115(23):3015-38.

8 Thienpont B, Mertens L, de Ravel T, Eyskens B, Boshoff D, et al. (2007) Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. Eur Heart J 28(22):2778-84.

9 Ogata T, Yoshida R (2005) PTPN11 mutations and genotype-phenotype correlations in Noonan and LEOPARD syndromes. Pediatr Endocrinol Rev 2(4):669-74.

10 Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, et al. (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 42(9):790-3.

11 Popp S, Schulze B, Granzow M, Keller M, Holtgreve-Grez H, et al. (2002) Study of 30 patients with unexplained developmental delay and dysmorphic features or congenital abnormalities using conventional cytogenetics and multiplex FISH telomere (M-TEL) integrity assay. Hum Genet 111(1):31-9.

12 Rooms L, Reyniers E, van Luijk R, Scheers S, Wauters J, et al. (2004) Subtelomeric deletions detected in patients with idiopathic mental retardation using multiplex ligation-dependent probe amplification (MLPA). Hum Mutat 23(1):17-21.

13 Bendavid C, Dubourg C, Pasquier L, Gicquel I, Le Gallou S, et al. (2007) MLPA screening reveals novel subtelomeric

rearrangements in holoprosencephaly. Hum Mutat 28(12):1189-97.

14 Koolen DA, Nillesen WM, Versteeg MH, Merkx GF, Knoers NV, et al. (2004) Screening for subtelomeric rearrangements in 210 patients with unexplained mental retardation using multiplex ligation dependent probe amplification (MLPA). J Med Genet 41(12):892-9.

15 Vissers LE, de Vries BB, Osoegawa K, Janssen IM, Feuth T, et al. (2003) Array-based comparative genomic hybridization for the genome wide detection of submicroscopic chromosomal abnormalities. Am J Hum Genet 73(6):1261-70.

16 Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, et al. (2004) Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. J Med Genet 41(4):241-8.

17 Schaeffer AJ, Chung J, Heretis K, Wong A, Ledbetter DH, et al. (2004) Comparative genomic hybridization-array analysis enhances the detection of aneuploidies and submicroscopic imbalances in spontaneous miscarriages. Am J Hum Genet 74(6):1168-74.

18 Le Caignec C, Boceno M, Saugier-Veber P, Jacquemont S, Joubert M, et al. (2005) Detection of genomic imbalances by array based comparative genomic hybridisation in fetuses with multiple malformations. J Med Genet 42(2):121-8.

19 Vialard F, Molina Gomes D, Leroy B, Quarello E, Escalona A, et al. (2009) Array comparative genomic hybridization in prenatal diagnosis: another experience. Fetal Diagn Ther 25(2):277-84.

20 Tyreman M, Abbott KM, Willatt LR, Nash R, Lees C, et al. (2009) High resolution array analysis: diagnosing pregnancies with abnormal ultrasound findings. J Med Genet 46(8):531-41.

21 Armengol L, Nevado J, Serra-Juhé C, Plaja A, Mediano C, et al. (2011) Clinical utility of chromosomal microarray analysis in invasive prenatal diagnosis. Hum Genet 131(3):513-523.

22 Momma K (2010) Cardiovascular anomalies associated with chromosome 22q11.2 deletion syndrome. Am J Cardiol 105(11):1617-24.

23 Webber SA, Hatchwell E, Barber JC, Daubeney PE, Crolla JA, et al. (1996) Importance of microdeletions of chromosomal region 22q11 as a cause of selected malformations of the ventricular outflow tracts and aortic arch: a three-year prospective study. J Pediatr 129(1):26-32.

24 Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, et al. (2008) Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat Genet 40(12):1466-71.

25 Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, et al. (2009) De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. Nat Genet 41(8):931-5.

26 Pietrzyk JJ. (1997) A search for environmental and genetic background for neural tube defects: twenty-five years of experience. Cent Eur J Public Health. 5(2):86-9.

27 Nora JJ, Nora AH. (1976) Recurrence risks in children having one parent with a congenital heart disease. Circulation 53:701–2.

28 Hardin J, Carmichael SL, Selvin S, Lammer EJ, Shaw GM. (2009) Increased prevalence of cardiovascular defects among 56,709 California twin pairs. Am J Med Genet A. 149A(5):877-86.

29 Butler TL, Esposito G, Blue GM, Cole AD, Costa MW, et al. (2010) GATA4 mutations in 357 unrelated patients with

congenital heart malformation. Genet Test Mol Biomarkers 14(6):797-802.

30 Sperling S, Grimm CH, Dunkel I, Mebus S, Sperling HP, et al. (2005) Identification and functional analysis of CITED2 mutations in patients with congenital heart defects. Hum Mutat 26(6):575-82.

31 Erdogan F, Larsen LA, Zhang L, Tümer Z, Tommerup N, et al. (2008) High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. J Med Genet 45(11):704-9.

32 Cusco I, del Campo M, Vilardell M, González E, Gener B, et al. (2008) Array-CGH in patients with Kabuki-like phenotype: identification of two patients with complex rearrangements including 2q37 deletions and no other recurrent aberration. BMC Med Genet 9:27.

33 Rodriguez-Santiago B, Brunet A, Sobrino B, Serra-Juhé C, Flores R, et al. (2010) Association of common copy number variants at the glutathione S-transferase genes and rare novel genomic changes with schizophrenia. Mol Psychiatry 15(10):1023-33.

34 Cooper GM, Coe1 BP, Girirajan S, Rosenfeld JA, Vu TH, et al. (2011) A copy number variation morbidity map of developmental delay. Nat Genet 43(9):838-46.

35 Rodriguez-Santiago B, Malats N, Rothman N, Armengol L, Garcia-Closas M, et al. (2010) Mosaic uniparental disomies and aneuploidies as large structural variants of the human genome. Am J Hum Genet 87(1):129-38.

36 González JR, Rodríguez-Santiago B, Cáceres A, Pique-Regi R, Rothman N, et al. (2011) A fast and accurate method to detect allelic genomic imbalances underlying mosaic rearrangements using SNP array data. BMC Bioinformatics 12:166.

37 Kamburov A, Wierling C, Lehrach H, Herwig R (2009) ConsensusPathDB--a database for integrating human functional interaction networks. Nucleic Acids Res 37:623-628.

38 Stankiewicz P, Sen P, Bhatt SS, Storer M, Xia Z, et al. (2009) Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. Am J Hum Genet 84(6):780-91.

39 van Bon BW, Mefford HC, Menten B, Koolen DA, Sharp AJ, et al. (2009) Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from non-pathogenic to a severe outcome. J Med Genet 46(8):511-23.

40 Lavallee G, Andelfinger G, Nadeau M, Lefebvre C, Nemer G, et al. (2006) The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. Embo J 25(21):5201-13.

41 Nagamani SC, Erez A, Bader P, Lalani SR, Scott DA, et al. (2011) Phenotypic manifestations of copy number variation in chromosome 16p13.11. Eur J Hum Genet 19(3):280-6.

42 Valduga M, Philippe C, Bach Segura P, Thiebaugeorges O, Miton A, et al. (2010) A retrospective study by oligonucleotide array-CGH analysis in 50 fetuses with multiple malformations. Prenat Diagn 30(4):333-41.

43 Ohtoshi A, Behringer RR (2004) Neonatal lethality, dwarfism, and abnormal brain development in Dmbx1 mutant mice. Molec Cell Biol 24(17):7548-58.

44 Chen LT, Gliman AG, Kozasa T (1999) A candidate target for G protein action in brain. J Biol Chem 274(38):26931-8.

45 Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P et al. (2010) A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. Nat Genet 42(3):203-9.

46 Hinton RB Jr, Martin LJ, Tabangin ME, Mazwi ML, Cripe LH, et al. (2007) Hypoplastic left heart syndrome is heritable. J Am Coll Cardiol 50(16):1590-5.

47 Iascone M, Ciccone R, Galletti L, Marchetti D, Seddio F, et al. (2011) Identification of de novo mutations and rare variants in hypoplastic left heart syndrome. Clin Genet.

48 Ashleigh R, Payne AR, Chang SW, Koenig SN, Zinn AR, et al. (2012) Submicroscopic Chromosomal Copy Number Variations Identified in Children With Hypoplastic Left Heart Syndrome. Pediatr Cardiol.

49 Bentham J, Bhattacharya S (2008) Genetic mechanisms controlling cardiovascular development. Ann N Y Acad Sci 1123:10-9.

50 Bittel DC, Butler MG, Kibiryeva N, Marshall JA, Chen J, et al. (2011) Gene expression in cardiac tissues from infants with idiopathic conotruncal defects. BMC Med Genomics 4(1):1.

SUPPLEMENTARY MATERIAL

Supplementary table 1
List of heart malformations present in the cohort of 33 studied fetuses with isolated congenital heart defect.

Supplementary table 2
Overview of the central nervous malformations in 26 of the analyzed fetuses.

Supplementary table 3
Type of renal malformations observed in 9 of the studied fetuses.

Supplementary table 4
Overview of the affected organs and systems in fetuses with multiple malformations.

Supplementary table 5
MLPA probes used to discard well-known genetic alterations related to MCA / MR.

Supplementary table 6
MLPA probes used to validate the alterations detected by CMA and to study parental samples.

Supplementary table 7
Summary of rare copy number variations >100kb detected in samples of 168 control subjects.

Sample	Tissue	Gender	Malformation
1	lung	female	RHH with tricuspid valve and pulmonary artery stenosis, hypoplastic pulmonary artery and branches, membranous IVC
2	heart	female	LHH
3	heart	female	LHH
4	heart	male	LHH
5	heart	female	tricuspid valve stenosis, RHH, VSD type ostium primum, muscular VSD, mitral valve dysplasia, aneurismal dilatation of posterior left atrium
6	liver	male	LHH
7	heart	male	LHH
8	heart	male	double outlet right ventricle, D-TGA, dual left anterior descending artery, hypoplastic pulmonary artery, overriding aorta over membranous IVC
9	heart	female	L-TGA, incomplete AVSD, LHH, overriding aorta over membranous IVC
10	heart	male	double outlet right ventricle with D-TGA, extreme hypoplasia of the aortic arch, IVC, moderate hypoplasia of the supraaortic trunk
11	heart	male	double outlet right ventricle with IVC, hypoplasia of the ascending aorta
12	liver	female	L-TGA with atrioventricular discordance, subpulmonar IVC
13	liver	male	dextrocardia, D-TGA, double outlet right ventricle with IVC
14	liver	female	truncus arteriosus type II
15	heart	male	LHH
16	heart	female	tetralogy of Fallot
17	heart	male	atrioventricular discordance, right ventricle with left ventricle morphology and viceversa
18	heart	female	LHH
19	heart	male	D-TGA, muscular VSD, hypoplastic ductus arteriosus, origin of both coronary arteries from the anterior leaflet of the aortic valve, bicuspid pulmonary valve
20	liver	male	D-TGA, IAC type ostium secundum, LHH with mitral and pulmonary valves atresia, hypoplastic pulmonary artery, permeable ductus
21	heart	male	large foramen ovale, tricuspid valve atresia, VSD, TGA, pulmonary atresia, hypoplastic pulmonary artery
22	lung	female	LHH with mitral and aortic valves atresia
23	heart	female	LHH with mitral and aortic valves atresia and hypoplasia of the ascending aorta
24	heart	male	membranous IVC
25	heart	male	LHH with mitral and aortic valves hypoplasia
26	heart	female	double outlet right ventricle, large foramen ovale, VSD, coarctation of the preductal aorta
27	heart	male	hypoplasia of the ascending aorta and muscular IVC
28	heart	female	LHH with mitral and aortic valves atresia, incarceration of the left ventricle and extreme hypoplasia of the ascending aorta
29	heart	female	LHH
30	lung	female	extreme LHH with mitral and aortic valves atresia and hypoplasia of the ascending aorta
31	heart	male	LHH with mitral and aortic valves atresia, aortic valve with two leaflets and tubular hypoplasia of the aortic arch
32	heart	female	truncus arteriosus type II, absent ductus arteriosus, truncal valve with three leaflets, VSD
33	heart	male	complex cardiopathy

Supplementary table 1. List of heart malformations present in the cohort of 33 studied fetuses with isolated congenital heart defect. RHH: right heart hypoplasia; IVC: interventricular communication; LHH: left heart hypoplasia; VSD: ventricular septal defect; D-TGA: dextro-transposition of the great arteries; L-TGA: levo-transposition of the great arteries; AVSD: atrioventricular septal defect.

Sample	Tissue	Gender	Malformation
34	lung	female	Holoprosencephaly
35	CNS	female	Sacral neural tube defect, Arnold-Chiari malformation
36	CNS	male	Lumbosacral neural tube defect, Arnold-Chiari malformation
37	CNS	female	Lumbosacral neural tube defect, Arnold-Chiari malformation
38	liver	female	Anencephaly
39	CNS	male	Sacral myelomeningocele, Arnold-Chiari malformation
40	liver	male	Holoprosencephaly
41	liver	female	Occipital encephalocele, Arnold-Chiari malformation
42	liver	female	Biventricular hydrocephalus
43	liver	male	Lumbosacral myelomeningocele, Arnold-Chiari malformation
44	CNS	male	Bilateral ventriculomegaly
45	CNS	female	Lumbosacral myelomeningocele, Arnold-Chiari malformation
46	CNS	male	Lumbar myelomeningocele, Arnold-Chiari malformation
47	CNS	male	Biventricular hydrocephalus
48	liver	male	Lumbosacral myelomeningocele, Arnold-Chiari malformation, brachycephaly
49	liver	female	Agenesis of the corpus callosum
50	liver	female	Biventricular hydrocephalus
51	CNS	female	Lumbosacral myelomeningocele, Arnold-Chiari malformation
52	liver	male	Sacral neural tube defect, Arnold-Chiari malformation
53	heart	female	Bilateral ventriculomegaly
54	CNS	female	Lumbosacral myelomeningocele, Arnold-Chiari malformation
55	CNS	female	Lumbosacral neural tube defect, Arnold-Chiari malformation
56	CNS	female	Lumbar myelomeningocele, Arnold-Chiari malformation
57	CNS	female	Holoprosencephaly
58	CNS	female	Bilateral ventriculomegaly
59	amniotic fluid	male	Neural tube defect

Supplementary table 2. Overview of the central nervous malformations in 26 of the analyzed fetuses. CNS: central nervous system.

Sample	Tissue	Gender	Malformation
60	adrenal gland	male	Bilateral renal agenesis, bilateral agenesis of the ureters. Oligohydramnios sequence
61	lung	male	Bilateral multicystic dysplastic kidney disease with ureter dysplasia and bladder hypoplasia
62	adrenal gland	female	Bilateral renal agenesis, bilateral agenesis of the ureters. Oligohydramnios sequence
63	liver	male	Bilateral renal agenesis, bilateral agenesis of the ureters. Oligohydramnios sequence
64	lung	male	Bilateral multicystic dysplastic kidney disease with ureter dysplasia and bladder hypoplasia
65	adrenal gland	male	Congenital nephronophthisis
66	liver	male	Bilateral renal agenesis, bilateral agenesis of the ureters. Oligohydramnios sequence
67	kidney	male	Bilateral obstructive renal dysplasia
68	adrenal gland	male	Bilateral renal agenesis, bilateral agenesis of the ureters. Oligohydramnios sequence

Supplementary table 3. Type of renal malformations observed in 9 of the studied fetuses.

Sample	Tissue	Gender	CHD	CNS	IUGR	Craniofacial	Skeletal	Urogenital	Others
69	heart	male	X						Esophageal atresia, anal imperforation (VATER)
70	heart	female	X						Asplenia
71	liver	female		X			X	X	Craniosynostosis
72	liver	male				X	X		Lymphangioma
73	liver	female	X					X	
74	spleen	female	X	X					Omphalocele
75	liver	male	X					X	
76	liver	female		X		X	X		
77	liver	male	X	X	X	X	X	X	
78	liver	male	X					X	
79	liver	male	X						Asplenia
80	liver	female	X	X		X			
81	liver	male			X	X		X	
82	lung	male				X	X	X	
83	liver	female		X		X			Hairy polyp, bilateral pieloureteral duplication
84	lung	male	X					X	
85	liver	female	X						Accessory spleen, pulmonary hypoplasia
86	liver	female		X				X	Accessory spleen, malformation of the adrenal gland
87	liver	female		X				X	
88	heart	male	X				X		
89	liver	female							
90	liver	male	X			X			Two accessory spleen, bilateral adrenal neuroblastoma
91	liver	female					X	X	OEIS complex
92	adrenal gland	male						X	Anal imperforation, rectovesical fistula
93	liver	female	X	X	X	X	X	X	
94	liver	female	X			X		X	Diaphragmatic hernia
95	liver	female	X					X	Omphalocele, anal imperforation

Supplementary table 4. Overview of the affected organs and systems in fetuses with multiple malformations. CHD: congenital heart defect; CNS: central nervous system; IUGR: intrauterine growth restriction.

Gene	Chr	Start	End	Syndrome related	Hybridization sequence
WBSCR1	7	73604582	73604629	Williams-Beuren sd	GGTAGCTCTCGAGAATCTAGAGGTGGATGGGATTCCCGGGATGACTTC
HIRA	22	19318971	19319021	DiGeorge sd	CCCTCAGGATGTCGAGCTGTTCCTGACACTCGGTGAAGAGGCGCTGGAATC
NSD1	5	176700712	176700762	Sotos sd	CCCAAAAAGAGCTAAGACAGCTGCAGGAAGACCGAAAGAATGACAAGAAGC
SNRNP	15	25200680	25200734	Prader-Willi sd	AGGGGGTGTTGAGCGCAGGTAGGTGTATAATAGTGACCACTGCGTGGTGGAGCAG
ARIH1	15	72855773	72855832	15q24 del sd	CCACCATGTTGTTAAAGTCCAATATCCTGATGCTAAACCTGTTCGCTGCAAATGTGGGCG
TRIP3	17	34851076	34851139	Mental retardation	CAACATTAAGAAGCTTATTGCTCAATCCACACCTCAGGCAGTTGATGGTCAACCTCGATCAGGG
PML	15	74290611	74290671	15q24 del sd	CAAAGAGTCGGCCGACTTCTGGTGCTTTGAGTGCGAGCAGCTCCTCTGCGCCAAGTGCTTC
COPS3	17	17179360	17179431	Smith-Magenis sd	CTTGCGAAGAACTTATCCCATCTGGACACTGTGCTCGGGGCTCTGGATGTACAAGAACACTCCTTGGGCGTC
BAZ1B	7	72925125	72925198	Williams-Beuren sd	AGTACTCTTGCACGTCCAAATGCGCTCACTGTACCTTTCCAAGCGGGCTTCATACTCTCTGTTGGCAGTAGTTC
MAPT	17	44102520	44102589	17q21.31 del sd	CTGGGACTTTAGGGCTAACCAGTTCTCTTTGTAAGGACTTGTGCCTCTTGGGAGACGTCCACCCGTTTC
BCL9	1	147083562	147083622	1q21.1 sd	CTGTTTCTGCTGCAACCCGAGAGGAACTCGGTGAGCCTGTCCCGTTTGTGACTGCAAGCTC
CDH9	5	26906872	26906944	Control probe	CATCTGTTATACAAGTAACTGCAACAGATGCAGATGACGCCAACTATGGAAATAGTGCCAAAGTGGTCTATAG
KIAA0427	18	46383961	46384012	Control probe	CTTGCAATCTCAGGATGTGAAGGAAGATGCTGTCCTTTGCTGCTCTATGGAG
GNB1L	22	19789603	19789662	DiGeorge sd	GAAGGTGTGCAGCCGCATCGCCTGCCATGAGGAGCCCGTCATGGACCTTGACTTTGACTC
FLJ20436	12	49065580	49065646	Control probe	GCCTAGAGCTTCATGTTCCACTTTGCGATTATGTAAGTATCGGCGCTTCTTCTCCTTGAGCAGATGC
SKI	1	2161109	2161169	1p36 syndrome	CGCTGCCTGGACGACGTGAAGGAGAAATTCGACTATGGCAACAAGTACAAGCGGCGGGTGC
TP73	1	3624031	3624090	1p36 syndrome	CCTGTAACAGGACACCTCCTAGACGGGACAGGACGACTGACTGTGTGTGTTTCCCCCTCC

Supplementary table 5. MLPA probes used to discard well-known genetic alterations related to MCA / MR. Hg19 assembly.

Gene	Chr	Start	End	Hybridization sequence
KLF13	15	31666123	31666182	GGTATCTTGCCAGAGATACCTGTTTTGATGAGTACCTATTTTGATGCAAAGGAACGGCCC
FOXC2	16	86602268	86602335	CAGCAGCAAACTTTCCCCAACGTGCGGGAGATGTTCAACTCCCACCGGCTGGGGATTGAGAACTCGAC
OSGIN	16	83982733	83982798	GAAAGTCCACGTCTCGCCCCACAGTTCCCCATATCAGCAAATCTTCAAATTCTAGAACAAGGTCAG
DIAPH3	13	60584710	60584762	CTTATTGGCCAAAGCCGTGGATCCCAGACACCCCAATATGATGACAGATGTGG
USP46	4	53468131	53468190	GGCCCTGCACCTAAAGCGGTTCAAGTACATGGAGCAGCTGCACAGATACACCAAGCTGTC
GPLD1	6	24447129	24447190	GCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCCCTTCACGG
PIK3C2G	12	18499652	18499715	GCAGATTTTCAGCCTGTAAATGTACCTAGATGCACTTCCTATCTAAATCCCGGGCTTCCTTCCC
CDH19	18	64212057	64212119	CTGAATCTGCACCCACTGGGACTTCTATAGGAACAATCATGGCATATGATAATGACATAGGAG
CNTN4	3	3081837	3081894	CGTGTTCAGGAATGAGAGCGTGCACCCCTTCTCTCCCTTTGAGGTTAAAGTAGGTGTC

Supplementary table 6. MLPA probes used to validate the alterations detected by CMA and to study parental samples. Hg19 assembly.

Sample	Gain/Loss	Region	Length (kb)	Start	End	Genes in the region	Control frequency (10,320)
1101	Gain	2p16.1	115.779	56606229	56722007	*CCDC85A*	0
1007	Gain	5q35.1	289.45	169441749	169731198	*DOCK2,FOXI1,LCP2,LOC133874*	0
101	Loss	6q22.33	171.412	127983009	128154420	*THEMIS*	0
1091	Gain	10p15.1	122.207	5016797	5139003	*AKR1C1, AKR1C2, AKR1C3, LOC100134257*	0
1087	Loss	4q12	312.257	53906416	54218672	*SCFD2*	0
	Gain	3p25.3	164.621	9900524	10065144	*CIDEC, CIDECP, CRELD1, IL17RC, IL17RE, JAGN1, LOC401052, PRRT3, TMEM111*	0
119	Gain	12p11.21	194.756	32871481	33066236	*DNM1L, PKP2, YARS2*	0
1024	Loss	9p22.1	680.069	18208852	18888920	*ADAMTSL1*	0
1030	Gain	10q23.33	134.131	96626682	96760812	*CYP2C9*	0
1032	Gain	8p23.1	428.039	8215349	8643387	*CLDN23, MFHAS1, PRAGMIN*	0
14	Gain	4p15.32	736.736	16842680	17579415	*CLRN2, LAP3, LDB2, QDPR*	1
1041	Gain	16p12.3	268.802	18792147	19060948	*ARL6IP1, RPS15A, SMG1, TMC7*	1
1010	Gain	22q11.23	538.044	23111199	23649242	*BCR, GNAZ, RAB36, RTDR1*	1
136	Gain	18q12.2	674.203	34279819	34954021	*BRUNOL4, C18orf10, FHOD3, KIAA1328*	1
1064	Gain	6q15	326.983	90590332	90917314	*BACH2, GJA10*	1
1006	Gain	13q12.11	144.699	20066736	20211434	*MPHOSPH8,TPTE2*	1
1082	Gain	4q31.3	402.135	151668146	152070280	*LRBA, RPS3A, SH3D19, SNORD73A*	1
1115	Gain	3q13.12	106.25	107441005	107547254	*BBX*	1
113	Gain	5q35.2	317.302	175467619	175784920	*C5orf25, FAM153B, KIAA1191*	1
1029	Gain	12q24.33	454.025	132527235	133037139	*DDX51, EP400, EP400NL, GALNT9, LOC100130238, NOC4L*	1
1014	Gain	5q11.2	504.677	55415101	55919777	*ARL15*	1
1015	Loss	13q12.13	200.368	26174679	26375046	*ATP8A2*	1
1106	Gain	2p21	564.149	45409988	45974136	*PRKCE,SRBD1*	2
1020	Gain	10q26.3	192.786	134568387	134761172	*C10orf92, C10orf93, INPP5A, NKX6-2*	2
152	Gain	5q22.3	437.6	115071808	115509419	*AP3S1, ATG12, CDO1, COMMD10, LVRN*	2
1099	Loss	11q21	110.245	95933123	96043367	*MAML2*	4
1045	Gain	7p15.1	327.356	29209352	29536707	*CHN2*	4
1061	Gain	8q21.3	156.332	87177637	87333968	*SLC7A13*	4
1080	Gain	1p22.1	351.535	92273644	92625178	*BRDT, BTBD8, EPHX4, TGFBR3*	4
1034	Gain	16q23.1	458.538	76453043	76911580	*CNTNAP4*	5
1049	Loss	17q24.3	141.176	67169500	67310675	*ABCA10, ABCA5*	5
	Loss	15q25.3	254.344	87870024	88124367	*NCRNA00052*	0

Supplementary table 7. Summary of rare copy number variations >100kb detected in samples of 168 control subjects. Hg19 assembly.

CHAPTER 2

Clinical utility of chromosomal microarray analysis in invasive prenatal diagnosis

Armengol L, Nevado J, Serra-Juhé C, Plaja A, Mediano C, García-Santiago FA, García-Aragonés M, Villa O, Mansilla E, Preciado C, Fernández L, Ángeles Mori M, García-Pérez L, Lapunzina PD, Pérez-Jurado LA.

Hum Genet. 2012 Mar;131(3):513-23

CMA has become an essential tool in genetic research as well as in diagnostics. Nevertheless, the implementation of new molecular techniques in some settings, like prenatal setting, requires specific and large studies to prove its efficiency and to deal with the difficulties. With the aim of exploring the clinical utility of CMA in prenatal setting and to compare its detection rate, sensitivity, specificity and VOUS detection with gold standard techniques, a multicentric study was designed.

900 pregnant women who had to undergo invasive prenatal sampling were offered to participate in this study. MLPA, standard karyotype, targeted CMA and QF-PCR were done in all cases and the detection rate, turn-around-time and costs were considered in order to compare different techniques.

Armengol L, Nevado J, Serra-Juhé C, Plaja A, Mediano C, García-Santiago FA, García-Aragonés M, Villa O, Mansilla E, Preciado C, Fernández L, Ángeles Mori M, García-Pérez L, Lapunzina PD, Pérez-Jurado LA. Clinical utility of chromosomal microarray analysis in invasive prenatal diagnosis. Hum Genet. 2012 Mar;131(3):513-23

CHAPTER 3

Genetic counseling with high throughput prenatal screening methodologies: identification of relevant factors in decision making

C.Serra-Juhé, L.Fernández García-Moya, I.Valenzuela, C.Giménez, N.Bosch, F.García-Santiago, E.Mansilla, N.Capdevila, M.Cendrós, M.Del Campo, J.Nevado, L.Armengol, P.Lapunzina, L.A.Pérez-Jurado

Under review in Clinical Genetics

Taking into consideration that the ultimate goal of research in genetic diseases is the improvement in patients' care, it is important to establish genetic counseling strategies at the same time that techniques and technologies are developed. The relevance of genetic counseling in prenatal screening and prenatal invasive testing has been proved in previous published studies. The lack of appropriate information has been identified as one of the factors related to the rejection of prenatal tests due to the perception of having not enough information to decide. It has been also reported that various demographic and social factors such as the educational level, the reproductive history or the religiousness influence in decision-making. Some of those factors also influence the comprehension and assimilation of the information received by pregnant women. Programs and strategies should be designed to minimize the current difficulties and ensure that women receive and comprehend the information about the situation.

During the studies described in the chapters of this thesis, genetic counseling was provided in relation to CMA and exome sequencing. In addition, a study to deeply analyze the acceptability and the most appropriate manner to provide genetic counseling in relation to those techniques was designed. CMA was the elected technique to perform the study considering that, nowadays, is the only one among the techniques used that has been broadly implemented in clinical setting. CMA shares several characteristics with other techniques taking into account the difficulty to counsel regarding variants of unknown significance. For this reason, the results of this study could also be considered in the context of other molecular techniques.

Genetic counseling with high throughput prenatal screening methodologies: identification of relevant factors in decision making

C.Serra-Juhé, L.Fernández García-Moya, I.Valenzuela, C.Giménez, N.Bosch, F.García-Santiago, E.Mansilla, N.Capdevila, M.Cendrós, M.Del Campo, J.Nevado, L.Armengol, P.Lapunzina, L.A.Pérez-Jurado

ABSTRACT

Pregnant women with an indication for prenatal invasive sampling due to risk for fetal anomalies, visited in Vall d'Hebron (Barcelona) and La Paz (Madrid) hospitals during 2009, were offered to participate in a study of the feasibility and cost-efficiency of novel methods for prenatal screening of chromosomal abnormalities and the factors influencing acceptance or rejection of these tests. We provided pre-test genetic counseling, including information on risks, expected detection rates and potential findings. A questionnaire about social and demographic characteristics and reasons to accept or refuse advanced prenatal testing was obtained from 402 women. 94% agreed to extend prenatal studies and 91.3% considered they had enough information to take decisions, provided by pre-test counseling. The average anxiety level was intermediate-high, related to the indication for testing and previous reproductive history. The academic level of progenitors influenced the reasons to participate and their understanding of the benefits and limitations of these techniques. Refusal of the study was related to anxiety, birth place and the indication for the invasive sampling. Our data show a high level of acceptance of novel high throughput molecular techniques in prenatal setting reinforcing the relevance of pre-test genetic counseling to minimize anxiety and help in decision making.

INTRODUCTION

The average risk in the general population of having a child with birth defects is ~2-3%. (1) Some of those birth defects and/or their cause can be detected during pregnancy, which is highly relevant to monitor gestation, determine fetal viability, and anticipate medical problems at birth or later. The identification of the causal genetic abnormality of birth defects is also of utmost importance for genetic counseling to the couple and other family members.

Currently, prenatal ultrasound and maternal blood screening in the first or second trimester of the pregnancy are usually offered to pregnant women in most countries, being an almost universal practice in Spain. Depending on the predicted risk derived from these screenings, maternal age and family history, invasive sampling of fetal cells through amniocentesis or chorionic villous biopsy may be offered for prenatal genetic diagnosis. Karyotyping and quantitative fluorescent PCR (QF-PCR) are commonly performed in all samples.

Conventional karyotyping can detect all aneuploidies and other chromosomal abnormalities such as translocations – balanced or unbalanced -, inversions and deletions or duplications with a resolution of ~5-10 Mb (2-4). Due to the requirement of cell culture, karyotype results usually take 2 or 3 weeks. The most common aneuploidies (13, 18, 21, X and Y), accounting >80% of the diagnosable clinically relevant

chromosomal abnormalities in the prenatal period, can also be detected by QF-PCR or interphase FISH (5-8) and the results are obtained in 2 days. In recent years, higher resolution molecular techniques such as Multiplex Ligation-dependent Probe Amplification (MLPA) or Chromosomal Microarray Analyses (CMA) have been developed with the aim of improving the detection rate of genetic alterations (9-15), which can be increased with next generation sequencing of part or the whole genome (16). Despite potential improvement in detection rate, higher resolution genomic analyses may also increase the detection of variants of uncertain clinical significance (VOUS) with doubtful or unclear clinical impact (17, 18). The detection of VOUS in prenatal diagnosis can be reduced, but not eliminated, by targeted designs to analyze only genomic regions whose alterations cause a well-established phenotype (19).

Several studies have proved the need for genetic counseling in prenatal screening and diagnosis to facilitate well informed decision-making without increasing maternal anxiety (20). In general, there is a lack of information in pregnant women who had decided to undergo a prenatal screening or diagnostic test in relation to relevant aspects of the process, such as the limitations of the tests used (20). This lack of appropriate information has been related to screening rejection because of the perception of not having enough information to decide. Skepticism about the usefulness of the test and misunderstanding of its purpose are among the main reasons for rejecting prenatal screening tests (21).

The emerging molecular techniques described above involve a challenge in genetic counseling considering that they extend the diagnostic capability and the potential impact of VOUS (22). As in all genetic counseling processes, the pregnant woman should have the choice to decide whether she is interested in extending prenatal diagnostic tests to make a decision about the continuation of the pregnancy (23, 24). The counselor should attempt to provide, in a non-directive manner, the necessary information to the woman to enable her to make a real informed choice (24-26).

Communication skills of the counselors in prenatal setting, including the manner in which the advice is provided (27) or the vocabulary used (20, 28), are important for risk perception of pregnant women. Demographic factors may also influence decision-making. The educational level, the country of origin, the reproductive history, the marital status, the personal experience with people suffering a genetic disease and the opinion regarding the interruption of pregnancy are other factors that have been identified as relevant in the decision process (20, 21, 29-31). Religion and also religiousness have been described as two of the most relevant factors influencing prenatal choices (21, 29, 30, 32). It is important to identify all these aspects in order to design programs and strategies to minimize the current drawbacks and ensure that women receive and comprehend the information and choices. In this manuscript we have investigated the acceptance or rejection of novel high-throughput genetic tests for prenatal screening among pregnant women and the factors influencing their decision.

MATERIAL AND METHODS

In January 2009 we launched a study in the Hospitals Vall d'Hebron in Barcelona and La Paz in Madrid, Spain, to define the clinical utility and cost-efficiency of CMA and MLPA in invasive prenatal diagnosis in comparison with the currently available techniques, QF-PCR and karyotype (19). Pregnant women who underwent amniocentesis or chorionic villus sampling in those hospitals during 2009 were offered to participate. We also explored the challenge that new technologies represent for genetic counseling and the psychosocial and demographic aspects related to the acceptance or rejection of novel techniques.

The study included a pre-test genetic counseling and informative session, which involved the signing of the informed consent, along with a survey about psychosocial and demographic issues. A post-test genetic counseling session was provided when the result of some of the tests was positive.

Pre-test genetic counseling

This genetic counseling session included the explanation and discussion with the pregnant women or couple, in an adequate level of understanding, of the most relevant aspects related to the study: 1) the nature of the four techniques to be performed (karyotype, MLPA, QF-PCR and CMA) and possible outcomes; 2) the limitations of the tests performed, the possible finding of VOUS with requirement of parental samples, and the time required to obtain the results; 3) the detection of genetic anomalies with clinical impact would be informed by a post-test genetic counseling session.

Relevant factors in decision making

At the end of the pre-test session, pregnant women had some time to decide before accepting or rejecting their participation in the study. Regardless of their decision, they were offered to answer a short survey to explore reasons related to their decision, including psychosocial and demographic aspects. The survey (see Supplemental material) was based on the already known relevant factors for decision-making in prenatal setting (20, 21, 29-31). The reasons for pregnant women to accept or reject the study were collected in the questionnaire as an open question.

Moreover, considering the experience gained during the counseling of all women in relation to new methodologies in prenatal screening, the main aspects for pre-test and post-test proper genetic counseling were identified.

Statistical analyses.

To explore the correlation between the level of anxiety and demographic characteristics and other parameters considered in the survey, the Mann-Whitney U and Kruskal-Wallis tests were used. To analyze whether demographic characteristics were significantly different between women who accepted to extend prenatal studies and the ones who rejected it, Fisher exact test was used. The same test was also used to explore the correlation between the reasons to extend prenatal studies and the level of education.

Post-test genetic counseling

When a genetic alteration with known clinical impact was detected, the mother/couple was informed of the anomaly by the counselor or doctor. During the post-test session, the genetic alteration, its possible clinical significance and the recurrence risk in future pregnancies were all explained. Options related to the pregnancy in progress were also discussed. When the identified alteration had unknown or unclear clinical impact, the couple was informed about the uncertainty while asking for parental samples to prove the inherited or de novo nature of the genetic alteration.

RESULTS

Demographic characteristics

A total of 402 pregnant women agreed to participate in this study answering the survey. Mean age was 35.6+/-4.8, being 44.5% of them between 36 and 40 years old. Among them, 94% decided to extend prenatal studies. The great majority, 91.3%, considered to have enough information for decision-making, while only 8.7% thought their information was not appropriate.

Most participants were Spanish (84%) with Southern Americans accounting for 11%; 61% were married and an additional 16% were non-married co-habiting couples. Regarding educational background, 52% had university studies, 36% high school studies and the remaining 12% middle school studies. Slightly more than 2/3 (69%) of women considered themselves religious

persons, while 31% were atheistic or agnostic. In table 1, there is an overview of the social and demographic characteristics of participants in the survey, including information about their reproductive history and the indication for invasive prenatal sampling in the current pregnancy.

Characteristics	Percentage of participants
Age	
<20years	1%
20-25 years	3%
26-30 years	11%
31-35 years	29%
36-40 years	44%
41-45 years	12%
Education level	
Middle school	12%
High school	36%
University	52%
Marital status	
Single	19%
Separated	1%
Widowed	1%
Divorced	2%
Married	61%
With partner	16%
Religious beliefs	
Atheistic	17%
Agnostic	14%
Christian	63%
Muslim	1%
Other	5%
Reproductive history	
Prior pregnancy	71%
Prior live births	53%
Prior miscarriage	27%
Prior termination of pregnancy	18%
Previous invasive test	22%
Assisted reproductive technologies	6%
Indication for invasive prenatal sampling	
Ultrasound finding	11%
Altered biochemical screening	26%
Maternal age	46%
Family history	17%
Anxiety level	
Very low	3%
Low	18%
Intermediate	42%
High	28%
Very high	9%

Table 1. Social and demographic characteristics of participants in the survey.

Reasons to extend prenatal studies

The reasons why pregnant women decided to accept the extension of prenatal tests were asked as open question. The answers were classified within one or more of the following groups: 1) to have more information (44.6%); 2) to contribute to scientific progress (48.1%); 3) to be less worried / to be sure that everything is correct (5.1%); and 4) to be grateful for the attention received (2.2%). As expected, the most relevant issues for pregnant women related to any novel genetic tests were the expected sensitivity and specificity of the potential findings in terms of accurate prognosis for the fetus, the turnaround timing for delivery of results and whether additional testing might be required. Avoiding repetition of invasive sampling was a priority.

We explored a possible correlation between the level of academic education and the reasons to participate in the study. Contribution to scientific progress was most frequent answer in the group of women with university studies (main reason in 69.3% of women of this group), and less frequent in the other groups (38.5% in women with high school studies and 28.2% in women with middle school studies, $p < 0.001$) . Likewise,

the answer "to be sure that everything is correct", was more frequent in women with middle school studies (10.3%) compared with women with high school studies (4.4%) or university studies (2%) but the differences in frequencies were not statistically significant.

Anxiety level

Several comparisons were done with the aim to identify factors that may influence maternal anxiety in prenatal setting. The average anxiety level was 3.19 in a 0-5 scale, with 42.3% of women considering their anxiety level as medium, 36.4% high or very high and 21.3% low or very low.

Anxiety was slightly higher, but not statistically significant, in women who had used assisted reproductive technologies (3.5 versus 3.32 for those who did not) and inversely correlated with the education level (3.33 average in middle or high school studies, 3.11 in the group with university studies, $p < 0.05$). Other factors, such as the reasons to extend prenatal studies or the indication for invasive prenatal sampling also correlated with anxiety level (figure 1 and table 2).

Reasons to reject additional prenatal studies

A total of 8/128 (6.3%) women, with mean age 33.8 years old, rejected additional prenatal studies. Reasons for rejecting additional studies were only explored in the cohort of 128 women attended in Vall d'Hebron hospital. Among them, there was an increased proportion of South American origin, and lower proportion with university degree (20%). The average anxiety level for this group was higher than the total group, 3.8, and the main reason for rejection was increased anxiety in 50% of them.

Indications for invasive prenatal sampling	**p-value**
Age *vs* altered biochemical screening	0.001
Age *vs* Previous medical history	0.001
Age *vs* Ultrasound finding	0.001
Reasons to extend prenatal studies	p-value
"Contribute to scientific progress" *vs* "have more information"	0.05
"Grateful for the attention received" *vs* "have more information"	0.05

Table 2. Significant correlations in anxiety levels between different socio-demographic characteristics.

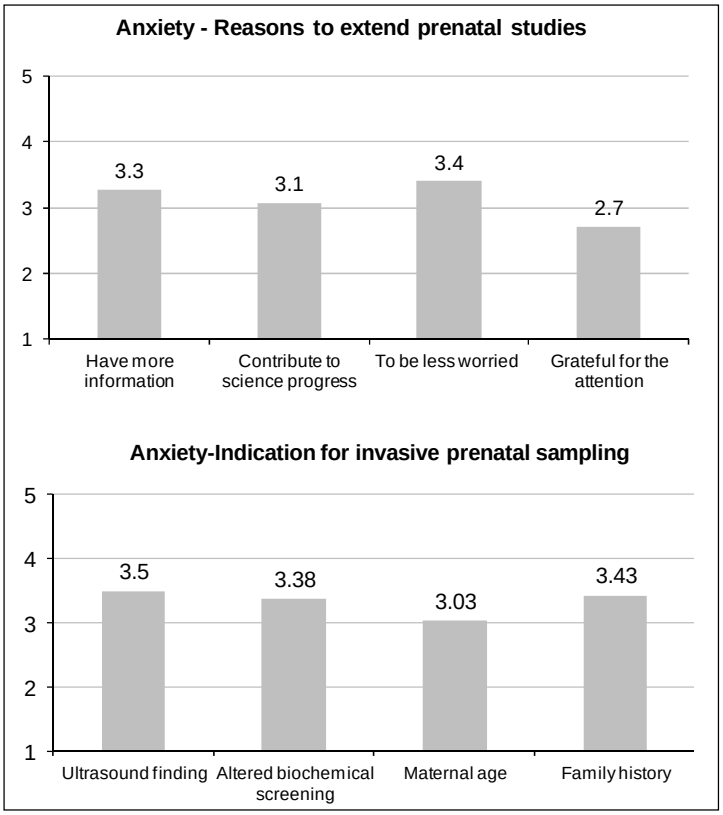


Figure 1. Anxiety level of pregnant women depending on reasons to extend prenatal studies and indication for invasive prenatal sampling

The possible need for invasive resampling was considered a potential cause for rejection while collection of parental samples was not regarded as conflictive.

DISCUSSION

We have shown here that there is a high acceptance rate for the use of novel genomic high throughput technologies in invasive prenatal diagnosis among pregnant women with at risk gestations. Our data reinforce the idea that detailed pre-test genetic counseling is crucial to help decision-making in prenatal setting. The script we used was apparently appropriate for the aim of the session, given that more than 90% of pregnant women considered to have received sufficient information to make a decision. Relevant issues to clearly explain during the session include the actual sensitivity and specificity of the applied tests, potential certainty (or lack of it) in the prognosis of the detectable alterations, and the turnaround timing for result delivery.

Given that the educational level is a key factor for decision-making in prenatal studies, it is important to consider it during genetic counseling sessions, in order to make sure that the explanations are understandable. This becomes even more relevant when a result of uncertain significance has to be reported. Therefore, it is recommendable to promptly assess the level of understanding of the recipients to provide a counseling session according to their level of comprehension. It is also important to allow them to set the pace of the session according to their

understanding and emotional situation, as well as to create a suitable atmosphere for them to feel comfortable expressing their doubts and concerns.

We have identified relevant factors influencing decision-making, some of them previously reported, such as educational level, place of origin and reproductive history (20, 21, 29-31). However, communication failure and misunderstanding of the purpose of the test, common reasons for rejecting prenatal studies, were rare in our study, since 91.3% had enough information to decide with no differences between those who accepted and those who refused additional testing.

The level of education emerged as the most important factor influencing anxiety and global understanding. Anxiety inversely correlated with education level, indicating that a better understanding reduces anxiety despite potentially increased uncertainty. The answers to the questionnaire also showed a greater sensitization of the participants with higher education towards scientific research and progress.

Anxiety of pregnant women was also influenced by the previous reproductive history and the perception of risk depending on the indication for invasive prenatal sampling. Slightly higher anxiety levels were found in women during first pregnancies and/or after assisted reproductive technology. The perception of risk was greater, and therefore also the anxiety, if invasive testing was indicated due to family history of disease or ultrasound findings. In those cases, some information related to the risk for the fetus had already been provided prior to referral.

Despite the small size of the group that rejected additional prenatal studies, there was a clear correlation with the countries of origin. Half of them were of South American origin compared to the 20% proportion of the entire cohort, with lower proportion having a higher education level. These data suggest that the success in making counseling understandable was worse for South American than for Spanish women despite having the same language, likely due to cultural differences. Interestingly, half of the women who refused to extend prenatal test but agreed to answer the survey had an indication of invasive testing due to ultrasound findings. In those cases, there was a likely denial of the possible unfavorable evolution of the pregnancy, considering that complications had already been detected and the additional studies were indicated for validation and prognosis of the findings: pregnant women might try to avoid testing that could confirm a bad prognosis for the gestation.

The level of religiosity and the religion of the mother had also been described as one of the most influential factors in prenatal decision-making with higher religiosity correlating with rejection of invasive prenatal sampling (21, 29-31). In our study, this aspect was not relevant, as the studied population had already considered undergoing invasive prenatal testing.

In summary, there is a high acceptance rate for expanded prenatal diagnosis with novel technologies among pregnant women with at risk gestations, assuming appropriate pretest genetic counseling is provided. We have also identified relevant factors that may condition women's decision and should be taken into account during counseling. Given the proven clinical utility and the predictable increased availability of high throughput prenatal genetic testing, further studies are granted. Studies addressed to evaluate opinion and satisfaction of women after receiving the results of this type of testing, either predictable alterations or VOUS, will significantly complement the current data. Understanding how all this information about detectable genetic alterations in the fetus during the pregnancy is received by the family, how they cope with the situation and what they think about its usefulness is of utmost importance for appropriate counseling and clinical implementation.

CONFLICT OF INTEREST

Lluís Armengol and Luis Pérez-Jurado are executive director and scientific advisor, respectively, of qGenomics, a privately held company that provide genomics services to the scientific and medical community.

AKNOWLEDGMENTS

We would like to thank the support of the medical geneticists and obstetricians in both hospitals, T. Vendrell, A. Cueto, S. García-Miñaúr, M.A. Sánchez, J. Sagalá, M.T. Higueras, E. Carreras, R. Rodríguez, A. González, F. Herrero, as well as all participants. We would also like to thank to Gabriela Palacios for statistical support and critical review. This work was supported by a grant from the Spanish Agency for Evaluation of Novel Sanitary Technologies (AETS-PI08/90056). CS-J was supported by a predoctoral fellowship (FI08/00365) of the Fondo de Investigación Sanitaria, respectively.

REFERENCES

1 Kalter H, Warkany J. Medical progress. Congenital malformations: etiologic factors and their role in prevention (first of two parts). N Engl J Med 1983: 308 (8): 424-31.

2 Steele MW, Breg WR, Jr. Chromosome analysis of human amniotic-fluid cells. Lancet 1966: 1 (7434): 383-5.

3 Caspersson T, Zech L, Johansson C, Modest EJ. Identification of human chromosomes by DNA-binding fluorescent agents. Chromosoma 1970: 30 (2): 215-27.

4 Lindsten J, Zetterstrom R, Ferguson-Smith M. Prenatal diagnosis of genetic disorders of the foetus. Acta paediatrica Scandinavica 1976: (259): 1-99.

5 Mansfield ES. Diagnosis of Down syndrome and other aneuploidies using quantitative polymerase chain reaction and small tandem repeat polymorphisms. Human molecular genetics 1993: 2 (1): 43-50.

6 Pertl B, Kopp S, Kroisel PM, et al. Quantitative fluorescence polymerase chain reaction for the rapid prenatal detection of common aneuploidies and fetal sex. American journal of obstetrics and gynecology 1997: 177 (4): 899-906.

7 Bui TH, Blennow E, Nordenskjold M. Prenatal diagnosis: molecular genetics and cytogenetics. Best practice & research 2002: 16 (5): 629-43.

8 Cirigliano V, Voglino G, Marongiu A et al. Rapid prenatal diagnosis by QF-PCR: evaluation of 30,000 consecutive clinical samples and future applications. Ann NY Acad Sci 2006: 1075: 288-298.

9 Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic acids research 2002: 30 (12): e57.

10 Vissers LE, de Vries BB, Osoegawa K et al. Array-based comparative genomic hybridization for the genome wide detection of submicroscopic chromosomal abnormalities. Am J Hum Genet 2003: 73 (6): 1261-70.

11 Hochstenbach R, Meijer J, van de Brug J et al. Rapid detection of chromosomal aneuploidies in uncultured amniocytes by multiplex ligation-dependent probe amplification (MLPA). Prenat Diagn 2005: 25: 1032-1039.

12 Sahoo T, Cheung SW, Ward P, et al. Prenatal diagnosis of chromosomal abnormalities using arraybased comparative genomic hybridization. Genet Med 2006: 8: 719-727.

13 Lu XY, Phung MT, Shaw CA et al. Genomic imbalances in neonates with birth defects: high detection rates by using chromosomal microarray analysis. Pediatrics 2008: 122 (6): 1310-8.

14 Boormans EMA, Birnie E, Wildschut HI et al. Multiplex ligation-dependent probe amplification versus karyotyping in prenatal diagnosis: the M.A.K.E. study. BMC Preg and Child 2008: 8:18.

15 Miller DT, Adam MP, Aradhya S 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 2010: 86 (5): 749-64.

16 Kitzman JO, Snyder MW, Ventura M et al. Noninvasive whole-genome sequencing of a human fetus. Sci Transl Med 2012: 4(137): 137.

17 Rosenberg C, Knijnenburg J, Bakker E et al. Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. J Med Genet 2006: 43 (2): 180-6.

18 Cooper GM, Coe1 BP, Girirajan S et al. A copy number variation morbidity map of developmental delay. Nat Genet 2011: 43 (9): 838-46.

19 Armengol L, Nevado J, Serra-Juhé C et al. Clinical utility of chromosomal microarray analysis in invasive prenatal diagnosis. Hum Genet 2011: 131 (3): 513-523.

20 Dahl K, Kesmodel U, Hvidman L, Olesen F. Informed consent: providing information about prenatal examinations. Acta Obstet Gynecol Scand. 2006: 85 (12): 1420-5.

21 Li DK, Karlberg K, Wi S, Norem C. Factors influencing women's acceptance of prenatal screening tests. Prenat Diagn 2008: 28 (12): 1136-43.

22 Darilek S, Ward P, Pursley A et al. Pre- and postnatal genetic testing by array-comparative genomic hybridization: genetic counseling perspectives. Genet Med 2008: 10 (1): 13-18.

23 Saal HM. When the Clinician Disagrees With the Patient's Decision. Cleft Palate Craniofac J 2002: 39 (2): 174-8.

24 Santalahti P, Hemminki E, Latikka AM, Ryynänen M. Women's decision-making in prenatal screening. Soc Sci Med 1998: 46 (8): 1067-76.

25 van den Berg M, Timmermans DR, Ten Kate LP, van Vugt JM, van der Wal G. Are pregnant women making informed choices about prenatal screening? Genet Med 2005: 7 (5): 332-8.

26 Agnieszka S, OElêzak R, Pesz K, Gil J, Ssiadek MM. Prenatal diagnosis - principles of diagnostic procedures and genetic counseling. Folia Histochemica et cytobiologica 2007: 45 (1): 11-16.

27 Gates EA. Communicating risk in prenatal genetic testing. J Midwifery Womens Health 2004: 49 (3): 220-7.

28 Bernhardt BA, Biesecker BB, Mastromarino CL. Goals, Benefits, and Outcomes of Genetic Counseling: Client and Genetic Counselor Assessment. Am J Med Genet 2000: 94 :189–197.

29 van den Berg M, Timmermans DR, Kleinveld JH, Garcia E, van Vugt JM, van der Wal G. Accepting or declining the offer of prenatal screening for congenital defects: test uptake and women's reasons. Prenat Diagn 2005: 25: 84–90.

30 Press N, Browner CH. Characteristics of women who refuse an offer of prenatal diagnosis: data from the California maternal serum alpha fetoprotein blood test experience. Am J Med Genet 1998: 78 (5): 433-45.

31 Dormandy E, Tsui EY, Marteau TM. Development of a measure of informed choice suitable for use in low literacy populations. Patient Educ Couns 2007: 66 (3): 278-95.

32 Bell M, Stoneman Z Reactions to prenatal testing: reflection of religiosity and attitudes toward abortion and people with disabilities. Am J Ment Retard 2000: 105 (1): 1-13.

SUPPLEMENTAL MATERIAL: SURVEY

Age:.....

Gender:.....

Place of birth:.....

Place of residence:.....

Education level:

1. Middle school
2. High school
3. University

Marital status:

1. Single
2. Married
3. Divorced
4. Separated
5. Widowed
6. Co-habiting partner

Religious beliefs:

1. Atheistic
2. Agnostic
3. Christian
4. Muslim
5. Other

Why will you undergo amniocentesis or chorionic villous sampling?

1. Ultrasound finding
2. Altered biochemical screening
3. Maternal age
4. Family history

Have you been pregnant before?

.....

Do you have children?

.....

Have you had a miscarriage?

.....

Have you ever decided to terminate a pregnancy?

.....

Before this pregnancy, have you ever undergone prenatal invasive sampling (amniocentesis or chorionic villous sampling)?

1. Yes
2. No

Do you have a family history of a genetic disease, congenital malformation or mental retardation?

.....
.....
.....

Have you ever got pregnant with the aid of assisted reproductive technologies? If the answer is yes, which technique has been used?

.....
.....
.....

Do you think you have enough information to make a decision regarding prenatal analysis extension?

.....
.....
.....

Why have you decided to participate or not to participate in this study regarding new molecular techniques in prenatal diagnosis?

.....
.....
.....

How would you define your anxiety level?

1. Very low
2. Low
3. Intermediate
4. High
5. Very high

CHAPTER 4

DNA methylation abnormalities in congenital heart disease

Clara Serra-Juhé, Ivon Cuscó, Aïda Homs, Raquel Flores, Benjamín Rodríguez-Santiago, Teresa Vendrell, Núria Torán, Luis A. Pérez-Jurado

In preparation

Epigenetics, including DNA methylation and histone modification, has been recently identified as one of the most promising fields in a broad number of diseases, including cancer, autism or neurodegenerative disorders. Considering the relevant role of epigenetics in the regulation of gene expression, alterations in those mechanisms might be also related to aberrant gene expression and altered development pathways leading to congenital malformations.

This work was designed in order to study the role of methylation alterations in congenital malformations using CHD as a disease model. In order to elucidate the role of methylation in heart development and heart disease, samples with isolated CHD as well as samples with Down syndrome and CHD were included, regarding the high prevalence of CHD in patients with trisomy 21. Down syndrome was considered an appropriate disease to explore methylation patterns because all Down syndrome patients share the same genetic alteration although distinct clinical features are present among them. Therefore, other factors may determine the presence or absence of some of the clinical manifestations, such as CHD.

To explore the role of methylation alterations in CHD, a genome-wide methylation array was used to analyze isolated CHD, as well as CHD associated to Down syndrome.

DNA methylation abnormalities in congenital heart disease

Clara Serra-Juhé, Ivon Cuscó, Aïda Homs, Raquel Flores, Benjamín Rodríguez-Santiago, Teresa Vendrell, Núria Torán, Luis A. Pérez-Jurado

ABSTRACT

Congenital heart defects (CHD) represent the most common malformation at birth occuring also in ~50% of Down syndrome (DS) patients and are assumed to have multifactorial etiology, but the main cuses are largely unknown. We have searched for potential abnormal methylation profiles on developing heart-tissue DNA in samples with syndromic (DS) and non-syndromic CHD compared to controls. Among the significant differences, the vast majority of sites were hypermethylated in CHD patients comparing to controls altering genes involved in growth regulation, apoptosis and folate pathway. A likely pathogenic epimutation in the *MSX1* gene, involved in outflow tract morphogenesis, was found in a fetus with isolated CHD. In addition, several patients with DS and CHD also showed epimutations in a few candidate genes. Finally, hypermethylation of the *GATA4* gene was present in all fetuses with DS with or without CHD as well as in fetuses with isolated CHD. In summary, our data indicates that relevant epigenetic alterations are present in the developing heart DNA in fetuses with CHD and may contribute to the pathogenesis of the malformation. Further work is granted to better define the role of epigenetics in CHD.

INTRODUCTION

Epigenetic mechanisms contribute to the regulation of multiple physiological processes concerning development or aging (1). Among epigenetic mechanisms, aberrant DNA methylation has been recently associated to several diseases including cancer (2, 3), diabetes (4, 5) or psychiatric disorders (6, 7, 8). DNA methylation, by the addition of a methyl group in 5' carbon of citosines, alters the structure of the DNA molecule resulting in possible modifications of gene expression patterns. Aberrant DNA methylation may result in changes in transcription and subsequently in gene expression and may have important role in human disease.

Congenital heart defects (CHD) represent a high percentage of clinically significant birth defects occurring in ~8 per 1,000 live births, making CHD the most common malformation and an important public health burden (9, 10). Although there is strong evidence that genetics plays a major role, environmental factors are also likely to contribute and the molecular bases remain unclear in the majority of cases. The incidence of congenital heart defects (CHD) in patients with trisomy 21 is between 43% and 58% (11, 12). Given that all patients with Down syndrome (DS) carry an identical chromosomal alteration, other factors, genetic or epigenetic, might be involved in the development of the CHD.

A possible implication of methylation abnormalities in CHD has been recently explored by studying the role of the methylation regulatory folate-pathway in relation to this type of malformations (13, 14). Several SNPs in the reduced folate carrier gene (*SLC19A1*) showed significant association with the incidence of CHD in DS patients. In addition, over transmission of functional *MTHFR* gene polymorphisms to DS cases with CHD with respect to controls, suggested that disruption of the folate

pathway may contribute to CHD in DS (13). The folate-pathway has also been related with isolated CHD by several studies, with a confirmatory meta-analysis showing association of *MTHFR* variants with increased risk for CHD (14).

However, methylation profiles have not been explored in relation to CHD to date. Considering the relevant role of epigenetics in the regulation of gene expression in development and the increasing evidence linking epigenetic alterations with congenital malformations, we have searched for potential abnormal methylation profiles on developing heart-tissue DNA in samples with syndromic and non-syndromic CHD compared to controls.

METHODS

Samples

We selected 22 heart tissue samples from four different groups of fetuses from medically terminated pregnancies: fetuses with normal development, fetuses with DS without CHD, fetuses with DS with CHD and fetuses with CHD not presenting DS.

To minimize the potential bias introduced by other variables known to influence methylation profiles, such as developmental age and gender, patient samples and control samples were matched by gestational age and gender (15, 16). Heart tissue was used in all cases.

It was not possible to match the DS and no-DS samples for the type of CHD in all cases. Since the most frequent heart malformations in DS are septal defects (11, 12) with relatively good prognosis, parents usually do not decide to terminate the pregnancy in the case of isolated malformation but do it in DS. For this reason, isolated CHD (iCHD) were generally more severe than the CHD in DS. In addition to fetal ultrasound data (including fetal echocardiography) a detailed necropsy report was available for all samples included in the study. It was useful in order to define precisely the CHD detected by ultrasound or to assure the absence of heart malformation in control samples.

Trisomy 21 was diagnosed or discarded using prenatal QF-PCR and/or GTG banding chromosome analysis. All patients presenting heart malformations were also analyzed using a chromosomal microarray analysis to discard the presence of large rearrangements related to CHD.

In table 1, there is an overview of the samples, weeks of pregnancy, gender and type of CHD.

DNA was extracted from heart tissue using the Gentra Puregene Blood kit (Qiagen) according to manufacturer's instructions.

Methylation array

All 22 samples were studied by using the Illumina Infinium Human Methylation array Platform which covers 27,000 cytosine positions located in genes, intergenic regions, CpG islands and miRNA. Samples from different groups were randomly distributed in the chip to avoid position bias. A biological replica and a technical replica were included in order to determine the quality of the experiment. Quality controls were performed according to manufacturer instructions.

Bioinformatic and statistical analysis

Two different types of data analysis were performed in order to detect methylation alterations between groups and also in individual samples, with and without previous data normalization. The results of both analyses were very similar.

The software provided by Illumina, Genome Studio, was used to analyze differences in methylation between groups and among samples. β values are used, which range from 0 (completely unmethylated) to 1 (completely methylated).

A second analysis was done by using an R package. Data normalization was performed using R 2.12.0, Cran HumMeth27QCReport, based on Lumi package quantile normalization. This package transforms beta values into M-values, which range from -6 (completely unmethylated) to 6 (completely methyalted). The differential methylation analysis was done using limma R Package with the statistical eBayes function (17, 18).

Quality control

Quality controls were performed according to manufacturer instructions. To check batch effects, the correlation between the technical replicates was analyzed and the correlation parameter was above 95% (r^2 = 99.7%). Bisulfite conversion was also considered and values for all samples were above the minimum required (4000) described in previous articles.

Differential methylation analyses

Considering the probable heterogeneity of causes underlying CHD, analyses comparing individual samples with control group were performed. Each sample with iCHD was compared with the group of 4 controls and each sample with CHD and DS was compared with the control group as well as with the group of DS fetuses without CHD. A CG was considered to be differentially methylated (DM) if the β value difference was greater than ±0.25 in not normalized data. In the analyses with normalized data, DM was considered significant if the log fold change was greater than ±1.8 and with a p-value<0.05. The CpG located in the same position as a described SNP were discarded because of the possibility that differences in methylation were a false positive due to the SNP.

Comparisons between groups were also done by using Genome Studio and R package. Genome Studio did not provide any DMCpG between groups. For this reason, comparisons between groups were done with normalized data. Only DMCpG with a p-value lower than 0.05 and a log Fold change greater than ±0.7 were considered.

Validations

To validate the predicted alterations using Infinium Human Methylation arrays Platform 27k two different techniques were used: a methylation specific Multiple Ligation-Dependent Probe Amplification (MS-MLPA) and EpiTYPER analysis.

MS-MLPA

A MS-MLPA assay was designed to validate methylation differences detected by the array. A total of 100 ng of genomic DNA from heart tissue was subject to MLPA using specific probes designed to detect methylation levels in target CpG (19, 20). In MS-MLPA to parallel assays are performed in order to detect copy number changes and methylation patterns in selected CpG. For the methylation analyses, a sensitive methylation enzyme (HhaI) was used. This enzyme differentiates methylation status of its target sequence ad only digests DNA if it is methylated. All MLPA reactions were analyzed on an ABI PRISM 3100 Genetic analyzer according to manufacturers' instructions. Each MLPA signal was normalized and compared to the corresponding peak height obtained in control samples. MS-MLPA probes used in validations are shown in Supplementary material table 1.

EpiTYPER

In order to validate the results obtained in Infinium Human Methylation arrays Platform 27k, Sequenom MALDI-TOF technology was used to analyze 9 regions (15 CpG). The first step in this procedure is the treatment of DNA with bisulfate in order to convert cytosine nucleotides into uracile. After the

conversion, an enzyme cleaves specifically at some nucleotides giving different patterns depending on the uracile and cytosine content, that it to say, depending on methylation state (21, 22). Those fragments are quantitatively measured by MALDI-TOF technology.

A technical replicate was done for all the samples and regions and the ones with a SD greater than 0.1 were removed. The procedure was performed according to manufacturer instructions.

Enrichment analyses

The genes nearby the DMCpG were analyzed using a computational resource, Consensus Path DB (23), to obtain an overview of the pathways which could be altered, with the aim of identifying overrepresented pathways. The p-value of each pathway was considered ($p<0.05$).

RESULTS

Whole genome methylation pattern

Globally, a total of 17,877 (64,8%) CpG were found to be hypomethylated, 5,370 (19,5%) were identified as hemimethylated and the remaining 4,331 (15,7%) hypermethylated, as shown in figure 1.

Dendograms were obtained in order to study the distribution of the samples (figure 2). A Manhattan dendogram with all 22 samples is presented showing that samples cluster regardless of phenotype or gender, after excluding the X chromosome. A Manhattan dendogram for groups is also shown. The most resembling samples were from DS patients, regardless the presence or absence of CHD, even when chromosome 21 was excluded for the data.

No clear methylation patterns were identified among the different groups analyzed; samples in the same group did not cluster closer than samples in different groups. Nevertheless, enough differences in methylation profiles were present between fetuses with trisomy 21 and fetuses with a normal karyotype, since the two groups with Down syndrome cluster closer than fetuses with normal development and fetuses with iCHD.

Differential methylation analyses: individual cases vs controls

Each subject with CHD was compared to the control group. For each DS-CHD subject, a comparison with the DS without CHD group was also done.

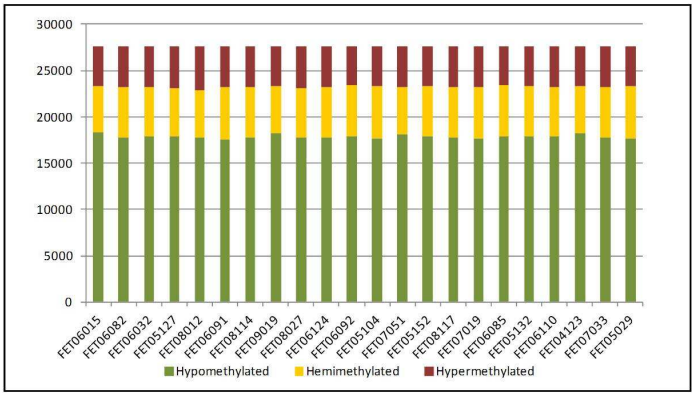


Figure 1. Proportion of hypomethylated, hemimethylated and hypermethylated CpG per samples considering the 27,578 CpG included in the Illumina Infinium Human Methylation arrays Platform 27k.

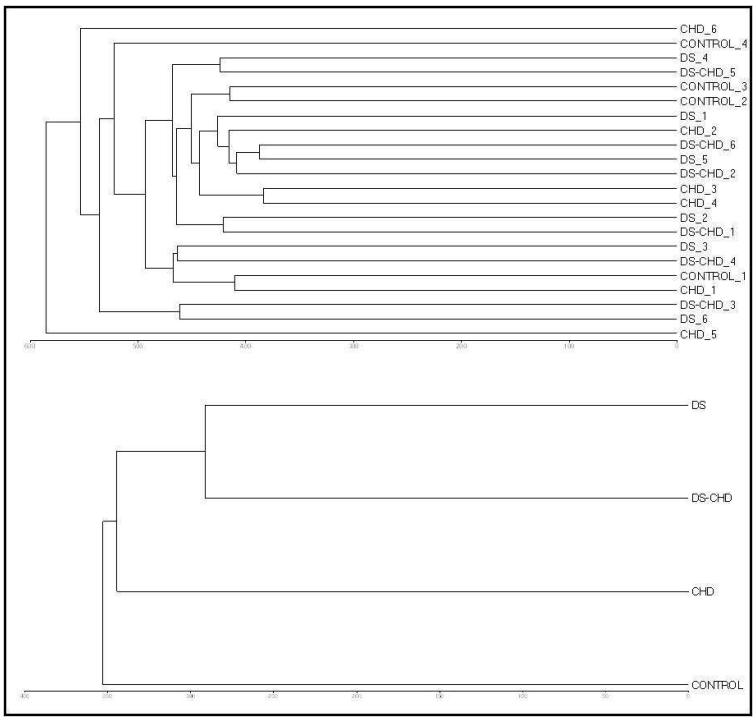


Figure 2. A: Manhattan dendogram with the 22 samples included in the study; samples cluster randomly considering gender and group (X chromosome was excluded from the analysis). B: Manhattan dendogram with the four different groups used in the study. DS samples (with and without CHD) cluster closer comparing to iCHD samples and fetuses with normal development.

A total of 26 DMCpGs corresponding to 18 regions were found in individual iCHD cases respect to the control group (Table 2). While 14 regions showed a hypermethyaltion status in iCHD, the remaining regions presented hypomethyaltion. One third (6/18) of the CpG sites were located in promoter regions and 12 in intragenic regions.

The comparison between individual DS-CHD cases with two control groups (normal development and DS without CHD) pointed out 21 DMCpG corresponding to 18 regions (Table 3): 13 regions showed a hypermethyaltion status in DS-CHD samples and the remaining presented hypomethyaltion. Regarding location, 6 lied in promoter regions, 8 in intragenic regions and 4 in both.

A total of 13 DMCpG corresponding to 13 regions were identified in more than one sample of fetuses with CHD: 7 regions showed a hypermethylation status in CHD samples compared to controls and the remaining ones presented a hypomethyaltion. 2 of them were located in promoter regions and 11 in intragenic regions. Results are shown in table 4.

Differential methylation analyses: cases vs controls

Fewer DMCpG were identified when compared different groups of patients with controls. DMCpG were only obtained when compared DS fetuses (DS and DS-CHD) versus controls and all patiens (DS, DS-CHD

and iCHD) versus controls. Only 10 DMCpG were identified between DS (with CHD and without CHD) and controls, corresponding to 10 different regions: 1 CpG was located in a promoter region and 9 in intragenic regions. For 7 of them patient samples showed hypermethylation with respect to controls and for the remaining ones, controls presented a hypermethylation (Table 5). Comparing all patients' samples (DS, DS-CHD and CHD) with controls 13 DMCpG corresponding to 8 different regions were identified, being 3 CpG in promoters and 10 intragenic. For 11 of them patient samples presented a hypermethylation state compared to control samples and for the remaining ones, controls showed a hypermethylation state (Table 6).

Validations

Two alterations identified by Infinium Human Methylation arrays Platform 27k could be validated by MS-MLPA. This multiple assay was used to check the hypermethylation in *MSX1* in one of the patients presenting double outlet right ventricle and the hypermethylation in two DS-CHD samples in the *CHES1/FOXN3* gene. Both methylation alterations were validated.

By using EpiTYPER, 8 regions corresponding to 14 CpG were analyzed using the same DNA samples. The differences in methylation patterns were clearly validated for 4 of the regions (*LYCAT*, *NALP2*, *ACRBP* and *MSX1*). In 3 cases (*MTHFD2*, *SERPINB5* and *KIAA0101*), the EpiTYPER results showed the same tendency as the array results, but the difference was not enough to definitively validate the alteration. For the 7 remaining region, it was not possible to conclude if the methylation pattern was different than in controls due to technical problems.

Enrichment analyses

Two different overrepresentation analyses were done with the genes closed to DMCpG in samples with iCHD and DS-CHD, respectively. For the genes in DS-CHD 4 pathways were found to be enriched and for the samples with iCHD 7 pathways were identified as enriched ones. The same analysis was also done with the genes identified in comparisons between groups. 5 pathways emerged as being enriched in patients (DS, DS-CHD, CHD) versus controls and 1 in DS patients (with and without CHD) versus controls. The gene ontology term, the genes contained in the pathway and the p-value are shown in table 7.

DISCUSSION

Although no clearly differentiated patterns of abnormal methylation have been identified in relation to CHD in this study, a few remarkable changes have been defining indicating that epigenetic abnormalities likely contribute to the etiology of some CHD. Interestingly, the vast majority of DMCpG showing differences were hypermethylated in CHD patients comparing to controls (73.2% hypermethylated versus 26.8% hypomethylated). Repression in gene expression is expected when the promoter region is methylated, as methylation is one of the mechanisms to silence DNA expression, while promoter hypomethylation is generally related to increased expression (24, 25). Given the global finding of increased hypermethylation in some gene promoters of our cohort of patients, reduced gene expression in developing heart tissue seems more relevant in CHD.

Overrepresentation analysis identified enriched pathways related to cell growth and cell death in iCHD samples. The regulation of both processes is essential in development, as the formation of several structures requires a precise control of cell and tissue growth. Therefore, the misregulation of pathways related to cell growth and cell death could cause an abnormal heart development. Mainly at early stages in development, cardiomyocytes have a high proliferation rate, which decreases progressively in late

gestation (26). As an example, a proliferating-center in the caudal coelomic wall has been proved to be crucial in the elongation of the heart tube at both its venous and arterial pole, providing a morphological mechanism for early heart formation; when this process fails, the formation of the atria and the right ventricle is impaired (27).

Some genes related to cell growth and apoptosis have already been reported in association to CHD. *Hes1* knockout mouse embryos display defects in proliferation at earlier developmental stages which induce a reduction in cardiac neural crest cells and failure to completely extend the outflow tract (28). Increased postmigratory neural crest cell apoptosis in a mice model defective for *Alk5* leads to severe outflow tract defects (29). The identification of pathways related to cell growth and cell death in fetuses with iCHD might provide insights into pathways and genes potentially related to heart development and heart disease.

Folate pathway was also overrepresented in the enrichment analysis. Folate pathway has been explored in relation to different types of malformations including CHD, as isolated malformations or in Down syndrome patients. Some SNPs have been reported in associations with a higher risk of CHD in patients with trisomy 21 (13). In our study, genes in this pathway were found to be hypo or hypermethylated in patients with CHD compared to fetuses with normal development, including alterations of the *MTHFS* in one case and *MTHFD2* in three. Our data further reinforce the hypothesis that deregulation of the folate pathway, either by genetic mutations or epigenetic mechanisms, can contribute to CHD.

Hypermethylation of a CpG island located in the gene *MSX1* has been documented in a single fetus with double outlet right ventricle with IVC and hypoplasia of the ascending aorta (figure 3). *MSX1* is a strong candidate for the heart malformation presented in the fetus given its role cardiac development. Chen *et al.* (30) described, by using a mice model, the relevant role of *Msx1* and *Msx2* in outflow tract morphogenesis by both protecting secondary heart field precursors against apoptosis and inhibiting excessive proliferation of cardiac neural crest, endothelial and myocardial cells in the conotruncal cushions. They also showed that the impairment of *Msx1* and *Msx2* was associated with reduced expression of *Hand1* and *Hand2* (31). As a result of Msx1 and Msx2 disruption, mice model presented outflow tract malalignment defects including

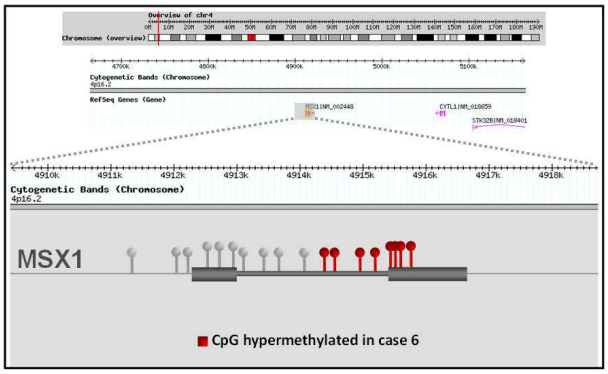


Figure 3. Hypermethylation of 8 CpGs located in *MSX1* (implicated in cardiac development in animal models) in case 17. The position of the gene with respect to the chromosomal band 4p16.2 is shown above. The location of the CpGs with respect to the two exons of the gene and the hypermethylated sites (in red) is shown below.

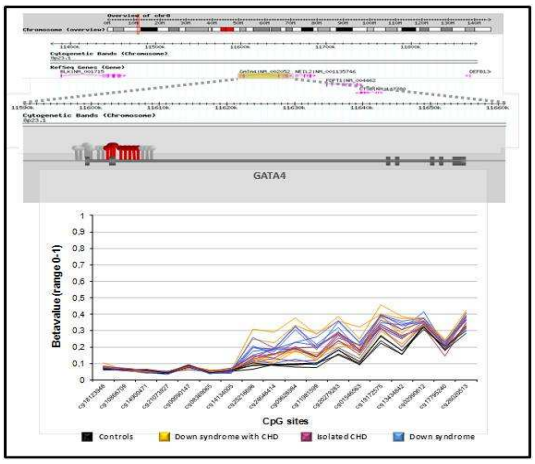


Figure 4. We observed hypermethylation of several CpGs in *GATA4* patients with respect to controls. A: the position of the gene with respect to the chromosomal band 8p23.1 is shown above and the *GATA4* CpGs analyzed below. In grey, *GATA4* CpGs analyzed; in red, the CpGs significantly altered in patients. *B:*Beta value of the CpGs in controls (black) with respect to the 3 groups of samples (colored lines).

double-outlet right ventricle, overriding aorta and pulmonary stenosis. The hypermethylation in the fetus is likely an epimutation causative of the phenotype of this patient through impaired *MSX1* expression leading to the abnormal heart development.

In two other fetuses with DS and CHD a hypomethylation has been identified in a CpG in the *KLF17* gene. This gene, from the Kruppel-like family, is a transcription repressor that binds to the promoter of target genes and prevents their expression acting as a negative regulator of epithelial-mesenchymal transition (32). Another gene related to tumor development (33), *FAM107A*, has been found to be altered in two different samples with iCHD. In one case there was a hypermethylation and in the other case a hypomethylation; one fetus had tetralogy of fallot and the other one right heart hypoplasia. In addition, a hypermethylation in another gene potentially related to cell growth control, *CHES1/FOXN3*, a checkpoint suppressor in cell cycle (34), was identified in two samples with DS and CHD.

In three samples with CHD, two DS fetuses and one fetus with normal karyotype, a hypermethylation was detected in the gene *TCL1B*. This gene interacts with *AKT1* and *AKT2* and enhances the phosphorilation and activation of both mentioned genes. Interestingly, *AKT1* and *AKT2* have been reported in relation to heart function and heart development. The animal model with a disruption in *Akt1* displays CHD and a reduction in cell proliferation (35).

When comparing the methylation patterns between groups (figure 4), *GATA4*, a known gene associated to CHD (36), was identified as being hypermethylated in all patients respect to controls. This hypermethylation was present in all fetuses with DS with or without CHD as well as in fetuses with iCHD. Considering these results, it might be hypothesized that *GATA4* is a gene whose methylation can be deregulated by the effects of the trisomy of chromosome 21 contributing to the high incidence of CHD in these patients, although the main modifier for the penetrance of the phenotype is still unknown. Abnormal hypermethylation was also identified in fetuses with iCHD; this alteration could be a consequence of a misregulation of expression patterns in cardiac tissue due to the underlying cause of CHD in each patient.

In summary, our data strongly indicates that some epigenetic alterations are present in the developing heart DNA in fetuses with CHD and may contribute to the pathogenesis of the malformation through deregulation of gene expression, providing novel genes relevant for heart development. Further work is granted to better define the role of epigenetics in CHD and its relationship with genetic sequence and environment.

Sample	Tissue	Gender	Weeks of pregnancy	Phenotype
1	Heart	Female	20	Normal development
2	Heart	Female	20	Normal development
3	Heart	Female	15	Normal development
4	Heart	Male	15	Normal development
5	Heart	Female	19	Trisomy 21
6	Heart	Female	22	Trisomy 21
7	Heart	Female	22	Trisomy 21
8	Heart	Female	22	Trisomy 21
9	Heart	Female	18	Trisomy 21
10	Heart	Male	20	Trisomy 21
11	Heart	Female	22	Trisomy 21 and ventricular septal defect
12	Heart	Female	22	Trisomy 21 and ventricular septal defect
13	Heart	Female	18	Trisomy 21 and auriculoventricular septal defect, coarctation of the preductal aorta
14	Heart	Female	19	Trisomy 21 and tetralogy of Fallot
15	Heart	Male	22	Trisomy 21 and ventricular septal defect
16	Heart	Female	18	Trisomy 21 and left heart hypoplasia, hypoplasia of the ascending aorta
17	Heart	Male	22	Double outlet right ventricle with IVC, hypoplasia of the ascending aorta
18	Heart	Female	22	Tetralogy of Fallot
19	Heart	Female	21	Left heart hypoplasia with mitral and aortic valves atresia and hypoplasia of the ascending aorta
20	Heart	Female	22	Double outlet right ventricle, large foramen ovale, VSD, coarctation of the preductal aorta
21	Heart	Female	22	Tricuspid valve stenosis, right heart hypoplasia, VSD ostium primum, muscular VSD, mitral valve dysplasia, aneurismal dilatation of the posterior left atrium
22	Heart	Female	22	Truncus arteriosus type II, absent ductus arteriosus, truncal valve with three leaflets, VSD

Table 1. List of heart malformations present in the cohort of 22 studied fetuses. VSD: ventricular septal defect.

CG	Coordinate	Location	Nearby gene	State controls	State patients
cg10331038	chr1:173642651	intragenic	*TNR*	Hyper	Hemi
cg17749961 and cg10995359	chr2:30523674	intragenic	*LYCAT*	Hypo	Hemi
cg07314414	chr2:128501393	intragenic	*SAP130*	Hypo	Hemi
cg22609784,cg09573795,cg27038439,cg24840099, cg09748975,cg03843978,cg20891301,cg01785568	chr4:4914579-4915734	intragenic	*MSX1*	Hemi	Hyper
cg23244913	chr6:30051380	intragenic	*HCG9*	Hypo	Hemi
cg14290451	chr6:35544114	promoter	*RPL10A, FANCE*	Hypo	Hemi
cg02123547	chr6:37245442	promoter	*PIM1*	Hypo	Hemi
cg22552684	chr7:42918309	intragenic	*C7orf25*	Hypo	Hemi
cg06713098	chr7:45926892	intragenic	*IGFBP3*	Hypo	Hemi
cg12876594	chr9:35781798	promoter	*NPR2*	Hypo	Hemi
cg01853981	chr12:51232749	intragenic	*KRT71*	Hyper	Hemi
cg26843567	chr12:103370411	promoter	*CHST11*	Hyper	Hemi
cg16826718	chr12:115803042	intragenic	*HRK*	Hypo	Hemi
cg07684353	chr15:57452091	intragenic	*MYO1E*	Hypo	Hemi
cg23855392	chr15:77976837	promoter	*MTHFS*	Hypo	Hemi
cg09953122	chr20:23419693	promoter	*CST8*	Hyper	Hemi
cg03697316	chr22:19666741	intragenic	*LZTR1*	Hypo	Hemi
cg12078929	chr22:41226632	intragenic	*SERHL*	Hypo	Hemi

Table 2. DMCpG identified comparing individually fetuses with iCHD with the control group. Hg coordinates location of the CpG, nearby genes and the methylation state in samples and controls are shown.

CG	Coordinate	Location	Nearby gene	State controls	State patients
cg20345840	chr1:43197004	intragenic	*SLC2A1*	Hypo	Hemi
cg17749961	chr2:30523367	promoter region	*LYCAT*	Hypo	Hemi
cg10353870	chr2:127581248	intragenic	*BIN1*	Hypo	Hemi
cg06509940	chr3:120761604	intergenic	*CD80*	Hyper	Hemi
cg18271969	chr3:185254193	intragenic	*HTR3C*	Hyper	Hemi
cg17960516	chr4:3434802	promoter region	*DOK7*	Hypo	Hemi
cg09339527	chr5:178255358	promoter region	*ZFP2*	Hypo	Hemi
cg09638834 and cg07579404	chr6:150388190-150388414	intragenic/promoter region	*RAET1L*	Hypo	Hemi
cg07579404	chr7:129380246	promoter region	*UBE2H*	Hypo	Hemi
cg25007250	chr7:150318645	promoter region	*NOS3*	Hyper	Hemi
cg08618113	chr15:39311174	intragenic	*CHP*	Hypo	Hemi
cg02065795 and cg19839691	chr15:62460508-62460843	intragenic/promoter region	*KIAA0101*	Hypo	Hemi
cg25172835	chr15:63458780	promoter region	*IGDCC3, IGDCC4*	Hyper	Hemi
cg15473868	chr16:55274122	intragenic	*MT1X*	Hypo	Hemi
cg08785922	chr17:829581	intragenic	*NXN*	Hypo	Hemi
cg02017155, cg23441676	chr19:40541233-40553702	intragenic/promoter region	*FFAR3/GPR42*	Hemi	Hypo
cg00695416	chr21:36364346	intragenic	*CBR1, SETD4*	Hypo	Hemi
cg22903370	chrX:152712054	intragenic, promoter region	*IDH3G, SSR4*	Hemi	Hyper

Table 3. DMCpG identified comparing individually fetuses with DS-CHD and two control groups (fetuses with normal development and fetuses with DS without CHD). Hg coordinates location of the CpG, nearby genes and the methylation state in samples and controls are shown.

CG	Coordinate	Location	Nearby gene	Samples with DM pattern	State controls	State patients
cg01446393	chr3:58538649	intragenic	*FAM107A*	2 CHD	Hemi	Hyper/Hypo
cg17711541	chr6:26232683	intragenic	*HIST1H2AC, H2A/1*	2 CHD, 2 DS	Hypo	Hemi
cg17607973	chr7:99865344	intragenic	*BCDIN3, MEPCE*	CHD, DS-CHD	Hypo	Hemi
cg12437481	chr16:360113	intragenic	*MRPL28*	2 CHD, DS-CHD, DS	Hypo	Hemi
cg24330042	chr22:22714159	intragenic	*GSTT1*	2 CHD, 2 DS, 1 DS-CHD	Hypo	Hemi
cg08679985	chr1:44356937	intragenic	*KLF17*	2 DS-CHD	Hyper	Hemi
cg17687883	chr2:74278975	promoter region	*MTHFD2*	2 CHD, 1 DS-CHD, 1 CONTROL	Hypo	Hemi
cg22022041	chr3:45903141	intragenic	*CCR9 i LZTFL1*	DS-CHD, DS	Hyper	Hemi
cg01962826	chr6:34208945	intragenic	*GRM4*	DS-CHD, CHD	Hyper	Hemi
cg01120761	chr12:7794437	promoter region	*CLEC4C*	4 DS-CHD, 1 DS, 1 CHD	Hyper	Hemi
cg22228134	chr14:24148449	intragenic	*GZMH*	DS-CHD, CHD	Hyper	Hemi
cg05923056	chr14:88952999	intragenic	*CHES1*	2 DS-CHD	Hypo	Hemi
cg13771579	chr14:95222519	intragenic	*TCL1B, TCL6, TML1*	2 DS-CHD, 1 CHD	Hemi	Hyper

Table 4. DMCpG identified in more than one sample with CHD when compared to controls. Hg coordinates location of the CpG, nearby genes and the methylation state in samples and controls are shown.

CG	Coordinates	Location	Nearby gene	Sample 1	Sample 2	Sample 3	Sample 4	Average patients
cg16970828	chr3:49826357	intragenic	*UBA7*	-1,535	-2,439	-1,990	-2,604	-3,221
cg03136712	chr5:140033252	intragenic	*DND1*	2,026	3,104	2,100	2,984	3,476
cg25683185	chr12:6626349	intragenic	*ACRBP*	-0,222	-0,158	-0,804	-0,337	0,830
cg08411049	chr18:59295230	intragenic	*SERPINB5*	1,775	1,602	1,477	1,129	2,122
cg02218324	chr19:51010279	intragenic	*RSHL1*	-1,991	-1,599	-2,000	-2,086	-1,122
cg04052038	chr21:30509994	intragenic	*CLDN8*	2,250	2,558	2,391	2,453	3,114
cg20423977	chr21:41479511	intragenic	*BACE2*	2,101	2,187	1,731	1,616	2,543
cg00747849	chr22:17545386	intragenic	*SLC25A1*	-3,372	-2,977	-3,306	-3,622	-2,465
cg07359545	chr22:18091327	intragenic	*GP1BB/SEPT5*	-2,292	-0,915	-1,897	-1,232	-0,614
cg03389133	chr22:35938557	promoter region	*SSTR3*	-2,500	-1,796	-2,249	-2,136	-1,475

Table 5. DMCpG identified comparing DS and DS-CHD patients as a group with control samples. Hg coordinates location of the CpG, nearby genes and the methylation state in samples and controls are shown.

CG	Coordinates	Location	Nearby gene	Sample 1	Sample 2	Sample 3	Sample 4	Average patients
cg19695867	chr1:109386913	promoter region	*WDR47*	-0,335	-1,310	-0,519	-1,274	-1,736
cg03584220	chr1:166780191	promoter region	*XCL2*	1,564	1,641	0,541	1,668	2,467
cg16970828	chr3:49826357	intragenic	*UBA7*	-1,535	-2,439	-1,990	-2,604	-3,198
cg16609957	chr5:132189766	intragenic	*SHROOM1*	-2,277	-2,016	-1,913	-1,408	-1,036
cg09626984	chr8:11603505	intragenic	*GATA4*	-3,820	-3,673	-3,693	-4,068	-1,939
cg20279283	chr8:11604272	intragenic	*GATA4*	-2,638	-2,123	-2,133	-2,618	-1,264
cg01546563	chr8:11604598	intragenic	*GATA4*	-3,292	-2,862	-3,258	-3,165	-2,072
cg19172575	chr8:11605130	intragenic	*GATA4*	-1,789	-1,413	-1,664	-1,476	-0,823
cg13434842	chr8:11605305	intragenic	*GATA4*	-2,436	-2,075	-2,382	-2,275	-1,282
cg02838492	chr9:115901109	intragenic	*KIF12*	-2,277	-2,215	-2,259	-2,340	-1,546
cg10318258	chr14:23879599	promoter region	*RIPK3*	0,191	0,528	-0,092	0,547	1,122
cg17826679	chr19:10597038	intragenic	*SLC44A2*	-4,504	-3,769	-4,228	-3,504	-2,862
cg21663431	chr19:10597355	intragenic	*SLC44A2*	-3,607	-2,960	-3,496	-3,184	-2,269

Table 6. DMCpG identified comparing DS, DS-CHD and CHD patients as a group with control samples. Hg coordinates location of the CpG, nearby genes and the methylation state in samples and controls are shown.

Gene ontology term	Genes contained	p-value
Patients (DS, DS-CHD and iCHS) vs controls		
type I interferon production	*UBA7, RIPK3*	0,0003
cytokine production	*GATA4, UBA7, RIPK3*	0,0003
immune response	*XCL2, UBA7, RIPK3*	0,006
Microtubule	*SHROOM1, KIF12*	0,007
Signal transduction	*GATA4, XCL2, UBA7, SLC44A2, RIPK3*	0,009
DS and DS-CHD versus controls		
Parkinson disease	*UBA7, RIPK3*	0,005
DS-CHD vs controls		
Oxidoreductase activity	*CBR1, GSTT1, IDH3G, NOS3, MTHFD2, NXN*	0,001
caveola	*NOS3, SLC2A1*	0,005
response to heat	*NOS3, MTHFD2*	0,006
plasma membrane part	*CD80, FFAR3, GPR42, GRM4, NOS3, SLC2A1, IGDCC3, CCR9, RAET1L*	0,007
CHS vs controls		
pteridine-containing compound metabolic process	*MTHFS, MTHFD2*	0,001
cell growth	*IGFBP3, MSX1, FAM107A, TNR*	0,002
programmed cell death	*GRM4, GZMH, IGFBP3, MSX1, PIM1, HRK, CHST11*	0,005
regulation of growth	*IGFBP3, MSX1, FAM107A, TNR*	0,006
post-embryonic development	*MYO1E, CHST11*	0,007
bone development	*MSX1, NPR2*	0,008
cell death	*GRM4, GZMH, IGFBP3, MSX1, PIM1, HRK, CHST11*	0,009

Table 7. Results of over representation analysis. Enriched pathways, genes included in the and p-value for each of them are shown.

Gene	Chr	Start	End	Hybridization sequence
CHES1/FOXN3	14	89883237	89883287	CGCCCGCGAAAGTTTGTCGGCTCCTGCTGAAGGGCAGCGGCGCCCCGCAC
MSX1	4	4863672	4863723	CCCTGGCGCCAGGGGTGAACGCGATCCAACAGAGGACTGAGACAATCTAAAG

Supplementary table 1. MLPA probes used to validate the alterations detected by Illumina 27K. Hg19 assembly.

CHAPTER 5

The role of point mutations in fetuses with left heart hypoplasia

Clara Serra-Juhé, Christian Gilissen, Teresa Vendrell, Marisol del Rosario, Marloes Steehouwer, Núria Toran, Joris A. Veltman, Luis Pérez-Jurado*, Alexander Hoischen*

* shared last author

In preparation

After CMA, the development of next generation sequencing technologies has induced a revolution in genetic research increasing substantially the knowledge about the genome and, consequently, about genetic diseases. Several publications have proved the utility and efficiency of NGS in the identification of genes causing mendelian diseases. In addition, some studies have also pointed out the usefulness of this technology to elucidate the genetic mechanisms underlying diseases considered, at least until now, oligogenic or mulifactorial disorders.

In the studies described up to now in this thesis, we analyzed the role of CNV and epigenetic alterations in congenital malformations. Although some interesting findings were done, most of the cases were not solved by using CMA or methylation arrays. Therefore, the role of point mutations in congenital malformations was explored.

LHH is a severe CHD with an incidence of 2/10,000. Some studies have described a high heritability of this malformation, also in relation to other left ventricular outflow tract malformations. In addition, the study of congenital malformations by CMA revealed a high frequency of deletion-type CNVs in fetuses with heart hypoplasia compared to the control cohort, pointing out a strong genetic component among the causes underlying the disease. A cohort of 14 fetuses with LHH, previously studied by CMA, was studied by exome sequencing in order to elucidate the contribution of point mutations in LHH.

The role of point mutations in fetuses with left heart hypoplasia

Clara Serra-Juhé, Christian Gilissen, Teresa Vendrell, Marisol del Rosario, Marloes Steehouwer, Núria Toran, Joris A. Veltman, Luis Pérez-Jurado*, Alexander Hoischen*

* shared last author

Left heart hypoplasia (LHH) is a severe form of congenital heart defect (CHD) which remains a leading cause of infant mortality and childhood morbidity. The causes underlying LHH are largely unknown although several studies suggest a main contribution of genetic factors in the etiology of this malformation. The high heritability of LHH and the identification of chromosomal aberrations and point mutations in some patients support this hypothesis.

We selected 14 fetuses with isolated LHH from the fetal tissue bank of Vall d'Hebron hospital (Barcelona); all samples showed normal G-banding karyotype and no relevant copy number variants (CNVs) in a chromosomal microarray analysis (CMA). Here we aimed for genome-wide detection of point mutations. Therefore exome sequencing was performed for 2 patients-parent trios as well as 12 individual samples.

We identified one de novo mutation in each of the trios analyzed in *USP32* and *NCAPD3* respectively. An additional variant in *USP32* was identified in another fetus from the cohort. Two additional cases presented with variants in *NCAPD2*, a gene encoding for another subunit in the same complex as *NCAPD3*. In two singleton cases we identified variants in genes previously linked to CHD, *DTNA* and *ACVR2B*. In other cases several novel non-synonymous variants in possible novel candidate genes such as *FBXO25* or *PROX1* were detected. Strikingly, we identified 3 genes in which possible mutations were identified in three different samples (*ATP3A5*, *SCAF11* and *RB1CC1*).

Our data reveal a significant contribution of point mutations in LHH suggesting several new candidates, which may act in a monogenic fashion while others as part of an oligogenic disease model.

INTRODUCTION

Left heart hypoplasia (LHH) is a severe congenital heart defect (CHD) affecting 2/10,000 liveborns and causing high infant mortality and childhood morbidity. The bases underlying this severe malformation remain unknown although several studies point out genetic factors as the main cause of this type of malformations. Hinton *et al.* described a high heritability of hypoplastic left heart alone (99%) and with associated cardiovascular malformations (74%) and a sibling recurrence risk of 8% and 22%, respectively, suggesting that LHH is largely determined by genetic factors (1).

Nonetheless, to date, genetic factors contributing to LHH are largely unknown. Chromosomal defects have been described in some patients with LHH (2, 3), as well as point mutations in single gene disorders such as Holt-Oram syndrome (4). In those cases, left heart hypoplasia is usually part of a more complex syndrome which includes other major malformations and/or intellectual disability, as it has been described in Jacobsen

syndrome (5) or in the deletion in 16q24.1 encompassing the *FOX* gene cluster (6).

Regarding isolated LHH, a multifactorial model has been proposed accounting for an important proportion of cases. McBride *et al.* suggested an oligogenic pattern of inheritance for LHH and other left ventricular outflow tract malformations, obtaining the highest maximum likelihood score in a model with 2 loci (7). Consistently, a model considering one or more minor loci with rare dominant alleles was hypothesized.

In order to explore the genetic basis of isolated LHH, we used a whole exome sequencing approach to identify *de novo* and inherited mutations in a cohort of fetuses with this malformation. We focus on the analysis of this CHD as an isolated malformation in fetuses. In a previous study (8) we analyzed the role of copy number variants (CNVs) in this cohort and identified a significantly higher rate of rare CNVs in LHH comparing to control population or other cohorts of fetuses with congenital malformations (47% *vs* 18% in a control cohort). All cases in which the underlying CNVs could not explain the phenotype, were subject of the present study.

MATERIALS AND METHODS

14 samples of fetuses with LHH, 6 male and 8 female fetuses, were included in the cohort, all born from apparently health parents. Induced abortion occurred between 17 and 22 weeks of pregnancy, based on severity of the ultrasound finding. Extensive necropsy report was available for all of them (detailed information in table 1). DNA was extracted from the affected tissue when available (n=10); liver or lung tissue was used for the remaining samples with insufficient heart tissue (n=4). Parental DNA was available for 6 fetal samples. All the samples showed normal chromosomes in standard karyotype and no CNVs clearly linked to the phenotype were detected in a chromosomal microarray analysis.

We sequenced the exomes of 2 patient-parent trios as well as 12 additional samples of fetuses with isolated left heart hypoplasia. On average, 5.1Gb of mappable sequence data was obtained per samples after exome enrichment using SureSelect 50Mb human exome kit (v2, Agilent, Santa Clara, CA, USA) and SOLiD 5500 sequencing. Color space reads were mapped to the hg19 reference genome with the SOLiD LifeScope software which utilizes an iterative mapping approach. Single nucleotide variants were subsequently called by the DiBayes algorithm using the high call stringency. Small insertions and deletions were detected using the SOLiD Small InDel Tool. Called SNV variants and indels were combined and annotated using a custom analysis pipeline. On average, 80.1% of reads mapped in the region and 86.4% of the exome was covered at least 10-fold (table 2). On average 44,390 variants were called per exome, 20,520 of which were coding variants. Detailed information per sample is shown in supplementary table 1.

In order to identify potentially causative variants three different approaches were used to analyze the results. The analysis workflow is summarized in supplementary figure 1.

Firstly, a prioritization strategy was used to identify possible *de novo* mutations in both trios (9). For that analysis, non-genic, intronic and synonymous variants were excluded, as well as likely benign variants present in dbSNP or in the Nijmegen inhouse database containing variants detected in 672 exomes. Next, all inherited variants were filtered out by comparison with parental exome data. This analysis resulted in 3 and 4 possible *de novo* variants, respectively; for all those variants Sanger sequencing validation was performed.

Subsequently, an unbiased approach for the entire cohort was considered to analyze the results regardless to the gene function. With this aim, only private variants were considered (not described in dbSNP or in the Nijmegen in-house database). An overlap

Sample	Tissue	Gender	Malformation
Case 1	heart	female	LHH
Case 2	heart	male	LHH
Case 3	heart	female	L-TGA, incomplete AVSD, LHH, overriding aorta over membranous IVC
Case 4	heart	male	LHH
Case 5	heart	female	LHH
Case 6	heart	female	LHH with mitral and aortic valves atresia and hypoplasia of the ascending aorta
Case 7	heart	male	LHH with mitral and aortic valves atresia, aortic valve with two leaflets and tubular hypoplasia of the aortic arch
Case 8	heart	male	LHH with hypoplasia of mitral and aortic valves
Case 9	liver	male	LHH
Case 10	liver	male	D-TGA, IAC type ostium secundum, LHH with mitral and pulmonary valves atresia, hypoplastic pulmonary artery, permeable ductus
Case 11	lung	female	LHH with atresia of mitral and aortic valves
Case 12	lung	female	extreme LHH with mitral and aortic valves atresia and hypoplasia of the ascending aorta
Case 13	heart	female	LHH
Case 14	heart	female	LHH with mitral and aortic valves atresia, incarceration of the left ventricle and extreme hypoplasia of the ascending aorta

Table 1. List of heart malformations present in the cohort of 14 studied fetuses with left heart hypoplasia. IVC: interventricular communication; LHH: left heart hypoplasia; VSD: ventricular septal defect; D-TGA: dextro-transposition of the great arteries; L-TGA: levo-transposition of the great arteries; AVSD: atrioventricular septal defect.

Statistic	Value
Total number of bases mapped	5,116,717,913
Total mapped bases in regions (%):	4,098,491,048 (80.1%)
% of bases with 10x coverage	86.4%
Average of reads per base	79.2
Median of reads per base	63.7

Table 2. Exome sequencing statistics.

analysis was performed in order to identify genes with a possible mutation in several samples (10). We prioritized the genes with variants in three or more samples and also some genes with a variant in two samples if both variants showed a high phyloP (basepair conservation) value (phyloP > 5.0). In order to exclude genes from this overlap analysis that frequently show private variants also in the general population, a cohort matched for size and ethnicity presenting with a different phenotype was used. Therefore, the list of genes obtained in the overlap analysis was compared to the genes with variants in a cohort of 14 samples of patients with neural tube defects. By using the same list of private variants, a second analysis was done regarding the most damaging variants, nonsense as well as frameshift. Among those variants, only the ones in genes that do not show similarly damaging variants in the Exome Variant Server (EVS) (11) were selected for further studies, to exclude an enrichment of non-functional or non-essential genes. The availability of parental samples was also considered for validations.

For the third approach, a candidate gene list consisting of 363 genes (supplementary table 2) for LHH was elaborated including genes previously described in humans related to a similar type of disease, as well as genes identified in animal models linked to the same type of malformation. For this approach, all variants with a frequency below 2% in dbSNP as well as the Nijmegen in-house database were considered. Among them, two different lists were generated regarding both possible inheritance pattern (recessive versus dominant). An overlapping analysis with those lists and the candidate gene list was done in order to select for variants in those lists.

Exome sequencing data of a cohort of 50 Spanish individuals was used to exclude variants not described in public databases but frequently found in Spaniards.

For each approach, the prioritized variants were validated by Sanger sequencing and parental samples were analyzed, when available.

RESULTS

De novo analysis

One *de novo* mutation was validated in both patient-parent trio. For trio 1 (case 14), a *de novo* missense mutation in *USP32* was identified in a fetus with LHH. The mutation causes an aminoacid change (N153S) and is located in a highly conserved nucleotide (phyloP 5.1) predicted to be probably damaging by PolyPhen2. A second private variant in the same gene (P417T) was detected in a second fetus with LHH, transposition of great arteries and atrio-ventricular septal defect; parental samples were not available to test for *de novo* occurrence. USP32 is supposed to be involved in the ubiquitin-proteosome pathway (12).

In the second trio (case 13), a *de novo* missense mutation was found in *NCAPD3* in

a highly conserved nucleotide (phyloP 6.3), predicted to be damaging by PolyPhen2.

The mutation causes a substitution of an alanine for a glycine (A718G). Interestingly, two additional highly conserved missense variants were identified in *NCAPD2* in 2 other LHH patients for whom no parental samples were available. *NCAPD2* is a gene encoding a protein from the same complex as NCAPD3, condensin complex II. Condensin complexes I and II play relevant roles in mitotic chromosome assembly and segregation. In each complex, there are 2 invariant structural maintenance of chromosome (SMC) subunits and different sets of non-SMC subunits. NCAPD3 and NCAPD2 are both non-SMC subuni ts of condensin complex II (13). The fetus with a mutation in *NCAPD3* had LHH, as one of the fetus with a variant in *NCAPD2*. The other fetus with a variant in *NCAPD2* presented a more complex malformation, with LHH, transposition of great arteries and atrio-ventricular septal defect.

All results are shown in table 3, and figures 1 and 2.

Genes with variants identified in multiple individual samples

By using the overlap analysis, only three genes were found with variants in 3 or more samples with LHH (*ATP13A5*, *RB1CC1* and *SCAF11*). For two samples with *RB1CC1* variants parental DNA was available and both variants were inherited from the unaffected mother.

Two more genes (*ENTHD1* and *PROX1*) were select ed regarding the fact that two samples with LHH carried a variant in a highly conserved nucleotide (phyloP>5). In *PROX1*, an identical missense variant (P195L) in a nucleotide with a high phyloP (6.4) was identified in two fetuses with LHH (case 12 and case 14). Parental samples were available for one of those two fetuses and the variant was inherited from the mother. The mother presents, as her brother, a Wolff-Parkinson-White syndrome, requiring a surgery in the brother.

Variants in those genes, as well as the inheritance pattern when available, are shown in table 4 and figure 3.

Nonsense and frameshift variants

13 nonsense and frameshift variants in genes without previously reported nonsense variants in EVS (11) were identified in six samples of fetuses with LHH for which parental samples were available. All of them were inherited. Those variants and the inheritance pattern are shown in table 5.

Candidate gene list approach

Several missense variants not described previously in dbSNP, EVS as well as Nijmegen inhouse database or described with a frequency less than 2% were identified in candidate genes for LHH, as it is shown in table 6. Among those variants, two were located in genes previously described related to LHH or a broader spectrum of heart malformations in humans, one in *DTNA* (case 6) and another one in *ACVR2B* (case 8) (14-16). The variant identified in *ACVR2B* was identical to a previously described disease causing point mutation in association to heterotaxy (14, 15). Nonetheless, in the data available in EVS (11) the same variant is described in 60 subjects out of 4,346 in the African American cohort and in 2 individuals out of 8,598 in the European American cohort.

A fetus with LHH carried a missense variant in the gene *DTNA*, in a nucleotide with a basepair conservation score phyloP of 4.9. No parental samples were available. *DTNA,* dystrobrevin gene, was described related to left ventricular non-compaction and additional heart malformations (14).

In table 7, an overview of all variants found is shown per sample.

Sample	Gene name	Type of mutation	cDNA level	Protein level	PhyloP	Grantham score	Inheritance
Case 13	*NCAPD3*	Missense	2153G>C	A718G	6.307	60	De novo
Case 3	*NCAPD2*	Missense	3635A>C	Q1212P	4.636	76	-
Case 6	*NCAPD2*	Missense	3130T>C	F1044L	2.005	22	-
Case 14	*USP32*	Missense	458T>C	N153S	5.098	46	De novo
Case 3	*USP32*	Missense	1249G>T	P417T	2.832	38	-

Table 3. De novo mutations identified in analyzed patient-parent trios. Variants identified in other samples with left heart hypoplasia in the same gene or in genes in the same complex are also shown.

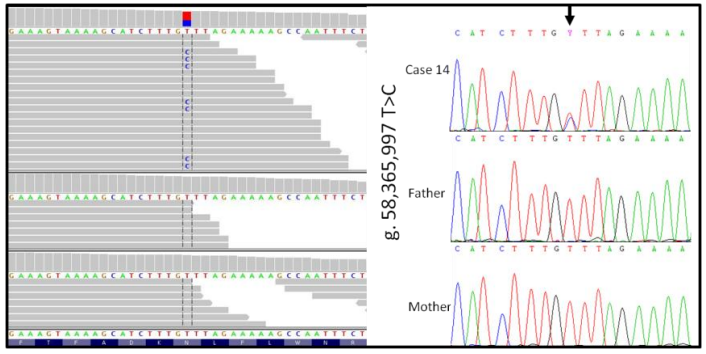


Figure 1: Heterozygous mutation in *USP32* identified by exome sequencing in case 14. On the left side, sequencing reads spanning the mutation in case 14; sequencing reads from the father and the mother are also shown (hg19). The right panel shows the confirmation of the *de novo* mutation by Sanger sequencing.

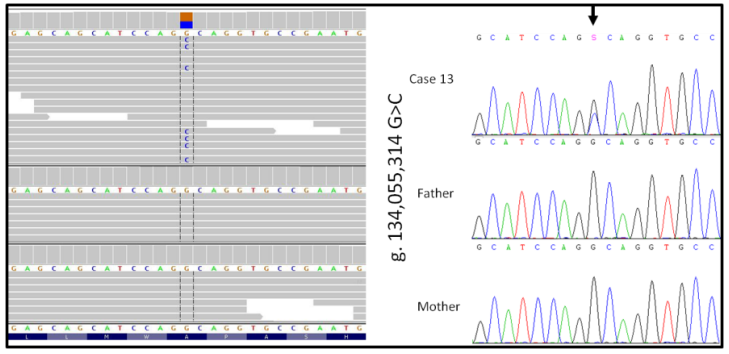


Figure 2: Heterozygous mutation in *NCAPD3* identified by exome sequencing in case 13. The left panel shows sequencing reads from case 13, father and mother (hg19); the right panel shows the confirmation of the *de novo* mutation by Sanger sequencing.

Sample	Gene name	Type of mutation	cDNA level	Protein level	PhyloP	Grantham score	Inheritance
Case 1	*ATP13A5*	Missense	1217G>A	A406V	2.445	64	Paternal
Case 8	*ATP13A5*	Missense	1814T>C	Q605R	4.558	43	-
Case 9	*ATP13A5*	Missense	967T>C	T323A	1.497	58	-
Case 1	*RB1CC1*	Missense	3005T>C	E1002G	3.024	98	Paternal
Case 2	*RB1CC1*	Missense	3977G>A	A1326V	4.624	64	Paternal
Case 3	*RB1CC1*	Missense	2821T>C	M941V	2.376	21	-
Case 2	*SCAF11*	Missense	3500C>T	R1167K	2.913	26	Paternal
Case 5	*SCAF11*	Missense	2356T>C	K786E	2.33	56	Paternal
Case 10	*SCAF11*	Missense	1486C>T	E496K	2.363	56	-
Case 8	*ENTHD1*	Missense	62A>T	V21D	5.119	152	-
Case 9	*ENTHD1*	Missense	206C>T	R69H	6.133	29	-
Case 12	*PROX1*	Missense	584C>T	P195L	6.375	98	-
Case 14	*PROX1*	Missense	584C>T	P195L	6.375	98	Maternal

Table 4. Variants identified by using the overlap analysis.

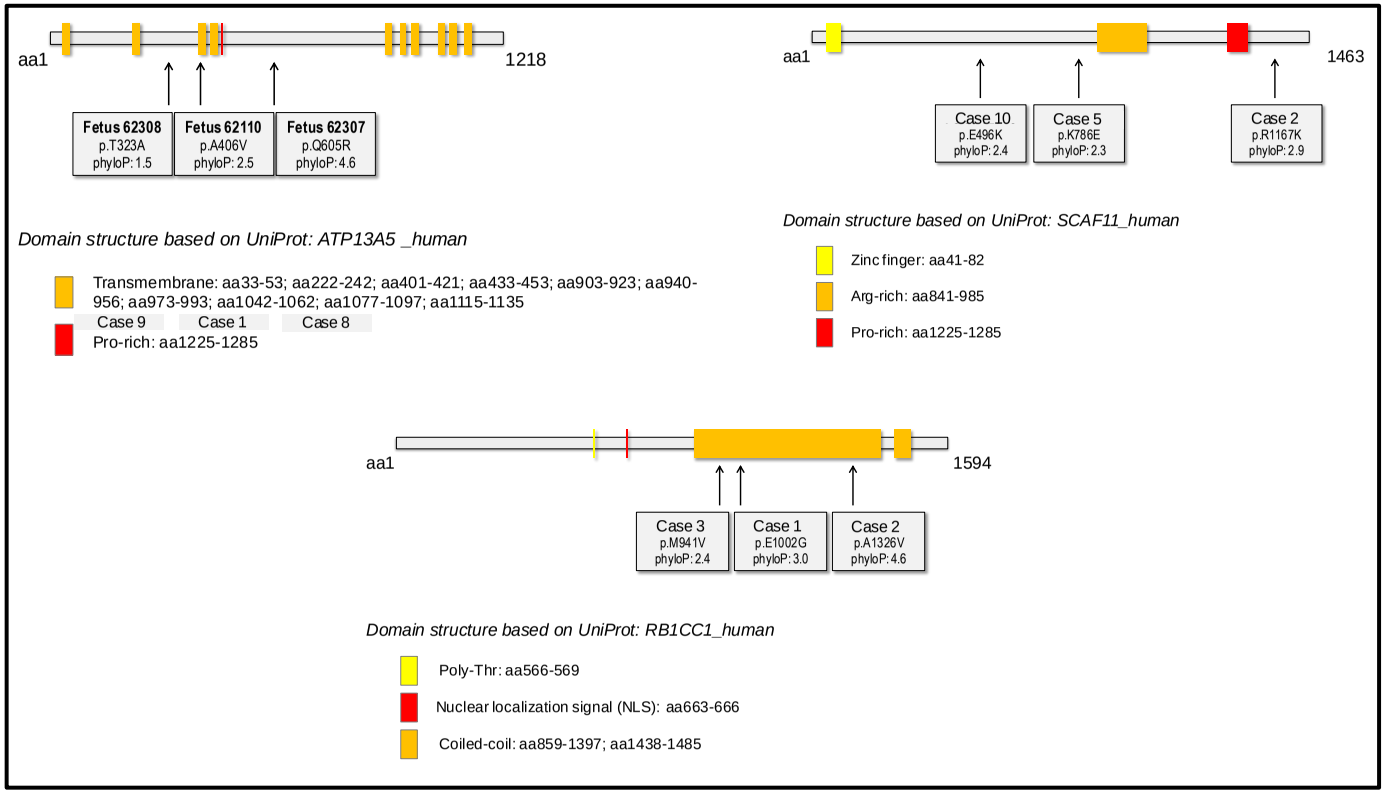


Figure 3: Genes identified in the overlap analysis with mutations in three samples with left heart hypoplasia. Gene structure and location of the mutations identified in *ATP13A5*, *SCAF11* and *RB1CC1* in the cohort of fetuses with left heart hypoplasia.

Sample	Gene name	Type of mutation	cDNA level	Protein level	Inheritance
Case 1	*PER3*	Frameshift	297TCTC>	I99IX	Paternal
Case 1	*IKZF5*	Nonsense	880G>A	Q294*	Paternal
Case 1	*TOMM20L*	Nonsense	252G>A	W84*	Paternal
Case 2	*STRA6*	Frameshift	2018>GG	R634RX R649RX R625RX R671RX R673RX	Maternal
Case 11	*YES1*	Nonsense	197G>C	S66*	Paternal
Case 11	*ANXA6*	Frameshift	1809G>	F571FX F603FX	Paternal
Case 13	*FCRL6*	Nonsense	267T>A	Y89*	Maternal
Case 13	*MS4A14*	Nonsense	1306C>T	Q419* Q436*	Paternal
Case 13	*MOGS*	Nonsense	1421C>T	W368* W474*	Paternal
Case 14	*PHB2*	Nonsense	391G>A	R131*	Paternal
Case 14	*CNKSR3*	Nonsense	573A>T	C191*	Maternal
Case 14	*SCG3*	Frameshift	1238>A	R181RX R413RX	Paternal
Case 14	*TWISTNB*	Frameshift	923T>	K308KX	Maternal

Table 5. Nonsense and frameshift variants identified in genes with none nonsense variants described in exome variant server in the 6 samples of left heart hypoplasia for which parental samples were available.

Sample	Gene name	Type of mutation	cDNA level	Protein level	PhyloP	Grantham score	Inheritance
Case 8	*ACVR2B*	Missense	119G>A	R40H	1.764	29	-
Case 10	*CHD7*	Missense	6577G>A	E2193K	2.326	56	-
Case 6	*DTNA*	Missense	1240C>T	R417W R126W R414W	4.881	101	-
Case 12	*ECE1*	Missense	840G>C	D284E D293E D296E D280E	-0.465	45	-
Case 3	*FBXO25*	Intron deletion	-	-			-
Case 3	*FBXO25*	Missense	451A>G	M218V, M151V	2.313	21	-
Case 7	*FOXO1*	Missense	1643T>C	H548R	4.908	29	-
Case 12	*FOXO1*	Missense	1312G>T	P438T	2.155	38	-
Case 12	*PROX1*	Missense	584C>T	P195L	6.375	98	-
Case 14	*PROX1*	Missense	584C>T	P195L	6.375	98	Maternal
Case 1	*RB1CC1*	Missense	3005T>C	E1002G	3.024	98	Paternal
Case 2	*RB1CC1*	Missense	3977G>A	A1326V	4.624	64	Paternal
Case 3	*RB1CC1*	Missense	2821T>C	M941V	2.376	21	-

Table 6. Variants identified in candidate genes for left heart hypoplasia.

Gene	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12	Case 13	Case 14	Total
ECE1												1			1
PROX1												1		1	2
NCAPD3													1		1
NCAPD2			1			1									2
FOXO1							1					1			2
STRA6		1													1
USP32			1											1	2
DTNA						1									1
ENTHD1								1	1						2
ACVR2B								1							1
FBXO25			2												2
RB1CC1	1	1	1												3
CHD7										1					1
SCAF11		1			1					1					3
ATP13A5	1							1	1						3
Total candidate variants per case	2	3	5	0	1	2	1	3	2	2	0	3	1	2	27

Table 7. Overview of candidate variants (based on tables 5, 6 and 8) identified per fetus with left heart hypoplasia.

DISCUSSION

Two patient-parent trios were analyzed and one *de novo* mutation was identified in each trio. Regarding *USP32*, a second missense variant was identified in another fetus with LHH, although the inheritance could not be established in that case. USP32 has been identified as one of the components of the ubiquitin-proteasome system (UPS). Specifically, it is an ubiquitin-specific protease whose function is the removal of ubiquitin from ubiquitin-conjugates. UPS recognizes specific protein substrates and places polyubiquitin chains on them for subsequent destruction by the proteasome. Therefore, UPS has an essential role in the regulation of protein degradation and its impaired functioning may cause misregulation in several pathways, such as the ones related to proliferation, adaptation to stress, regulation of cell size, and cell death (17). The relevant role of UPS in vascular development and cardiovascular diseases has been described in several articles (18-20), and the role of UPS has been studied in heart development (21-25). So far, *UBR1* and *UBR2*, components of UPS, have been identified in mice model as key effectors of UPS whose impaired functioning leads to a defective heart development (21-23). Moreover, defects in posttranslational small ubiquitin-like modifier (SUMO) conjugation-deconjugation pathway have also been shown in association to cardiac defects, because of the impairment in cardiomyocyte proliferation (24, 25). Accordingly to this data, UPS may play an important role in heart development.

Strikingly the same fetus who carried a variant with unknown inheritance in USP32 presented with additional variants in genes of the UPS pathway. Particularly, two variants were found in the gene *FBXO25*, a 6-bp deletion in intron 6 near the exon boundary and a missense variant in a nucleotide with a phyloP of 2.1 not described previously in dbSNP or in the Nijmegen inhouse database. Though no parental DNA was available, allele specific PCR showed that these two variants affected different alleles each. Jang *et al.* described FBXO25 as a cardiac-specific protein which acts as an ubiquitin E3 ligase for cardiac transcription factors such as NKX2-5, ISL1 and HAND1 (26). They revealed a high expression of *FBXO25* in cardiomyocytes and a higher expression level in fetal heart than in adult heart (26). Mutations in *NKX2-5* (27-30) have been described in patients with CHD, such as septal defects and tetralogy of Fallot. Missense mutations in *NKX2-5* have also been described in two unrelated patients with LHH (27, 31). Regarding *HAND1*, mutations in cardiac tissue of patients with septation defects and heart hypoplasia were found (32, 33). Considering the role of *NKX2-5* and *HAND1* in heart defects it seems plausible that mutations in *FBXO25* lead to heart defects by the impaired regulation of those transcription factors.

A *de novo* mutation in fetus 13 was identified in *NCAPD3*, a non-SMC subunit of condensin complex II. In addition, two missense variants were also detected in *NCAPD2* in two other patients, another non-SMC subunit of the same complex. *NCAPD3* is a gene located in 11q25 and it is included in the region deleted in Jacobsen syndrome. This syndrome is a rare chromosomal disorder caused by deletions in distal 11q. The deletion size ranges from ~7 to 20 Mb, with the proximal breakpoint within or telomeric to 11q23.3 and the deletion extending usually to the telomere (5).

To date, over 200 cases of Jacobsen syndrome have been reported. The most common clinical features include pre- and postnatal growth retardation, psychomotor retardation, characteristic facial dysmorphism, abnormal platelet function, thrombocytopenia and congenital malformations. One of the most common congenital malformations is CHD, which occurs in 56% of cases. The most frequent heart defects (2/3 of patients that have CHD) are ventricular septal defects or left heart obstructive malformations including abnormalities of aortic or mitral valves,

coarctation of the aorta or hypoplastic left heart. So far, a critical region for CHD in Jacobsen syndrome has been identified and contains 40 annotated genes, including *NCAPD3* (5, 34, 35).

Furthermore, Ye *et al.* reported a patient with a paracentric inversion in distal 11q who had hypoplastic left heart and congenital thrombocytopenia. The distal breakpoint mapped to *JAM3*, a gene previously identified as a candidate gene for causing LHH in Jacobsen syndrome patients and included in the critical region for CHD. Nevertheless, the mouse model for *Jam3* revealed no cardiac abnormalities. *NCAPD3* is the adjacent gene to *JAM3*; therefore, *NCAPD3* emerges as a candidate gene to explain LHH in Jacobsen syndrome and in the fetus studied. (36)

The identification of *de novo* mutations itself has already led to the identification of novel candidate genes in several diseases including defined syndromes (37), or similarly heterogeneous disorders like intellectual disability (9), schizophrenia (38) and autism (39, 40). The amount of identified non-synonymous *de novo* mutations in our study is in line with the expected human *de novo* mutation rate and other findings in exome sequencing studies. The combination of *de novo* mutation occurrence, mutation severity (nucleotide conservation) and potential gene function, and importantly additional possible mutation in the same or similar genes in our cohort make both genes *USP32* and *NCAPD3* new candidate genes for LHH, either causing the phenotype itself, or are at least involved in the pathogenesis of this phenotype. The ultimate proof can only be delivered by identifying additional *de novo* mutations in fetuses with similar phenotype, such a study requires a much larger follow-up cohort.

Two private variants in genes previously described as being related to CHD were identified. A missense variant in *DTNA* (R417W), the dystrobrevin gene, was detected in a fetus with LHH. A mutation in the same gene was previously described in a family with left ventricular non-compaction and additional heart malformations co-segregating with the disease (14). In that family, there were 6 relatives affected and several of them also presented additional heart malformations. 2 of them had ventricular septal defects, one patent ductus arteriosus and 2 LHH. The finding of a private missense variant in a highly conserved nucleotide in a fetus with LHH reinforces the role of *DTNA* in heart development and CHD

A variant in *ACVR2B* was identified in one of the samples studied. This gene had been described previously in association with heterotaxy defects and additional heart malformations (15-16), but not linked to LHH. *ACVR2B* had been related to heterotaxy and additional heart malformations because of the phenotype in the animal model (14) -including atrial and ventricular septal defects and right-sided morphology malformations- and the identification of mutations in patients with this type of congenital defect (15, 16). Variants in this gene were found to be inherited from a healthy parent in most cases (15, 16). The identification of one of the mutations described before (R40H) in one of the fetus with LHH without heterotaxy defects in the studied cohort might indicate a broader spectrum of malformations in relation to mutations in *ACVR2B*. Nevertheless, this variant has shown to be quite frequent in African American population, considering the EVS (11) which describes the same variant in 60 subjects out of 4,346 in the African American cohort and in 2 individuals out of 8,598 in the European American cohort. In the light of the latter, it seems more likely that this variant represents at most a predisposing factor. This is one example that one has to carefully consider previous literature stating disease causality for a given genetic variant, but this also shows that population frequency information can be valuable to distinguish likely from unlikely disease candidate variants.

By using the unbiased approach and the candidate gene approach, other potentially related genes were identified. *PROX1* and *RB1CC1*, as examples, emerged as two candidates for LHH.

Regarding *PROX1*, exactly the same variant – not present in dbSNP, Nijmegen inhouse database, EVS (11) and 50 Spanish exomes- was identified in two fetuses with LHH. In one of them, the variant was inherited from the mother, affected by Wolff-Parkinson-White syndrome. Wolff-Parkinson-White syndrome is a disorder of the conduction system of the heart and it is caused by the presence of an abnormal accessory electrical conduction pathway between atria and ventricles, causing supraventricular tachycardia. An important proportion of patients remain asymptomatic although they have a higher risk of sudden death. Mutations in *PRKAG2* have been reported related to Wolff-Parkinson-White syndrome (41); mutations in this gene were discarded in the mother of the fetus analyzed. Parental samples for the second case with a variant in *PROX1* were not available. The homozygous knockout mouse model for *Prox1* dies between E14.5 and E15 because of the reduction in heart size (up to 50%) and myocardial disarrays (42, 43). *Prox1* heterozygous mice present heart reduced in size (by an average of 30%) as compared with those of wild-type littermate controls (42) and hypoplastic ventricular walls, loss of muscle striation, a disorganized interventricular septum and abnormally persistent muscle surrounding the aorta. Interestingly, a patient with LHH and a *de novo* balanced rearrangement interrupting the regulatory region of *PROX1* was described by Gill *et al.* (44). Published data concerning *PROX1*, as well as the variant identified in two fetuses with LHH, positions this gene as one of the best novel candidates regarding LHH.

Finally, three missense variants were identified in *RB1CC1* in three different samples of fetuses with LHH. Homozygous knockout mice for *Rb1cc1* died because of the lost of the normal trabecular and external compact myocytes and a significantly thinner wall which contained fewer cells when compared with the wild-type littermates (45). Considering the high detection rate of variants -not described previously- in *RB1CC1* in our cohort of fetuses with LHH (3/14) and data regarding mouse model, a role of that gene in heart development and heart defects might be considered.

The results presented in this article highlight the heterogeneity of LHH and the difficulties to face regarding the interpretation of the results. Nevertheless, the work done point out several interesting genes related to this severe heart malformation, such as *USP32*, *NCAPD3* or *PROX1*. At least in some cases, more than one variant in a potentially related gene has been identified perhaps as contributors in an oligogenic model for the disease. The study of larger cohorts as well as functional studies regarding those variants identified will provide valuable data in order to clearly establish the role of these genes in heart development and in heart disease. In a broader sense this study emphasizes efforts to expand the clinical spectrum of several disorder by even more severe phenotypes detected in aborted fetuses, miscarriages and stillborns (46).

REFERENCES

(1) Hinton RB Jr, Martin LJ, Tabangin ME, Mazwi ML, Cripe LH, Benson DW. (2007). Hypoplastic left heart syndrome is heritable. J Am Coll Cardiol 50(16):1590-5.

(2) van Egmond H, Orye E, Praet M, Coppens M, Devloo-Blancquaert A. (1988). Hypoplastic left heart syndrome and 45X karyotype. Br Heart J 60(1):69-71.

(3) Morris CA, Sutton TM, Olson TM. (2006). Congenital heart defects in siblings with partial aneuploidy of chromosome 2p. Cardiovasc Pathol 15(1):33-5.

(4) Glauser, T. A., Zackai, E., Weinberg, P., Clancy, R. (1989). Holt-Oram syndrome associated with the hypoplastic left heart syndrome. Clin Genet 36: 69-72.

(5) Mattina T, Perrotta CS, Grossfeld P. (2009). Jacobsen syndrome. Orphanet J Rare Dis 4:9.

(6) Stankiewicz P, Sen P, Bhatt SS, Storer M, Xia Z, Bejjani BA, Ou Z, Wiszniewska J, Driscoll DJ, Maisenbacher MK, Bolivar J, Bauer M, Zackai EH, McDonald-McGinn D, Nowaczyk MM, Murray M, Hustead V, Mascotti K, Schultz R, Hallam L, McRae D, Nicholson AG, Newbury R, Durham-O'Donnell J, Knight G, Kini U, Shaikh TH, Martin V, Tyreman M, Simonic I, Willatt L, Paterson J, Mehta S, Rajan D, Fitzgerald T, Gribble S, Prigmore E, Patel A, Shaffer LG, Carter NP, Cheung SW, Langston C, Shaw-Smith C. (2009). Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. Am J Hum Genet 84(6):780-91.

(7) McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, Lam W, Leal SM, Kaplan N, Schliekelman P, Towbin JA, Belmont JW. (2005). Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: Segregation, multiplex relative risk, and heritability. Am J Med Genet A 134A(2):180-6.

(8) Serra-Juhé C, Rodríguez-Santiago B, Cuscó I, Vendrell T, Camats N, Torán N, Pérez-Jurado LA. Contribution of Rare Copy Number Variants to Isolated Human Malformations. In press.

(9) Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. (2010). A de novo paradigm for mental retardation. Nat Genet 42(12):1109-12.

(10) Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, Steehouwer M, de Vries P, de Reuver R, Wieskamp N, Mortier G, Devriendt K, Amorim MZ, Revencu N, Kidd A, Barbosa M, Turner A, Smith J, Oley C, Henderson A, Hayes IM, Thompson EM, Brunner HG, de Vries BB, Veltman JA. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat Genet 42(6):483-5.

(11) Exome Variant Server, NHLBI Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [date June 2012 accessed].

(12) Akhavantabasi S, Akman HB, Sapmaz A, Keller J, Petty EM, Erson AE. (2010). USP32 is an active, membrane-bound ubiquitin protease overexpressed in breast cancers. Mamm Genome 21(7-8):388-97.

(13) Ono T, Losada A, Hirano M, Myers MP, Neuwald AF, Hirano T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell 115: 109-121.

(14) Ichida F, Tsubata S, Bowles KR, Haneda N, Uese K, Miyawaki T, Dreyer WJ, Messina J, Li H, Bowles NE, Towbin JA. (2001). Novel gene mutations in patients with left ventricular noncompaction or Barth syndrome. Circulation 103: 1256-1263.

(15) Kosaki, R., Gebbia, M., Kosaki, K., Lewin, M., Bowers, P., Towbin, J. A., Casey, B. (1999). Left-right axis malformations associated with mutations in ACVR2B, the

gene for human activin receptor type IIB. Am J Med Genet 82: 70-76.

(16) Ma L, Selamet Tierney ES, Lee T, Lanzano P, Chung WK. (2012). Mutations in ZIC3 and ACVR2B are a common cause of heterotaxy and associated cardiovascular anomalies. Cardiol Young 22(2):194-201.

(17) Willis MS, Townley-Tilson WH, Kang EY, Homeister JW, Patterson C. (2010). Sent to destroy: the ubiquitin proteasome system regulates cell signaling and protein quality control in cardiovascular development and disease. Circ Res 106(3):463-78.

(18) Tetzlaff MT, Yu W, Li M, Zhang P, Finegold M, Mahon K, Harper JW, Schwartz RJ, Elledge SJ. (2004). Defective cardiovascular development and elevated cyclin E and Notch proteins in mice lacking the Fbw7 F-box protein. Proc Natl Acad Sci U S A 101:3338–3345.

(19) Arai T, Kasper JS, Skaar JR, Ali SH, Takahashi C, DeCaprio JA. (2003). Targeted disruption of p185/Cul7 gene results in abnormal vascular morphogenesis. Proc Natl Acad Sci U S A 100:9855–9860.

(20) Marfella R, D'Amico M, Di Filippo C, Baldi A, Siniscalchi M, Sasso FC, Portoghese M, Carbonara O, Crescenzi B, Sangiuolo P, Nicoletti GF, Rossiello R, Ferraraccio F, Cacciapuoti F, Verza M, Coppola L, Rossi F, Paolisso G. (2006). Increased activity of the ubiquitin-proteasome system in patients with symptomatic carotid disease is associated with enhanced inflammation and may destabilize the atherosclerotic plaque: effects of rosiglitazone treatment. J Am Coll Cardiol 47: 2444–2455.

(21) An JY, Seo JW, Tasaki T, Lee MJ, Varshavsky A, Kwon YT. (2006). Impaired neurogenesis and cardiovascular development in mice lacking the E3 ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway. Proc Natl Acad Sci U S A 103:6212–6217.

(22) Kwon YT, Xia Z, Davydov IV, Lecker SH, Varshavsky A. (2001). Construction and analysis of mouse strains lacking the ubiquitin ligase UBR1 (E3alpha) of the N-end rule pathway. Mol Cell Biol 21:8007–8021.

(23) Zenker M, Mayerle J, Lerch MM, Tagariello A, Zerres K, Durie PR, Beier M, Hulskamp G, Guzman C, Rehder H, Beemer FA, Hamel B, Vanlieferinghen P, Gershoni-Baruch R, Vieira MW, Dumic M, Auslender R, Gil-da-Silva-Lopes VL, Steinlicht S, Rauh M, Shalev SA, Thiel C, Ekici AB, Winterpacht A, Kwon YT, Varshavsky A, Reis A. (2005). Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). Nat Genet 37:1345–1350.

(24) Wang J, Chen L, Wen S, Zhu H, Yu W, Moskowitz IP, Shaw GM, Finnell RH, Schwartz RJ. (2011). Defective sumoylation pathway directs congenital heart disease. Birth Defects Res A Clin Mol Teratol 91(6):468-76.

(25) Kim EY, Chen L, Ma Y, Yu W, Chang J, Moskowitz IP, Wang J. (2011). Expression of sumoylation deficient Nkx2.5 mutant in Nkx2.5 haploinsufficient mice leads to congenital heart defects. PLoS One 6(6).

(26) Jang JW, Lee WY, Lee JH, Moon SH, Kim CH, Chung HM. (2011). A novel Fbxo25 acts as an E3 ligase for destructing cardiac specific transcription factors. Biochem Biophys Res Commun 410(2):183-8.

(27) Stallmeyer, B., Fenge, H., Nowak-Gottl, U., Schulze-Bahr, E. (2010). Mutational spectrum in the cardiac transcription factor gene NKX2.5 (CSX) associated with congenital heart disease. Clin Genet 78: 533-540.

(28) Wang, J., Xin, Y.-F., Liu, X.-Y., Liu, Z.-M., Wang, X.-Z., Yang, Y.-Q. (2011). A novel NKX2-5 mutation in familial ventricular septal defect. Int J Mol Med 27: 369-375.

(29) Watanabe, Y., Benson, D. W., Yano, S., Akagi, T., Yoshino, M., Murray, J. C. (2002).

Two novel frameshift mutations in NKX2.5 result in novel features including visceral inversus and sinus venosus type ASD. J Med Genet 39: 807-811.

(30) Goldmuntz, E., Geiger, E., Benson, D. W. (2001). NKX2.5 mutations in patients with tetralogy of Fallot. Circulation 104:2565-2568.

(31) McElhinney, D. B., Geiger, E., Blinder, J., Benson, D. W., Goldmuntz, E. (2003). NKX2.5 mutations in patients with congenital heart disease. J Am Coll Cardiol 42: 1650-1655.

(32) Reamon-Buettner, S. M., Ciribilli, Y., Traverso, I., Kuhls, B., Inga, A., Borlak, J. (2009). A functional genetic study identifies HAND1 mutations in septation defects of the human heart. Hum Molec Genet 18:3567-3578.

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

(34) Leegte B, Kerstjens-Frederikse WS, Deelstra K, Begeer JH, van Essen AJ. (1999). 11q- syndrome: three cases and a review of the literature. Genet Couns 10(3):305-13.

(35) Grossfeld PD, Mattina T, Lai Z, Favier R, Jones KL, Cotter F, Jones C. (2004). The 11q terminal deletion disorder: a prospective study of 110 cases. Am J Med Genet 129A:51-61.

(36) Ye M, Hamzeh R, Geddis A, Varki N, Perryman MB, Grossfeld P. (2009). Deletion of JAM-C, a candidate gene for heart defects in Jacobsen syndrome, results in a normal cardiac phenotype in mice. Am J Med Genet A 149A(7):1438-43.

(37) Rivière JB, van Bon BW, Hoischen A, Kholmanskikh SS, O'Roak BJ, Gilissen C, Gijsen S, Sullivan CT, Christian SL, Abdul-Rahman OA, Atkin JF, Chassaing N, Drouin-Garraud V, Fry AE, Fryns JP, Gripp KW, Kempers M, Kleefstra T, Mancini GM, Nowaczyk MJ, van Ravenswaaij-Arts CM, Roscioli T, Marble M, Rosenfeld JA, Siu VM, de Vries BB, Shendure J, Verloes A, Veltman JA, Brunner HG, Ross ME, Pilz DT, Dobyns WB. (2012). De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. Nat Genet 26;44(4):440-4, S1-2.

(38) Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, Gogos JA, Karayiorgou M. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat Genet 7;43(9):864-8.

(39) O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, Eichler EE. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat Genet 43(6):585-9.

(40) Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, Polak P, Yoon S, Maguire J, Crawford EL, Campbell NG, Geller ET, Valladares O, Schafer C, Liu H, Zhao T, Cai G, Lihm J, Dannenfelser R, Jabado O, Peralta Z, Nagaswamy U, Muzny D, Reid JG, Newsham I, Wu Y, Lewis L, Han Y, Voight BF, Lim E, Rossin E, Kirby A, Flannick J, Fromer M, Shakir K, Fennell T, Garimella K, Banks E, Poplin R, Gabriel S, DePristo M, Wimbish JR, Boone BE, Levy SE, Betancur C, Sunyaev S, Boerwinkle E, Buxbaum JD, Cook EH Jr, Devlin B, Gibbs RA, Roeder K, Schellenberg GD, Sutcliffe JS, Daly MJ. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature. 2012 Apr 4;485(7397):242-5. doi: 10.1038/nature11011.

(41) Gollob, M. H., Seger, J. J., Gollob, T. N., Tapscott, T., Gonzales, O., Bachinski, L., Roberts, R. (2001). Novel PRKAG2 mutation responsible for the genetic syndrome of ventricular preexcitation and conduction system disease with childhood

onset and absence of cardiac hypertrophy. Circulation 104: 3030-3033.

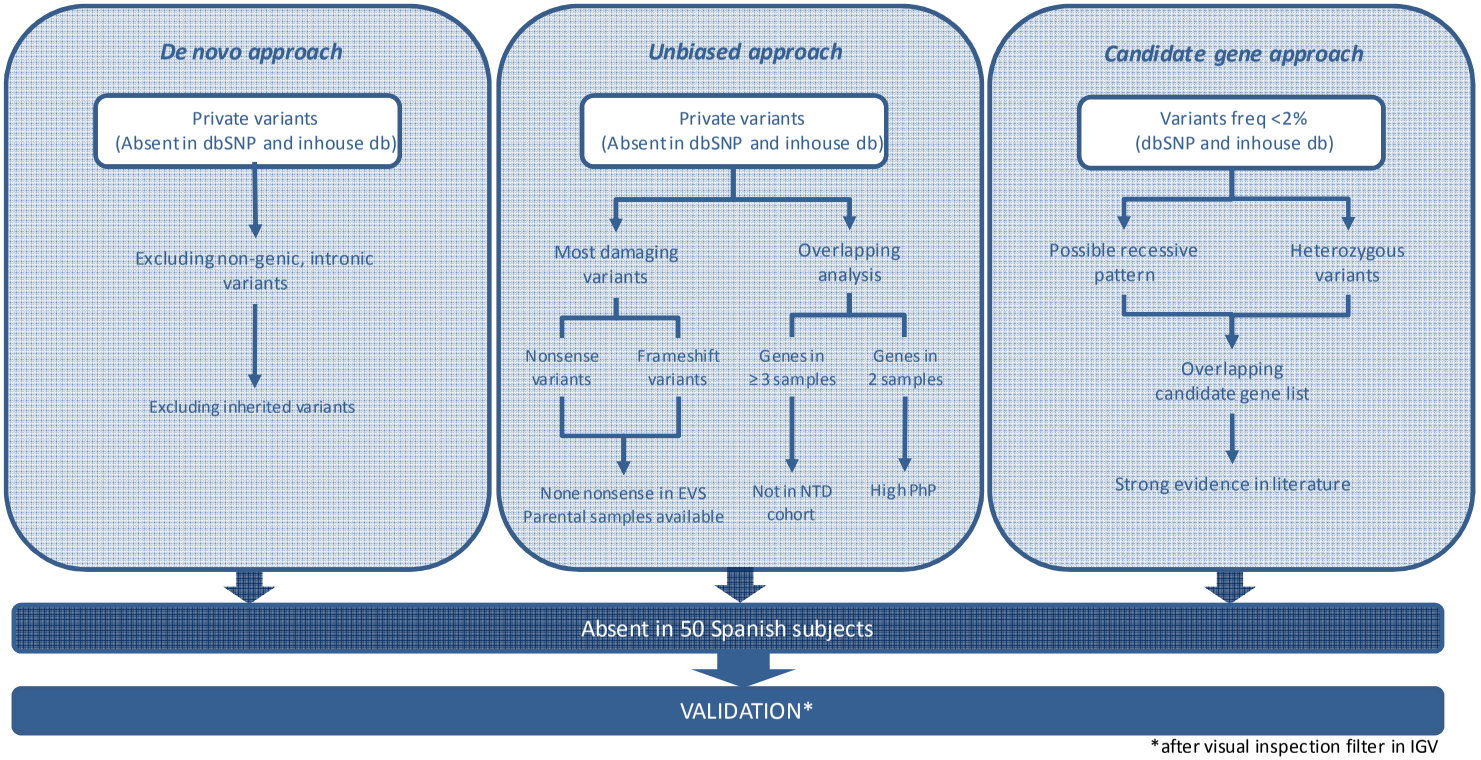
(42) Wigle, J. T., Chowdhury, K., Gruss, P. and Oliver, G. (1999). Prox1 function is crucial for mouse lens-fibre elongation. Nat. Genet. 21,318 -322.

(43) Risebro CA, Searles RG, Melville AA, Ehler E, Jina N, Shah S, Pallas J, Hubank M, Dillard M, Harvey NL, Schwartz RJ, Chien KR, Oliver G, Riley PR. (2009). Prox1 maintains muscle structure and growth in the developing heart. Development 136(3):495-505.

(44) Gill HK, Parsons SR, Spalluto C, Davies AF, Knorz VJ, Burlinson CE, Ng BL, Carter NP, Ogilvie CM, Wilson DI, Roberts RG. (2009). Separation of the PROX1 gene from upstream conserved elements in a complex inversion/translocation patient with hypoplastic left heart. Eur J Hum Genet 17(11):1423-31.

(45) Gan B, Peng X, Nagy T, Alcaraz A, Gu H, Guan JL. (2006). Role of FIP200 in cardiac and liver development and its regulation of TNF-alpha and TSC-mTOR signaling pathways. J Cell Biol 175: 121-133.

(46) Opitz JM. (2012). 2011 William Allan Award: development and evolution. Am J Hum Genet 90(3):392-404.



Supplementary figure 1: Analysis workflow to analyze exome sequencing data in left heart hypoplasia. Overview of the three strategies used to identified variants potentially related to left heart hypoplasia in the cohort analyzed.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12	Case 13	Case 14	Average	St dev
% of bases with less than 1 read	4.7%	4.7%	5.2%	6.0%	6.3%	6.3%	5.6%	5.6%	5.3%	5.3%	4.6%	4.7%	5.6%	5.3%	5.3%	0.58
% of bases with less than 5 read	8.9%	9.1%	9.7%	11.3%	11.6%	11.9%	10.3%	9.2%	9.9%	9.8%	8.6%	8.8%	10.5%	9.6%	9.9%	1.06
% of bases with less than 10 read	12.1%	12.7%	13.2%	15.9%	16.0%	16.4%	13.9%	12.5%	13.6%	13.5%	11.7%	12.0%	14.3%	12.9%	13.6%	1.53
% of bases with less than 20 read	17.5%	18.9%	19.2%	23.8%	23.3%	24.1%	20.0%	18.0%	19.9%	19.8%	16.9%	17.5%	20.2%	18.7%	19.9%	2.35
Average of reads per base	91.8	78.4	83.5	61.4	62.9	61.1	82.0	88.1	80.0	76.4	91.1	84.6	83.8	84.2	79.2	10.37
Median of reads per base	73.5	62.1	66.8	49.1	51.3	49.4	66.5	71.8	64.2	61.8	72.6	67.8	67.5	67.6	63.7	8.24
Total mapped bases in regions (%):	82.4%	84.3%	80.3%	76.8%	77.8%	77.8%	76.6%	80.3%	75.8%	79.6%	84.0%	82.7%	80.4%	82.6%	80.1%	2.83
Total mapped bases near regions (%):	7.9%	7.0%	8.6%	9.1%	9.4%	9.4%	9.4%	8.7%	8.3%	9.0%	7.2%	7.8%	8.1%	7.4%	8.4%	0.9
Total mapped bases outside regions (%)	9.8%	8.7%	11.2%	14.1%	12.8%	12.7%	14.0%	10.9%	15.9%	11.5%	8.8%	9.5%	11.5%	10.0%	11.5%	2.1
Total mapped bases (Mb):	5,767.2	4,742.7	5,409.0	4,220.4	4,244.1	4,126.1	5,600.2	5,664.4	5,460.4	5,005.8	5,521.9	5,289.5	5,385.6	5,196.2	5,116.7	563.7

Supplementary table 1. Overview of exome-sequencing performance per sample.

Candidate gene list for left heart hypoplasia												
ACP6	*BMPR1A*	*COL6A1*	*EDN1*	*FOG2*	*GNAQ*	*INVS*	*MAP3K7*	*NCKAP1*	*PCSK6*	*RARB*	*SMAD6*	*TFAP2B*
ACTC1	*BMPR2*	*COL6A2*	*EDNRA*	*FOLBP1*	*GPC3*	*IQGAP1*	*MAP3K7IP1*	*NCOA6*	*PDGFA*	*RARG*	*SMAD7*	*TGFB2*
ACTIVIN	*BRAF*	*CREBBP*	*EFNB2*	*FOLR1*	*GSTT1*	*IRX4*	*MDM2*	*NF1*	*PDGFC*	*RBL2*	*SMC1A*	*TGFB3*
ACVR1	*BTC*	*CRELD1*	*EGFR*	*FOXA2*	*GTF2IRD1*	*ISL1*	*MED1*	*NFAT*	*PDGFRA*	*RBM24*	*SMO*	*TGFBR2*
ACVR2B	*C17ORF40*	*CRKL*	*EGLN1*	*FOXC1*	*GYS1*	*JAG1*	*MED13L*	*NFATC1*	*PDLIM3*	*RBP4*	*SMYD1*	*TGFBR3*
ACVRL1	*CALR*	*CST3*	*ELN*	*FOXC2*	*HAND1*	*JARID2*	*MEF2C*	*NKX2-5*	*PDPK1*	*RCAN1*	*SOX11*	*TH*
ADAM12	*CARP*	*CTBP2*	*ENG*	*FOXG1*	*HAND2*	*JAZF1*	*MEK1*	*NKX2-6*	*PITX2*	*RERE*	*SOX4*	*THBS1*
ADAM17	*CAV1*	*CX40*	*EP300*	*FOXH1*	*HAS2*	*JMJD6*	*MEK2*	*NKX2-7*	*PKD1*	*RFX3*	*SOX9*	*THRAP4*
ADAM19	*CAV3*	*CX43*	*EPHB4*	*FOXJ1*	*HBEGF*	*JNR*	*MEN1*	*NKX6-2*	*PKD2*	*ROR1*	*SRF*	*TLL1*
ADAM9	*CBP*	*CXADR*	*ERBB2*	*FOXM1*	*HDAC5*	*JUN*	*MESP1*	*NODAL*	*PLCE1*	*ROR2*	*SSR1*	*TMOD*
ADRBK1	*CCND1*	*CXCL12*	*ERBB3*	*FOXO1*	*HDAC7A*	*JUP*	*MESP2*	*NOG*	*PLXND1*	*RTTN*	*SUFU*	*TWIST1*
AGTR1	*CCND2*	*CXCR4*	*ERBB4*	*FOXP1*	*HDAC9*	*KDR*	*MGAT1*	*NOS3*	*PNPLA2*	*RUNX1*	*SUMO1*	*TXNRD2*
AGTR1B	*CCND3*	*CYP26A1*	*ETS-1*	*FOXP4*	*HEY1*	*KIF3A*	*MGRN1*	*NOTCH1*	*POR*	*RXRA*	*SUMO2*	*UGDH*
ALDH1A2	*CCNE1*	*DAND5*	*EVI1*	*FREM2*	*HEY2*	*KIF3B*	*MIB1*	*NOTCH2*	*POSTN*	*SALL1*	*SUMO3*	*VCAM1*
ALK2	*CCNE2*	*DISP1*	*F2R*	*FURIN*	*HGS*	*KL*	*MIXL1*	*NPPA*	*PPARBP*	*SALL4*	*SUZ12*	*VCAN*
ALK3	*CDH2*	*DLL1*	*FBN1*	*FXN*	*HHEX*	*KRIT1*	*MKL1*	*NR2F2*	*PPARG*	*SEMA3C*	*TAL1*	*VCL*
APC	*CDH7*	*DLL4*	*FBXO25*	*FZD2*	*HIF1A*	*KRT19*	*MKL2*	*NRG1*	*PRKAB2*	*SGSH*	*TBRI*	*VEGF*
APOE	*CDK1*	*DNAH11*	*FGF10*	*GAB1*	*HIRA*	*LATS2*	*MOSPD3*	*NRP1*	*PRKAR1A*	*SHC1*	*TBRIII*	*VEGFA*
ATE1	*CDK2*	*DNAHC11*	*FGF15*	*GAS1*	*HOXA1*	*LEFTA*	*MSX1*	*NRP2*	*PRKCA*	*SHH*	*TBX1*	*VEZF1*
ATP2A2	*CDK4*	*DNMT3B*	*FGF2*	*GATA3*	*HOXA3*	*LEFTY2*	*MSX2*	*NSD1*	*PROSIT240*	*SHP2*	*TBX18*	*WASF2*
ATRX	*CENTA2*	*DTNA*	*FGF8*	*GATA4*	*HRAS*	*LOX*	*MURF1*	*NT3*	*PROX1*	*SIRT1*	*TBX2*	*WHSC2*
AXIN1	*CFC1*	*DVL1*	*FGF9*	*GATA6*	*HSPG2*	*LRP6*	*MYCN*	*NTF3*	*PRRX1*	*SLC6A4*	*TBX20*	*WNT3A*
BCL9	*CHD1L*	*DVL2*	*FGFR1*	*GBX2*	*HTR1B*	*LRRC10*	*MYH10*	*NTRK3*	*PRRX2*	*SLC8A1*	*TBX3*	*WNT5A*
BCOR	*CHD7*	*DYNC2H1*	*FGFR2*	*GDF1*	*ID2*	*LY6E*	*MYH11*	*OFD1*	*PSEN1*	*SMAD1*	*TBX4*	*WRN*
BMP1	*CHMP5*	*DYNC2LI1*	*FKBP1A*	*GJA1*	*IFT172*	*MAB21L2*	*MYH6*	*OSR1*	*PSEN2*	*SMAD2*	*TBX5*	*ZFPM1*
BMP10	*CHRD*	*E2F1*	*FKBP1B*	*GJA5*	*IFT57*	*MAFBX*	*MYL2*	*OVOL2*	*PTEN*	*SMAD3*	*TCEB3*	*ZFPM2*
BMP2	*CITED2*	*ECE1*	*FLNA*	*GJA7*	*IFT88*	*MAP2K2*	*MYL7*	*PAX3*	*PTPN11*	*SMAD4*	*TCFAP2A*	*ZIC3*
BMP4	*COL2A1*	*ECE2*	*FMO5*	*GNA11*	*IGF1*	*MAP2K5*	*MYST3*	*PBRM1*	*RARA*	*SMAD5*	*TEF1*	

Supplementary table 2. List of the genes included in the candidate gene list for left heart hypoplasia. Genes described in animal models related to congenital heart defects, in patients with this type of malformations or genes identified as playing an important role in heart development have been considered.

CHAPTER 6

Exome sequencing in neural tube defects

Clara Serra-Juhé, Christian Gilissen, Teresa Vendrell, Marisol del Rosario, Marloes Steehouwer, Núria Toran, Joris A. Veltman, Luis Pérez-Jurado*, Alexander Hoischen*

* shared last author

In preparation

Neural tube defects are one of the most common congenital malformations affecting 1/1,000 newborns. Despite its high incidence, the mechanisms underlying the disorder are poorly understood. It is well-established that folate intake during pregnancy reduces its incidence although the basis of this decrease is not known. Several studies in genes in the folate pathway have been done, as well as studies in animal models. Studies of folate pathway have reported some associations between SNPS in genes within the pathway and the incidence of NTDs. Using animal models several genes have been identified as causing NTDs and, among them, some pathways as planar cell polarity pathway or glycine cleavage system have been described. Recently, variants not present in control population in genes from those pathways have been reported in cohorts of patients with NTD.

With the aim of improving the knowledge about described genes and discovering new candidates, a cohort of 14 fetuses with a severe form of neural tube defect (mainly myelomenginocele plus Arnold-Chiari malformation) was sequenced.

Exome sequencing in neural tube defects

Clara Serra-Juhé, Christian Gilissen, Teresa Vendrell, Marloes Steehouwer, Marisol del Rosario, Núria Toran, Joris A. Veltman, Luis Pérez-Jurado*, Alexander Hoischen*

* shared last author

Neural tube defects (NTDs) are the second most common congenital malformation with a frequency of 1/1,000 in live births. Recently, many advances in the understanding of NTDs have been described, the identification of genes and pathways related to the disorder underscore a strong genetic component. Nevertheless, the etiology of the majority of cases of NTDs remains unclear.

We selected 14 fetuses with severe forms of NTD from the fetal tissue bank of Vall d'Hebron hospital (Barcelona). All samples had been studied previously by G-banding karyotype and chromosomal microarray analysis (CMA) with no chromosomal aberration or relevant copy number variants (CNVs) identified. Exome sequencing was performed for 2 fetus-parent trios as well as 12 individual fetal samples.

No *de novo* mutations were identified in any of the analyzed trios. However, several potentially pathogenic variants affecting protein function were detected in genes previously related to NTDs in humans, such as *PRICKLE1*, *DACT1*, *GLDC1* or *HSPG2*. Additionally, disruptive variants in a number of novel NTD candidate genes were identified (*FZD1*, *PTK7*, *TSC2* and *PAMR1)*. Interestingly, a global enrichment of variants in genes of the planar cell polarity pathway was identified in this cohort. Also, evidence for a possible digenic inheritance model involving *FZD1* and *DACT1* genes was observed in a fetus with myelomeningocele and Arnold-Chiari malformation.

Our data suggest a significant contribution of point mutations in NTDs, and further support a relevant role of the planar cell polarity pathway in the etiology of the malformation.

INTRODUCTION

Neural tube defects (NTD) are one of the most common birth defects which occur very early in human development as a result of the failure in neural tube closure. The incidence of NTD is approximately one in 1,000 live births (1, 2). There are several types of neural tube defects depending on the placement and the presence or lack of cover. Open neural tube defects, which are more common, occur when the brain and/or spinal cord are exposed at birth through a defect in the skull or vertebrae without being covered; examples of open NTDs are anencephaly, encephalocele, hydranencephaly and spina bifida. Some of these defects, as anencephaly, are lethal during the pregnancy or just after birth; in other cases, such as spina bifida, infants may survive with severe disabilities (1, 2). One of the most severe NTD is myelomeningocele with Arnold-Chiari malformations, in which the unfused portion of the spinal column allows the spinal cord to protrude through the opening and, as a result, the nerves that emerged at that level of the cord are damaged or not properly developed.

It has been well-established that maternal periconceptional supplementation of folic acid reduces the incidence of neural-tube defects by 50 to 70% (3) but the mechanisms underlying this effect have not been clearly elucidated.

Valuable information provided by animal model as well as by studies in patients with NTD (4-7) suggests a strong genetic component. Nevertheless, as most variants in those genes detected in patients are inherited from healthy parents a complex multigenic cause is suggested in the vast majority of patients with NTD, something that is confirmed by heritability studies (8). However there might also be a more severe end to this phenotype, which may even involve less complex or even monogenic forms.

With the purpose of exploring the genetic mechanisms underlying NTD and the role of point mutations in the pathophysiology of the disease, we performed exome sequencing in a cohort of 14 fetuses with severe forms of open neural tube defects, mainly myolomeningocele in combination with Arnold Chiari malformations.

MATERIALS AND METHODS

A cohort consisting of 14 samples of fetuses with neural tube defects, 6 male and 8 female, were included in this study. For all cases an induced abortion occurred before the 22nd week of pregnancy. All of them presented myelomeningocele and Arnold-Chiari malformationn except for one who had an encephalocele plus Arnold-Chiari malformation; a necropsy report was available for all the cases included in the study (detailed information in table 1).

DNA was extracted from central nervous system tissue in ten cases and liver tissue was used for the remaining four samples because of insufficient central nervous system tissue. Parental DNA was available for 6 fetal samples. All samples showed a normal karyotype by G-banding and no CNVs clearly linked to the phenotype regarding previous published data in a chromosomal microarray analysis.

We sequenced the exomes of 2 fetus-parent trios as well as 12 additional fetal samples with neural tube defect plus Arnold-Chiari malformation.

On average, 5.2 Gb of mappable sequence data was obtained per sample after exome enrichment using SureSelect 50Mb human exome kit (v2 in one trio and v4 in another trio, Agilent, Santa Clara, CA, USA) and SOLiD 5500 sequencing. Color space reads were mapped to the hg19 reference genome with the SOLiD LifeScope software which utilizes an iterative mapping approach. Single nucleotide variants (SNVs) were subsequently called by the DiBayes algorithm using the high call stringency. Small insertions and deletions were detected using the SOLiD Small InDel Tool. Called SNVs and indels were combined and annotated using a custom analysis pipeline. On average, 80.9% of reads mapped in the region and 86.6% of the exome was covered at least 10-fold (table 2). On average 44,358 variants were called per exome, 20,467 of which were coding variants. Detailed information per sample is shown in supplementary table 1.

Three different approaches were used to analyze the results. The analysis workflow is summarized in supplementary figure 1.

Firstly, a prioritization strategy was used to identify possible *de novo* mutations in both trios (9). For that analysis, non-genic, intronic and synonymous variants were excluded, as well as likely benign variant described in dbSNP or in the Nijmegen inhouse database which contains exome sequencing data for 672 subjects.

Afterwards, by comparison with data from parental samples all inherited variants were discarded. This analysis resulted in 5 and 3 possible *de novo* variants respectively; Sanger sequencing validation was performed for all those variants. No *de novo* mutations were validated by Sanger sequencing in this study.

Sample	Tissue	Gender	Malformation
1	liver	male	Lumbosacral myelomeningocele, Arnold-Chiari malformation, brachycephaly
2	liver	male	Sacral neural tube defect, Arnold-Chiari malformation
3	central nervous system	female	Sacral neural tube defect, Arnold-Chiari malformation
4	central nervous system	male	Lumbosacral neural tube defect, Arnold-Chiari malformation
5	central nervous system	female	Lumbosacral neural tube defect, Arnold-Chiari malformation
6	central nervous system	female	Lumbosacral neural tube defect, Arnold-Chiari malformation
7	central nervous system	male	Lumbar myelomeningocele, Arnold-Chiari malformation
8	central nervous system	female	Lumbar myelomeningocele, Arnold-Chiari malformation
9	liver	male	Lumbosacral myelomeningocele, Arnold-Chiari malformation
10	liver	female	Occipital encephalocele, Arnold-Chiari malformation
11	central nervous system	male	Sacral myelomeningocele, Arnold-Chiari malformation
12	central nervous system	female	Lumbosacral myelomeningocele, Arnold-Chiari malformation
13	central nervous system	female	Lumbosacral myelomeningocele, Arnold-Chiari malformation
14	central nervous system	female	Lumbosacral myelomeningocele, Arnold-Chiari malformation

Table 1. List of neural tube defects present in the cohort of 14 fetuses.

Statistic	Value
Total number of bases mapped	5,172,767,655
Total mapped bases in regions (%):	4,184,251,756 (80.9%)
% of bases with 10x coverage	86.6%
Average of reads per base	80,2
Median of reads per base	64,0

Table 2. Exome sequencing statistics considering all samples except for case 13 which was enriched using a different version (v4), SureSelect 50Mb human exome kit v4, Agilent, Santa Clara, CA, USA.

Subsequently, an unbiased approach for the entire cohort was considered to analyze the results regardless to the gene function. With this aim, only private variants were considered (not described in dbSNP or in the Nijmegen in-house database). An overlap analysis was performed in order to identify genes with a possible mutation in several samples (10). We prioritized the genes with variants in three or more samples and also some genes with a variant in two samples if both variants showed a high PhyloP (basepair conservation) value (phyloP > 5.0). For the overlap analysis, data from a cohort of the same sample size but resenting a different phenotype was used to exclude an enrichment of genes displaying an excess of genetic variation. The list of genes obtained in the overlap analysis was compared to the genes with variants in a cohort of 14 samples of patients with congenital heart defects (CHD). By using the same list of private variants, a second analysis was done regarding the most damaging variants, nonsense as well as frameshift. Among those variants, only the ones in genes with no nonsense variants described in Exome Variant Server (EVS) (11) were selected for further studies. The availability of parental samples was also considered for validations.

For the third approach, a candidate gene list consisting of 390 genes (supplementary table 2) for neural tube defects was elaborated including genes previously described in humans related to a similar type of disease, as well as genes identified in animal models linked to the same malformation. For this approach, all variants with a frequency lower than 2% in dbSNP or in-house database were considered, allowing for the detection of recessive mutations, for which carriers in the general population can be expected. Among them, two different lists were generated regarding the possible inheritance pattern (recessive versus dominant). An overlapping analysis with those lists and the candidate gene list was done.

Exome sequencing data of a cohort of 50 Spanish individuals was used to exclude those variants not described in public databases but frequently found in Spaniards.

The prioritized variants were validated by Sanger sequencing and parental samples were analyzed, when available.

RESULTS

De novo analysis

The *de novo* analysis resulted in 3 and 5 possible *de novo* variants per trio; for all those variants Sanger sequencing validation was performed and no *de novo* mutations were validated.

Genes with variants identified in multiple individual samples

In the overlapping analysis only one gene was identified with variants in three different

samples of fetuses with NTD after excluding the ones also mutated in the cohort of fetuses with CHD used as a control cohort: *PAMR1*. All variants were missense. For two of the samples parental DNA was available; one variant was inherited from the father and the other one from the mother. Variants in this gene and the inheritance pattern when parental samples were available are shown in table 3 and figure 1.

Nonsense and frameshift variants

10 nonsense and frameshift variants in genes without previously reported (different) nonsense variants in EVS (11) were identified in six fetal samples with neural tube defects and Arnold-Chiari malformation for which parental samples were available to be able to establish the inheritance pattern. All of them were inherited. Those variants and the inheritance pattern are shown in table 4.

Candidate gene list approach

Several missense variants not described previously in dbSNP, EVS (11) and in the Nijmegen inhouse database or described with a frequency less than 2% were identified in candidate genes for NTD, as it is shown in table 5. Variants in 5 genes previously related to NTD in humans were identified in the cohort of fetus analyzed in this study.

In the only fetus (case 10) with encephalocele included in the cohort two missense variants not described previously in the gene *HSPG2* were identified (A2869P, R2609Q). No parental samples were available to establish the inheritance pattern in this case. This gene has been related to two severe polymalformation syndromes: Schwartz-Jampel syndrome and dyssegmental dysplasia, Silverman-Handmaker type (12, 13). Some patients with these syndromes suffer from encephalocele.

Two missense variants in the same fetus (case 1) in *GLDC* were detected (R410K, M107V), both inherited from the healthy mother. Variants in this gene have been linked to NTD in humans (4).

In fetus 62295 a missense variant (R463W) in a highly conserved nucleotide (PhyloP=6,4) was identified in *MTHFD1*, a gene largely explored in relation to NTD (14, 15). This variant has not been described in available databases.

Mutations in genes of the planar cell polarity (PCP) pathway described previously in humans as being related to NTD were identified in this cohort. Specifically, variants in *PRICKLE1* (5) (T275M), *CELSR1* (6) (R958Q) and *DACT1* (7) (S542C) were detected. The variant identified in *PRICKLE1* (figure 2) was identical to previously described point mutations in patients with NTD (5). Additional variants in PCP pathway in genes not described previously related to NTD in humans were also identified (*PTK7*, *INTU*, *CELSR2*, *FZD1*). Overall, 7 variants that were never described previously affecting genes in the PCP pathway were detected in the cohort of 14 fetuses with NTD; in a ethnically and size matched control cohort of fetuses with CHD only one private variant was identified in the genes from the PCP pathway. Except for the variant in *CELSR2*, all the variants detected in genes in the PCP pathway were predicted to be damaging or probably damaging by PolyPhen2.

In table 6, an overview of all variants found is shown per sample.

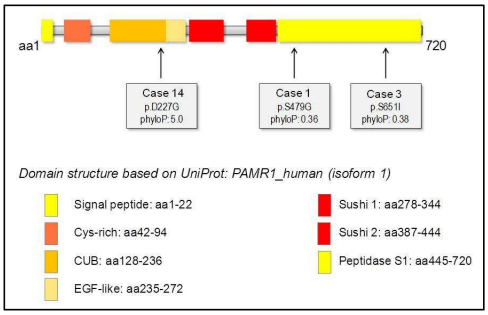


Figure 1. Heterozygous mutations in *PAMR1* identified in three fetuses with NTD. Scheme of the gene showing domain structure and the location of the mutations.

Sample	Gene name	Type of mutation	cDNA level	Protein level	PhyloP	Grantham score	Inheritance
Case 1	*PAMR1*	Missense	1486T>C	S479G S496G	0,358	56	Maternal
Case 3	*PAMR1*	Missense	2003C>A	S651I S668I	0,375	142	-
Case 14	*PAMR1*	Missense	680T>C	D227G	4,989	94	Paternal

Table 3. Variants identified by using the overlap analysis.

Sample	Gene name	Type of mutation	cDNA level	Protein level	Inheritance
Case 1	*RHBDF2*	Nonsense	865G>A	R318* R289*	Maternal
Case 1	*GHR*	Nonsense	476T>A	L159*	Paternal
Case 6	*FBXO36*	Frameshift	183C>	F61FX	Maternal
Case 6	*ESYT3*	Nonsense	379C>T	Q127*	Maternal
Case 2	*FAM96A*	Frameshift	444>A	L148LX	Maternal
Case 14	*C20orf114*	Frameshift	31TG>	C11CX	Maternal
Case 14	*NDRG3*	Frameshift	522CA>	V186VX V174VX	Maternal
Case 12	*FAM24B*	Frameshift	83T>	N28NX	Maternal
Case 12	*UEVLD*	Frameshift	1341T>	K447KX	Maternal
Case 13	*FZD1*	Nonsense	1333C>T	Q445*	Maternal

Table 4. Nonsense and frameshift variants identified in genes with none nonsense variants described in exome variant server in the 6 samples of neural tube defects for which parental samples were available.

Sample	Gene name	Type of mutation	cDNA level	Protein level	PhyloP	Grantham score	Inheritance
Case 1	*GLDC*	Missense	1229C>T	R410K	5,805	26	Maternal
Case 1	*GLDC*	Missense	319T>C	M107V	1,733	21	Maternal
Case 1	*CELSR2*	Missense	6886C>G	R2296G	2,297	125	Maternal
Case 2	*TSC2*	Missense	1318G>A	G440S	5,626	56	Paternal
Case 6	*TSC2*	Missense	4795G>A	V1576M V1532M V1599M	2,627	21	Paternal
Case 5	*INTU*	Missense	1027G>A	V343M	3,913	21	-
Case 6	*LMX1A*	Missense	979G>A	H327Y	2,409	83	Paternal
Case 8	*CELSR1*	Missense	2873C>T	R958Q	3,717	43	-
Case 9	*CECR2*	Missense	2048C>G	P541R P682R P683R	4,181		-
Case 9	*PTK7*	Missense	1264G>A	A552T A512T A422T	2,657	58	-
Case 10	*HSPG2*	Missense	8605C>G	A2869P	0,405	27	-
Case 10	*HSPG2*	Missense	7826C>T	R2609Q	0,883	43	-
Case 9	*MTHFD1*	Missense	1387C>T	R463W	6,358	101	-
Case 12	*PRICKLE1*	Missense	824G>A	T275M	5,984	81	Paternal
Case 13	*DACT1*	Missense	1624A>T	S505C S542C	2,455	112	Paternal
Case 13	*FZD1*	Nonsense	1333C>T	Q445*	5,649	na	Maternal
Case 14	*FBXW11*	Missense	926G>A	T296M T275M T309M	6,003	81	Paternal

Table 5. Variants identified in candidate genes for neural tube defects.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12	Case 13	Case 14	Total
GLDC	2														2
CECR2									1						1
LMX1A						1									1
CELSR1								1							1
MTHFD1									1						1
INTU					1										1
TSC2		1				1									2
HSPG2										2					2
CELSR2	1														1
PTK7									1						1
FBXW11														1	1
PRICKLE1												1			1
FZD1													1		1
DACT1													1		1
PAMR1	1		1											1	3
Total candidate variants per case	4	1	1	0	1	2	0	1	3	2	0	1	2	2	20

Table 6. Overview of LHH candidate variants (based on tables 5 and 7) identified in per fetuses with neural tube defects.

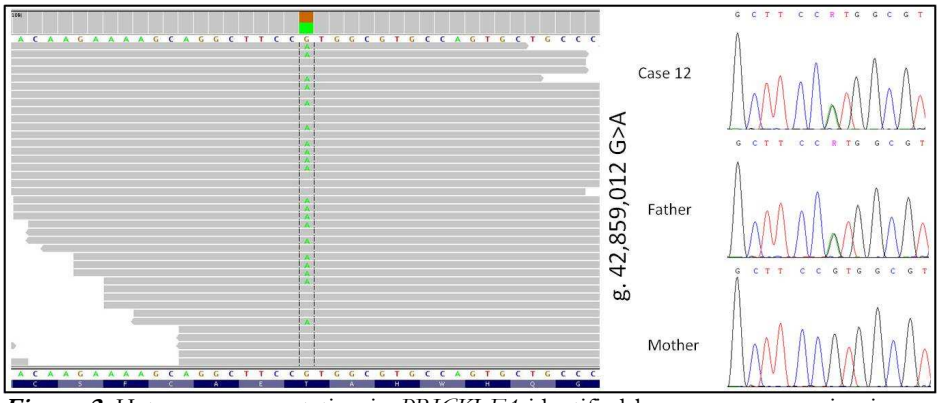


Figure 2: Heterozygous mutation in *PRICKLE1* identified by exome sequencing in case 12. The left panel shows sequencing reads from case 12 (hg19); the right panel shows the confirmation of the mutation by Sanger sequencing in the case and the presence of the same mutation in the father.

DISCUSSION

Though we were unable to detect any *de novo* mutations by the trio analysis, and we also were not able to show any *de novo* nonsense mutation, especially the candidate gene approach' delivered several interesting findings in this study.

The variant that we have identified in *PRICKLE1* (case 12), can be considered as a proof-of-principle. The same variant, causing the substitution T275M, was described in a patient with lumbosacral myelomeningocele, hydrocephalus, Chiari type II malformations and tethered cord. The aminoacid is located in a LIM domain in an absolutely conserved "HWHAT" motif. The substitution is supposed to be damaging as it is not conservative because it reduces hydrophilicity. In the same article they showed that this variant in zebrafish acts antagonistically with the wild type protein (5). Bosoi *et al.* sequenced *PRICKLE1* in a cohort of 810 NTD patients affected with nonsyndromic or isolated NTDs and identified 7 potentially damaging variants in this gene. The inheritance pattern was tested in 4 of them and all were inherited from an apparently healthy progenitor. By studying the exomes of 14 fetuses with NTD several variants in genes described previously related to NTD in humans were identified (4-7, 12, 13, 17-18). Some of these variants were found in genes from the planar-cell-polarity (PCP) pathway, one of the most explored pathways in relation to NTD.

Various PCP genes in mouse models have been described as causing NTD and several rare variants in those genes have also been reported in humans. PCP pathway is a non-canonical Wnt signaling pathway implicated in embryonic patterning, including cell shape, movement and polarity within the plane of an epithelial sheet. In vertebrates, PCP genes are required for convergent extension, a morphogenetic process in which cells within an epithelial sheet rearrange in a coordinated manner to produce a narrowing of the tissue along one axis and an extension along the orthogonal axis. In neural tube closure, these movements lead to the progressive narrowing of the folding neural plate. A group of core PCP genes were identified by Simons *et al.*: frizzled, disheveled, prickle-spiny legs, van gogh, flamingo/starry nights, and diego in Drosophila. The PCP pathway is highly conserved in vertebrates where all these core PCP genes have corresponding orthologs and homologs. Vertebrate PCP genes include Fzd3, Fzd6, Vangl1, Vangl2, Dvl1-3, Celsr1-3, Prickle1 and Dact1 among others (17).

In our study a fetus with NTD (case 8) we identified a missense variant in *CELSR1* not described previously. Robinson *et al.* described 3 missense variants not described previously in *CELSR1* in a cohort of 72 patients with craniorachischisis, one of the most severe forms of neural tube defect in which both the brain and spinal cord remain open to varying degrees. The inheritance pattern could was explored in two patients and both variants were inherited form a healthy progenitor. They showed that the protein subcellular localization was modified by the missense variants identified in *CELSR1* (6). An additional missense variant was also identified in *CELSR2* in the studied cohort.

In another gene of the PCP pathway, *INTU*, a missense variant was identified in a nucleotide with a phyloP of 3.9 in fetus 62291. Park *et al.* reported defects in neural tube closure, including exencephaly, spina bifida, and holoprosencephaly-like features in Xenopus embryos with a disruption in Xit, the homolog gene of *INTU* (22).

A missense variant in *PTK7*, another PCP pathway gene, was found in a fetus with NTD and Arnold-Chiari malformation (case 9), this gene was previously described as causing NTD in mice. Paudyal *et al.* demonstrated that mutant embryos for the homolog of *PTK7* fail to undergo initiation of neural tube closure, and had characteristics consistent with defective convergent extension (23, 24).

The complexity of the possible disease mechanism might be demonstrated by the interesting finding in another fetus (case 13). Here even two variants in genes of the PCP pathway were identified. A nonsense variant in *FZD1* was inherited from the healthy mother and a missense variant in *DACT1* in a nucleotide with a PhP of 2.5 was inherited from the father (figure 3). Variants in *DACT1* were described in patients with NTD by Shi *et al.* (7). Five missense heterozygous variants were identified in a cohort of 167 stillborn or miscarried fetuses with NTD. For two of the variants they proved a functional impairment, as DACT1 showed loss-of-function or reduced activity in inducing Dishevelled degradation and inhibiting JNK phosphorilation (18). As it was mentioned, a stop variant in *FZD1* was also identified in the same samples. No nonsense variant in *FZD1* has been reported in EVS (11). Fzd1 was shown to be associated to tissue fusion processes in mice models, concerning the closure of the palate and ventricular septum and neural tube. Mice model with homozygous knockouts for Fzd1 presented palate clefting as well as cardiac defects; nonetheless, NTD were not

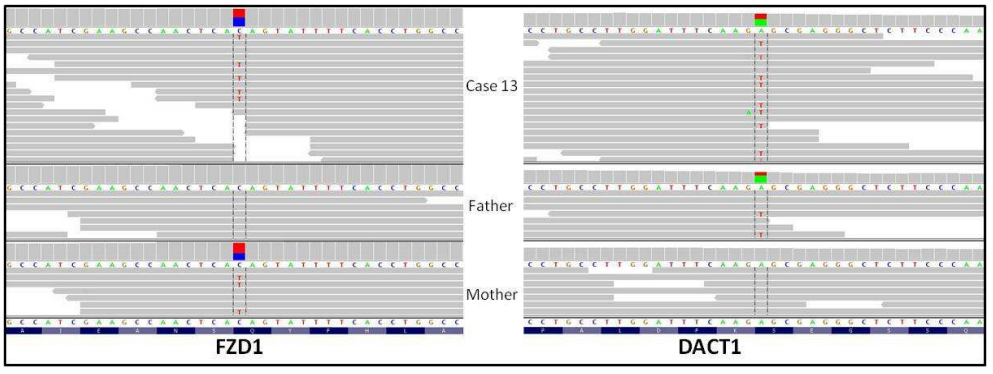


Figure 3: Heterozygous mutations in *FZD1* and *DACT1* identified by exome sequencing in case 13. The left panel shows sequencing reads from case 13, father and mother for the nonsense mutation in *FZD1* (hg19); the right panel shows sequencing reads of a missense mutation in *DACT1* in the same sample. *FZD1* is inherited from the mother and *DACT1* is inherited from the father.

reported. Interestingly, compound heterozygote combinations for Fzd1 and Vangl2 – a PCP gene – showed a high frequency of NTD (20). Suriben *et al.* showed in an animal model that Dact1 forms a protein complex with Vangl2 (21). Regarding the results of Yu *et al.* and Suriben *et al.*, it could be hypothesized that a possible digenic inheritance for a nonsense variant in *FZD1* and a missense variant in *DACT1* was the cause underlying NTD in the fetus studied.

Regarding the data shown, it seems evident that the PCP pathway has a relevant role in NTD also in humans. Seven variants in genes related to this pathway have been identified in six samples of the analyzed cohort. In one of the fetuses, a possible digenic inheritance similar to the one described in mice has been detected and in another fetus a variant previously described in humans has been identified, placing PCP as one of the most interesting pathways regarding mechanisms underlying NTD in humans. The significance of this finding was further shown by the pathway enrichment analysis, another cohort of 14 fetuses with congenital heart defects showed only 1 private variant in the same genes.

Another pathway, folate one-carbon metabolism (FOCM) has also been strongly implicated in the causes underlying NTD due to the association of folic acid intake and the incidence of NTD. Nonetheless, the mechanisms have not been yet elucidated. FOCM comprises a network of enzymes required for the synthesis of purines and thymidylate (for the DNA synthesis) and methionine (for the methylation of biomolcules). FOCM operates in the cytosol as well as in the mitochondria. So far, several genes funcionally related to folate metabolism have been analyzed. Some SNPs in genes like *MTHFR*, *DHFR*, *MTHFD1*, *MTRR* and *TYMS* have been found in association to NTD although none strong findings have been reported (14-16). Recently, a mitochondrial component of FOCM, the glycine cleavage system (GCS) has also been explored and variants not described previously in two genes (*AMT* and *GLDC*) were identified (4). Moreover, the homozygous mice model that lacked GCS activity showed NTD (exencephaly and craniorachischisis). In our cohort of 14 fetuses with NTD one missense variant in *MTHFD1* not present in dbSNP134, in-house database nor in EVS (11) was identified in fetus 62295. In another gene related to this pathway, *GLDC,* 2 missense variants in the same sample (62104) in both inherited from the healthy mother. Strikingly, several variants described by Narisawa *et al.* in *GLDC* were inherited from a healthy progenitor as well.

We have also observed private non-synonymous variants in other genes that have been previously been implicated in related disorders. Two variants in *HSPG2* were identified in the same fetus (case 10). *HSPG2* has been related to two different syndromes: Silverman-Handmaker type of dyssegmental dysplasia and Schwartz-Jampel syndrome; in both with an autosomal recessive pattern of inheritance. Patients with Silverman-Handmaker type of dyssegmental dysplasia present dwarfism with anisospondylic micromelia and, in some cases, encephalocele; patients with Schwartz-Jampel syndrome have osteoarticular abnormalities, short stature as well as myotonic myopathy. The mice model lacking Hspg2 have a severe chondrodysplasia with dyssegmental ossification of the spine and show radiographic, clinical and chondro-osseous morphology similar to a lethal autosomal recessive disorder in humans termed dyssegmental dysplasia, Silverman-Handmaker type (13). Only a few patients have been reported with Silverman-Handmaker syndrome and some of them presented exencephalocoele. The fetus with two hits in HSPG2 had exencephalocele. Nevertheless, the mutations reported in Silverman-Handmaker syndrome are frameshift or variants in splice site causing skipping of exons. In Schwartz-Jampel syndrome, missense mutations have been described. Both variants in the fetus analyzed

in this study were missense variants (13). Considering the phenoype of the fetus and the data published in relation to the gene, it seems plausible that those variants were related to the maformation. Nevertheless, the study of parental sampels to determine the inheritance pattern would be valuable in order to obtain more information supporting this hypothesis. Especially considering possible recessive disorders need to be considered carefully, as respective parents might have a 25% recurrence risk.

One gene, *PAMR1*, emerged in the overlap analysis because three samples had a variant not described previously contained in this gene. PAMR1 has a high expression in brain and, particularly, in spinal cord (25). The combination of rare/private nature of these variants, the finding in 3 of 14 cases, and the neural and spinal cord expression might be sufficient to speculate that this is a new candidate gene for NTD.

Among the genes which showed rare/private variants in more than 1 case, there was *TSC2*. Two missense variants in *TSC2*, not described previously, were identified in 2 fetuses with NTD (case 2 and case 6). The tuberous sclerosis gene can act as a tumor suppressor whose mutations in an heterozygote state predispose to benign and malignant lesions in a variety of tissues, known as tuberous sclerosis disease. In rats, the heterozygous inactivation of TSC2 leads to the same phenotype as in humans. Nevertheless, the homozygous state mutation in rats is lethal in mid-gestation when *TSC2* mRNA is highly expressed in embryonic neuroepithelium. This rat model shows dysraphia and papillary overgrowth of the neuroepithelium (26). A similar phenotype has been reported in homozygous mutations in *TSC1* (27). One might speculate whether heterozygous variants might predispose to NTD in humans, or whether our current analysis might have missed a 2nd variant.

The results shown in this study proves the relevance of PCP pathway and FOCM pathway in NTD and provides a valuable example regarding a putative oligogenic model in NTD. Moreover the identification of previously known variants like the mutation in *PRICKLE1* support that the severity of the phenotypes analyzed here allows speculations on monogenic forms, or at least major genetic contributors, with potentially candidate genes emerging from this study, as *PAMR1* or *TSC2*, in the pathophysiological basis of NTD. In order to elucidate the contribution of all the genes indentified in this study as potential candidates in NTD, the analysis of larger cohorts will be essential. The study of the inheritance pattern will also be valuable, not only to identify *de novo* mutations but also to explore a possible oligogenic model. In general, we believe that systematic fetal pathology and a thorough genetic workup will improve future genetic counseling and add to fundamental genetic and biological knowledge and can be considered as one of the new frontiers in human genetics (28).

REFERENCES

(1) Northrup H, Volcik KA. Spina bifida and other neural tube defects. Curr Probl Pediatr. 2000 Nov-Dec;30(10):313-32.

(2) Sandler AD. Children with spina bifida: key clinical issues. Pediatr Clin North Am. 2010 Aug;57(4):879-92.

(3) van der Linden IJ, Afman LA, Heil SG, Blom HJ. (2006). Genetic variation in genes of folate metabolism and neural-tube defect risk. Proc Nutr Soc 65:204-215.

(4) Narisawa A, Komatsuzaki S, Kikuchi A, Niihori T, Aoki Y, Fujiwara K, Tanemura M, Hata A, Suzuki Y, Relton CL, Grinham J, Leung KY, Partridge D, Robinson A, Stone V, Gustavsson P, Stanier P, Copp AJ, Greene ND, Tominaga T, Matsubara Y, Kure S. (2012). Mutations in genes encoding the glycine cleavage system predispose to neural tube defects in mice and humans. Hum Mol Genet 21(7):1496-503.

(5) Bosoi CM, Capra V, Allache R, Trinh VQ, De Marco P, Merello E, Drapeau P, Bassuk AG, Kibar Z. Identification and characterization of novel rare mutations in the planar cell polarity gene PRICKLE1 in human neural tube defects. Hum Mutat. 2011 Dec;32(12):1371-5. doi: 10.1002/humu.21589. Epub 2011 Sep 23.

(6) Robinson A, Escuin S, Doudney K, Vekemans M, Stevenson RE, Greene ND, Copp AJ, Stanier P. (2012). Mutations in the planar cell polarity genes CELSR1 and SCRIB are associated with the severe neural tube defect craniorachischisis. Hum Mutat 33(2):440-7.

(7) Shi Y, Ding Y, Lei YP, Yang XY, Xie GM, Wen J, Cai CQ, Li H, Chen Y, Zhang T, Wu BL, Jin L, Chen YG, Wang HY. (2012). Identification of novel rare mutations of DACT1 in human neural tube defects. Hum Mutat.

(9) Vissers LE, de Vries BB, Osoegawa K, Janssen IM, Feuth T, et al. (2003) Array-based comparative genomic hybridization for the genome wide detection of submicroscopic chromosomal abnormalities. Am J Hum Genet 73(6):1261-70.

(10) Hoischen, A., van Bon, B. W. M., Rodriguez-Santiago, B., Gilissen, C., Vissers, L. E. L. M., de Vries, P., Janssen, I., van Lier, B., Hastings, R., Smithson, S. F., Newbury-Ecob, R., Kjaergaard, S., and 11 others. De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. Nature Genet. 43: 729-731, 2011.

(11) Exome Variant Server, NHLBI Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [date June 2012 accessed].

(12) Nicole, S., Davoine, C.-S., Topaloglu, H., Cattolico, L., Barral, D., Beighton, P., Ben Hamida, C., Hammouda, H., Cruaud, C., White, P. S., Samson, D., Urtizberea, J. A., Lehmann-Horn, F., Weissenbach, J., Hentati, F., Fontaine, B. (2000). Perlecan, the major proteoglycan of basement membranes, is altered in patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia). Nature Genet 26: 480-483.

(13) Arikawa-Hirasawa, E., Wilcox, W. R., Le, A. H., Silverman, N., Govindraj, P., Hassell, J. R., Yamada, Y. (2001). Dyssegmental dysplasia, Silverman-Handmaker type, is caused by functional null mutations of the perlecan gene. Nature Genet 27: 431-434.

(14) Blom, H.J., Shaw, G.M., Den Heijer, M. and Finnell, R.H. (2006). Neural tube defects and folate: case far from closed. Nat. Rev. Neurosci 7, 724–731.

(15) Molloy, A.M., Brody, L.C., Mills, J.L., Scott, J.M. and Kirke, P.N. (2009). The search for genetic polymorphisms in the homocysteine/folate pathway that contribute to the etiology of human neural tube defects. Birth Defects Res A Clin. Mol. Teratol 85, 285–294.

(16) Beaudin, A.E. and Stover, P.J. (2009). Insights into metabolic mechanisms underlying folate-responsive neural tube

defects: a mini review. Birth Defects Res. A Clin. Mol. Teratol 85, 274–284.

(17) Wu G, Huang X, Hua Y, Mu D. Roles of planar cell polarity pathways in the development of neural [correction of neutral] tube defects. J Biomed Sci. 2011 Aug 24;18:66. Review.

(18) Wen J, Chiang YJ, Gao C, Xue H, Xu J, Ning Y, Hodes RJ, Gao X, Chen YG. (2010). Loss of Dact1 disrupts planar cell polarity signaling by altering dishevelled activity and leads to posterior malformation in mice. J Biol Chem 285(14):11023-30.

(19) Simons M, Mlodzik M. (2008). Planar cell polarity signaling: from fly development to

human disease. Annu Rev Genet 42:517–540.

(20) Yu H, Smallwood PM, Wang Y, Vidaltamayo R, Reed R, Nathans J. (2010). Frizzled 1 and frizzled 2 genes function in palate, ventricular septum and neural tube closure: general implications for tissue fusion processes. Development 137(21):3707-17.

(21) Suriben R, Kivimäe S, Fisher DA, Moon RT, Cheyette BN. (2009). Posterior malformations in Dact1 mutant mice arise through misregulated Vangl2 at the primitive streak. Nat Genet 41(9):977-85.

(22) Park, T. J., Haigo, S. L., Wallingford, J. B. Ciliogenesis defects in embryos lacking inturned or fuzzy function are associated with failure of planar cell polarity and Hedgehog signaling. Nat. Genet. 38: 303-311, 2006.

(23) Paudyal A, Damrau C, Patterson VL, Ermakov A, Formstone C, Lalanne Z, Wells S, Lu X, Norris DP, Dean CH, Henderson DJ, Murdoch JN. The novel mouse mutant, chuzhoi, has disruption of Ptk7 protein and exhibits defects in neural tube, heart and lung development and abnormal planar cell polarity in the ear. BMC Dev Biol. 2010 Aug 12;10:87.

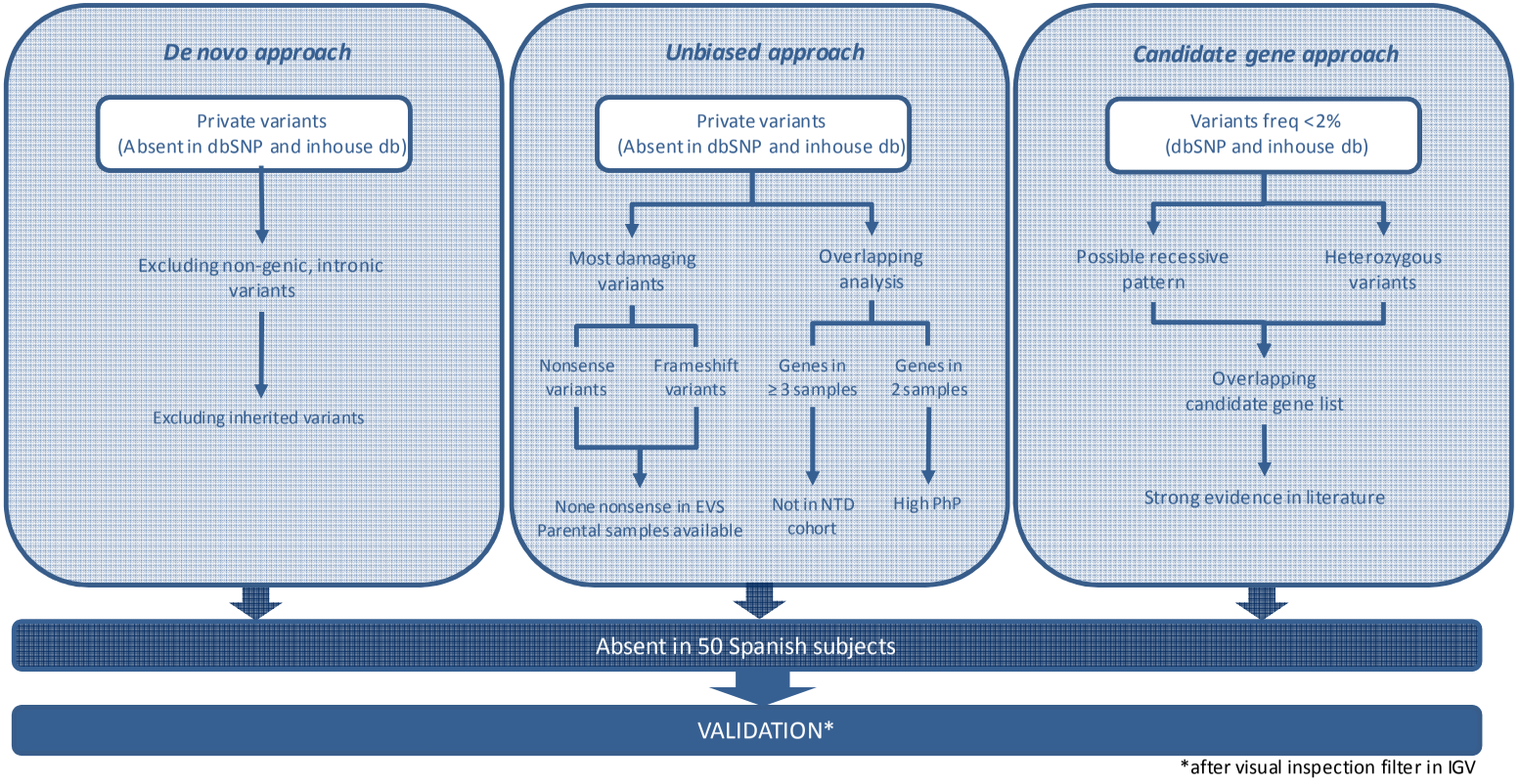
(24) Savory JG, Mansfield M, Rijli FM, Lohnes D. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. Development. 2011 Apr;138(7):1361-70. Epub 2011 Feb 24.

(25) Nakayama Y, Nara N, Kawakita Y, Takeshima Y, Arakawa M, Katoh M, Morita S, Iwatsuki K, Tanaka K, Okamoto S, Kitamura T, Seki N, Matsuda R, Matsuo M, Saito K, Hara T. Cloning of cDNA encoding a regeneration-associated muscle protease whose expression is attenuated in cell lines derived from Duchenne muscular dystrophy patients. Am J Pathol. 2004 May;164(5):1773-82.

(26) Rennebeck G*, Kleymenova EV†, Anderson R†, Yeung RS‡, Artzt K*, Walker CL†§. (1998). Loss of function of the tuberous sclerosis 2 tumor suppressor gene results in embryonic lethality characterized by disrupted neuroepithelial growth and development. Proc. Natl. Acad. Sci. USA 95: 15629–15634.

(27) Kobayashi T, Minowa O, Sugitani Y, Takai S, Mitani H, Kobayashi E, Noda T, Hino O. (2001). A germ-line Tsc1 mutation causes tumor development and embryonic lethality that are similar, but not identical to, those caused by Tsc2 mutation in mice. Proc Natl Acad Sci USA 98(15):8762-7.

(28) Opitz JM. (2012). 2011 William Allan Award: development and evolution. Am J Hum Genet 90(3):392-404.



Supplementary figure 1: Analysis workflow to analyze exome sequencing data in neural tube defects. Overview of the three strategies used to identified variants potentially related to neural tube defects in the cohort analyzed.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12	Case 13*	Case 14	Average	St dev
% of bases with less than 1 read	4,6%	4,4%	5,6%	5,3%	5,4%	5,8%	5,9%	6,1%	4,9%	5,1%	4,8%	5,5%	1,8%	4,9%	5,3%	0,52
% of bases with less than 5 read	8,9%	8,4%	10,3%	9,9%	10,1%	10,4%	10,6%	11,1%	9,32%	9,6%	9,1%	10,0%	4,5%	9,2%	9,8%	0,76
% of bases with less than 10 read	12,3%	11,5%	14,0%	13,7%	13,75%	14,2%	14,3%	14,95%	12,8%	13,1%	12,5%	13,6%	7,7%	12,5%	13,3%	0,97
% of bases with less than 20 read	18,3%	17,0%	20,4%	20,5%	20,2%	20,5%	20,8%	21,49%	19,0%	18,9%	18,3%	19,5%	15,4%	18,1%	19,5%	1,31
Average of reads per base	79,1	91,8	75,2	71,3	78,6	76,5	76,8	75,0	80,4	88,8	88,5	80,0	69,4	89,3	80,9	6,55
Median of reads per base	62,6	72,5	60,5	57,0	62,7	61,1	61,7	60,6	64,3	70,0	70,4	64,6	56,3	70,6	64,5	4,83
Total mapped bases in regions (%):	83,5%	85,2%	78,7%	79,4%	80,1%	78,2%	77,3%	78,3%	81,7%	81,8%	82,8%	81,4%	86,4%	83,4%	80,9%	2,44
Total mapped bases near regions (%):	7,3%	6,6%	9,0%	8,5%	8,6%	9,3%	9,6%	9,4%	8,2%	8,1%	7,7%	7,5%	5.8%	7,4%	8,2%	0,92
Total mapped bases outside regions (%):	9,2%	8,2%	12,4%	12,1%	11,4%	12,5%	13,2%	12,3%	10,1%	10,1%	9,6%	11,1%	7,7%	9,2%	10,9%	1,58
Total mapped bases (Mb):	4895,2	5457,9	5037,5	4700,7	5093,0	5116,2	5216,5	5031,9	5065,5	5620,4	5491,0	5036,5	4257,7	5483,7	5172,8	267,58

Supplementary table 1. Overview of exome-sequencing performance per sample. *This sample was sequenced after exome enrichment using SureSelect 50Mb human exome kit v4, Agilent, Santa Clara, CA, USA.

Candidate gene list for neural tube defects											
ABI1	*BIRC5*	*CRABP1*	*DNMT1*	*FREM2*	*HDAC4*	*LEPR*	*MLIN41*	*NOS2A*	*PIP5K1C*	*RYBP*	*SOBP*	*TRPM6*
ABL1	*BMP2*	*CRABP2*	*DNMT3A*	*FTHFS*	*HECTD1*	*LGL*	*MLP*	*NOS3*	*PIP5KIΓ*	*SALL1*	*SOD2*	*TRX2*
ABL2	*BMP4*	*CRC*	*DNMT3B*	*FUZ*	*HES1*	*LIM1*	*MRJ*	*NOTCH3*	*PITX2*	*SALL2*	*SP2H*	*TSC1*
ADNP	*BMP5*	*CREBBP*	*DNMT3L*	*FZD1*	*HES3*	*LMO4*	*MSGN1*	*NUMB*	*PKA*	*SALL4*	*SP8*	*TSC2*
AGO2	*BMP7*	*CRSH*	*DVL1*	*FZD2*	*HIC1*	*LRP6*	*MSHMT*	*NUP50*	*POR*	*SARDH*	*SPCA1*	*TULP3*
AHCY	*BPH*	*CSHMT*	*DVL2*	*FZD3*	*HIF1A*	*LUZP1*	*MSX1*	*OPB2*	*PRICKLE1*	*SCRB1*	*SPHK1*	*TWIST1*
AICAR	*BRD2*	*CSK*	*DVL3*	*FZD6*	*HIPK1*	*MACSK*	*MSX2*	*OPM*	*PRICKLE2*	*SCRIB*	*SPHK2*	*TXN2*
AICARFT	*BRG1*	*CSNK2A1*	*DYNC2H1*	*GADD45*	*HIPK2*	*MAP3K4*	*MTHFC*	*OTX2*	*PRKACA*	*SCY*	*SPINT2*	*TYMS*
ALDH1A2	*BST*	*CTH*	*DYNC2LI1*	*GAMT*	*HMX1*	*MAP3K7*	*MTHFD*	*OVOL2*	*PRKACB*	*SEC24B*	*SRG3*	*UCP2*
AMBRA1	*C2CD3*	*CTHRC1*	*EDG2*	*GAPDH*	*HOGG1*	*MAPK8*	*MTHFD1*	*P38IP*	*PRMT1*	*SFRP1*	*SS18*	*VANGL1*
AMD1	*CALR*	*CTNNBIP1*	*EFNA5*	*GARFT*	*HOX*	*MAPK9*	*MTHFD1L*	*PALLD*	*PRMT2*	*SFRP2*	*SUFU*	*VANGL2*
AMT	*CART1*	*CUBN*	*EIF2C2*	*GART*	*HSPG2*	*MARCKS*	*MTHFR*	*PAR1*	*PTCH1*	*SFRP5*	*SUZ12*	*VASP*
AP2	*CASP3*	*CX40*	*ENAH*	*GCN5*	*ICAT*	*MARCKSL1*	*MTOR*	*PAR2*	*PTDSR*	*SHH*	*SYT*	*VCL*
AP2A	*CASP9*	*CX43*	*ENPP2*	*GCPII*	*ICMT*	*MAT1A*	*MTR*	*PAX1*	*PTK7*	*SHMT*	*TAK1*	*WNT*
APAF1	*CAT*	*CYCS*	*EP300*	*GCS*	*IFT122*	*MAT2A*	*MTRR*	*PAX2*	*PTPN11*	*SHMT1*	*TCFAP2A*	*WNT11*
APEX1	*CBS*	*CYP26A*	*EPHA7*	*GCSH*	*IFT172*	*MAT2B*	*MTTP*	*PAX3*	*PTPN9*	*SHRM*	*TCNII*	*WNT3A*
APOB	*CC2D2A*	*CYP26A1*	*EPOR*	*GJA1*	*IKBKB*	*MDM4*	*MTU*	*PAX5*	*RAB23*	*SHROOM3*	*TCOF1*	*WNT5A*
ARL13B	*CCL2*	*CYP26B1*	*F2R*	*GLDC*	*INKA1*	*MED12*	*MUT*	*PAX6*	*RAC1*	*SIRT1*	*TEAD2*	*XNP*
ARNT	*CECR2*	*CYPOR*	*F2RL1*	*GLI2*	*INPP5E*	*MEG2*	*NAP1L2*	*PAX7*	*RALDH2*	*SKI*	*TERC*	*XPD*
ATIC	*CELSR1*	*DACT1*	*FDFT1*	*GLI3*	*INTU*	*MEKK4*	*NAT1*	*PAX8*	*RARA*	*SLC25A19*	*TFAP2A*	*XRCC1*
ATP2C1	*CELSR2*	*DBF*	*FGF8*	*GLUT1*	*ITGA3*	*MEN1*	*NAT2*	*PCFT*	*RARG*	*SLUG*	*TFT*	*XRCC2*
ATR2	*CELSR3*	*DEAF1*	*FGFR1*	*GLUT4*	*ITGB1*	*MENA*	*NCAM1*	*PCMT1*	*RBPSUH*	*SMAD5*	*TGIF1*	*XRCC3*
ATRX	*CEP290*	*DHF*	*FKBP1A*	*GNB1*	*ITPK1*	*MFT*	*NDST1*	*PCYT1A*	*RERE*	*SMARCA4*	*THF*	*YBX1*
ATX	*CFL1*	*DHFR*	*FKBP8*	*GRHL2*	*JARID2*	*MFTHFS*	*NEUROG2*	*PDGFRA*	*RFAP2A*	*SMARCC1*	*THM1*	*YY1*
AXD	*CHKA*	*DLC1*	*FOLR1*	*GRHL3*	*JNK1*	*MGMT*	*NF1*	*PDN*	*RFC1*	*SMO*	*TP53*	*ZFP36L1*
AXIN1	*CHUK*	*DLD*	*FOLR2*	*GRIP1*	*JNK2*	*MIB2*	*NNMT*	*PEMT*	*RGMA*	*SMURF1*	*TRAF4*	*ZIC1*
BBS4	*CITED2*	*DLX5*	*FOLR3*	*GRLF1*	*KAT2A*	*MKS1*	*NOD*	*PFN1*	*RING1*	*SMURF2*	*TRAF6*	*ZIC2*
BCL10	*COBL*	*DMBO*	*FOXB1*	*GTF2I*	*KIF3B*	*MKS1KRC*	*NOG*	*PGN*	*RNMT*	*SNX1*	*TRDMT1*	*ZIC3*
BHMT	*COMT*	*DMGD*	*FPN1*	*GTF2IRD1*	*KLKB1*	*MKS2*	*NOGGIN*	*PHACTR4*	*RPGRIP1L*	*SNX13*	*TRIM71*	*ZIC4*
BHMT2	*COQ3*	*DNAJB6*	*FREM1*	*GB1*	*LAMA5*	*MKS3*	*NOS1*	*PIGA*	*RPL24*	*SNX2*	*TRP53*	*ZIC5*

Supplementary table 2. List of the genes included in the candidate gene list for neural tube defects. Genes described in animal models related to NTD, in patients with this type of malformations or genes identified as playing an important role in neural tube development have been considered.

DISCUSSION

CMA in the study of isolated congenital malformations: a research perspective

A cohort of 95 fetuses with congenital malformations was analyzed by CMA with the aim of indentifying rearrangements related to the disease. Considering previously published papers, CMA is a valuable tool to find the cause in patients with multiple congenital anomalies and/or mental retardation, as detection rates of 15-24% have been reported (34, 35). Rare CNV not described previously in more than 10,000 or described with a frequency less than 1/2000 were identified in several of the cases analyzed (21,05%); 50% of the rearrangements were duplications and 50% deletions. The high detection rate of rare deletion-type CNVs in fetuses with congenital malformations when compared to control population (4,2% *vs* 11,6%) indicates a relevant role of CNV in birth defects.

The results were significantly different among the groups analyzed. While the frequency of CNVs was similar in fetuses with multiple malformations or isolated renal malformations compared to controls, higher detection rates of rearrangements, especially deletions, were identified in fetuses with CNS malformations (15,4% vs 4,1%) and cases with iCHD (15,2% vs 4,1%). Among fetuses with iCHD, the frequency of rearrangements was substantially different depending on the type of malformation. The frequency of segmental aneusomies in fetuses with heart hypoplasia was remarkable (29,4% vs 4,2% in controls), while the frequency was similar to controls for the other heart anomalies excluding heart hypoplasia. These results highlight the genetic component underlying some congenital defects and point out the differences among malformations affecting different organs and systems.

The inheritance pattern of the rearrangements identified in the study was explored in 9 cases and, except for one, all the rearrangements were inherited from an apparently healthy progenitor. Despite the rarity and the gene content of some of the rearrangements, the presence of the same alteration in a healthy progenitor hinders the causality proof. Although the findings in this study indicate a clearly role of deletion-type CNVs in some types of malformations, it is complicated to prove causality for individual cases.

In our study, two of the rearrangements found, both deletions, could be strongly linked to the malformation detected due to previous findings in other patients. A *de novo* deletion encompassing the cluster of FOX genes in 16q24.1 (60, 124) (*FOXF1*, *FOXC2*, *MTHFSD*, *FLJ30679* and *FOXL1*) was identified in a fetus with LHH; this deletion has been reported in association

to heart malformations, including LHH. A deletion in 15q13.3 in a fetus with RHH was also detected inherited from the healthy mother. Based on reported epidemiological evidence (125, 126), it is plausible to associate the 15q13.3 deletion with CHD, although some others factors might be involved in the physiopathology of the malformation considering the incomplete penetrance in this and other families.

In summary, we can conclude that only 2 out of 95 cases (2,1%) had a confirmatory diagnosis of the probable cause of the malformation using CMA, 2 out of 33 (6,1%) when considering only iCHD. Thus, the diagnostic yield of CMA in individual cases was rather low in our cohort (2/95, 2,1%), better for iCHD (2/33, 6,1%). Two main factors could be related to the lower detection rate of CMA in congenital malformations of our study compared to previous reports (49, 50, 55, 127). First, we had discarded cytogenetic and most well-known rearrangements (i.e. subtelomeric and recurrent) by MLPA. Second, inclusion criteria were more stringent in our study while previous reports had included more polymalformation syndromes.

However, besides the two cases clearly related to the disease, additional alterations potentially linked to the malformations were identified. In a fetus with holoprosencephaly two novel deletions were identified each one inherited from a healthy parent (father and mother, respectively). Both deletions contained genes functioning in the same pathway-phosphatidylinositol pathway- giving more relevance to the finding. Considering the heterogeneity of congenital malformations it is quite difficult to find exactly the same combination of deletions in another case of holoprosencephaly, which would be the manner to prove the causality of the rearrangements. A second case with a possible double-hit mechanism underlying the disease was identified: a fetus with hydrocephalus harbored two duplication-type CNVs, both containing candidate genes. One of the duplications encompassed the gene *DMBX1*, a diencephalon-mesencephalon homeobox associated to brain development (128), while the other duplication included *GPRIN2*, a gene related to neurite overgrowth (129).

In summary, we can conclude that there is a clear role of deletion-type CNVs in some isolated congenital malformations, like heart hypoplasia as well as central nervous system malformations, as well as in polymalformation syndromes. Some rearrangements are strongly linked to the disease and can be used as predictive or diagnostic tools. However, other rearrangements are also likely to contribute to the phenotype but with incomplete penetrance

and through more complex models, in conjunction with point mutations or other rare CNV. From a diagnostic perspective and until more evidence is obtained, these findings should still be classified as variants of unknown significance (VOUS).

CMA in the study of at-risk pregnancies: a clinical perspective

The multicentric study launched to explore the feasibility and efficiency of CMA in a clinical setting provided valuable data regarding the acceptance and clinical utility of the test. The acceptability of informed pregnant women of the use of novel techniques for prenatal diagnosis was very high (94%). Technical performance was excellent for karyotype, QF-PCR, and CMA with a failure rate of 1%, but relatively poor for MLPA (10% failure). Mean turn-around time was 7 days for CMA or MLPA, 25 for karyotype, and two for QF-PCR, with similar combined costs for the different approaches. The detection rate of CMA, when compared to other techniques, was significantly higher, with CMA yielding the highest detection rate (32% above other methods). Overall, 57 clinically significant chromosomal aberrations were found (6.3%).

The main concern for CMA before starting the study was the detection of VOUS and how to deal with them. Indeed, the identification of VOUS by CMA (17, 1,9%) tripled that of karyotype and MLPA, but most alterations could be classified as likely benign after proving they all were inherited. Evidently, the design of CMA used in the study was targeted to specific regions in the genome, reported in association to genetic diseases. Even though the design of CMA can reduce the chances to identify VOUS, it is not possible to avoid completely the detection of rearrangements with unknown clinical significance. In order to interrogate some specific syndromes it is necessary to use probes located in variable regions and, by the time the array was designed, the knowledge of the genome distribution of CNVs was limited. Comparing the data from the study of congenital malformations by whole-genome CMA and the results from targeted CMA, a higher detection rate of VOUS in the first study due to the higher probe density was evident. From a clinical perspective, several rearrangements (6/95, 6,3%) found by CMA in the study of fetuses with congenital malformations would have been classified as VOUS, as those alterations were not described as CNVs in control population. Therefore, the chance to find VOUS is higher in whole genome CMA.

Not only VOUS entailed difficulties in clinical interpretation. During the study, 6 cases of recurrent microduplication syndromes (0.7% of our series), 3 inherited from a phenotypically normal parent and three *de novo*, were reported. The inclusion of probes to detect microdeletions syndromes with well-established clinical features involves the chance to detect the reciprocal microduplication, usually with uncertain clinical repercussion (133, 134). In those cases, we dealt with the difficulty to counsel about genomic rearrangements associated with variable phenotypes and incomplete penetrance with scarce data in literature. We attempted to expose to the couple the limited information available in an appropriate manner to enable them to understand the situation and the possible implications for the child. Following a 20-week normal ultrasound evaluation, couples decided to continue the pregnancy.

Besides the VOUS detection differences between both CMA studies performed, another difference can be noticed. The detection rate of clinically significant chromosomal rearrangements was higher in the multicentric study compared to the study of isolated congenital malformations by high resolution CMA, 6,2% vs 2,1%. This difference is due to the alterations discarded prior to CMA in fetuses with congenital malformations; all samples in the first study presented a normal G-banding karyotype and some recurrent genomic disorders, as well as subtelomeric imbalances, had been discarded before CMA.

We can conclude from our data that the high acceptability among pregnant women, the significantly higher detection rate and the lower turn-around time of targeted CMA compared to karyotype and other standard methods could justify its the relatively higher cost to implement CMA as the first-tier method for detection of chromosomal rearrangements in all at-risk pregnancies after invasive prenatal sampling.

Genetic counseling with novel high-throughput molecular techniques

Among the new molecular techniques used in this thesis, CMA is the only one broadly implemented in clinical setting, although next generation sequencing is starting to be used in clinical practice with some specific indications. Taking advantage of the multicentric study to evaluate CMA in prenatal setting, we decide to also explore the utility of genetic counseling in relation the use of these technologies, including the acceptability of the

technique by pregnant women, the information which should be provided and the socio-demographic factors influencing decision-making.

The acceptability rate of CMA in prenatal setting was high, as 94% of women decided to extend prenatal testing. 91% of them also considered that they had enough information to take decisions. Lack of information has been described in several articles as one of the reasons to reject genetic tests, as patients do not feel capable to make a decision. Therefore, it is essential in genetic counseling to ensure that patients understand properly the information provided so it enables them to decide the most appropriate course of action considering the risk, family goals and the ethical and religious convictions (135-138). Education level plays an important role in comprehension, considering that several concepts used in genetic counseling sessions are difficult to understand and not familiar for most patients. As it was mentioned, some findings in the study entailed difficult situations to counsel, such as VOUS or chromosomal abnormalities with unclear clinical repercussion for the child. A higher education level also provided couples more tools to cope with the uncertainty of the result.

In the study presented in this thesis, education level was also identified as a relevant factor influencing anxiety level and reasons given by pregnant women to extend prenatal studies. Anxiety inversely correlated with education level, indicating that a better understanding reduces anxiety. Pregnant women with high education level also showed a greater sensitization towards scientific research and progress. Besides education level, indication to undergo prenatal invasive sampling and previous reproductive history also influenced anxiety. Thus, it is important to consider the education level and the anxiety of each woman to provide them an appropriate genetic counseling, according to her understanding and her emotional situation to encourage them to express their doubts and concerns.

Among women who rejected to extent prenatal tests, two factors –country of origin and indication for invasive prenatal sampling- were identified as influencing their decision. Half of them were of South American origin compared to the 20% proportion of the entire cohort. This data suggests that, likely due to cultural differences, genetic counseling was not as understandable for South American than for Spanish women despite having the same language. Cultural differences should be taken into account to ensure that the rejection of new molecular techniques is not due to a misunderstanding of the information provided. In addition to the country of origin, indication for prenatal test was also identified as a factor influencing

CMA rejection. Half of the women who refused CMA but agreed to answer the survey had an indication of invasive testing due to ultrasound findings. In those cases, there was already a complication in the pregnancy and the prenatal invasive sampling was indicated to try to find a molecular diagnosis and to define better the prognosis; the higher resolution of CMA could have provided relevant data. Therefore, in those women there was probably a denial of the possible unfavorable evolution of the gestation, trying to avoid testing which confirms a bad prognosis for the pregnancy.

In the course of this study, the relevance of appropriate pre-test genetic counseling was noticed concerning findings of uncertain clinical significance. One of the aspects including in the pre-test genetic counseling session was the possibility to detect clinically relevant findings but also the chance to detect rearrangements of unknown clinical repercussion. Explaining in advance possible uncertain findings reduced the anxiety when VOUS were detected and, therefore, it was easy for the couple to cope with the finding. Otherwise, to communicate a VOUS without previous explanation would have generated a higher level of anxiety in pregnant women.

In summary, there was a high acceptance rate for novel technologies in prenatal setting among pregnant women with at-risk gestations, when appropriate pretest genetic counseling was provided. Relevant factors as education level, anxiety or country of origin were identified influencing women's perception and decision-making and should be considered during genetic counseling. It is indispensable to develop genetic counseling strategies simultaneously to technique improvements and discoveries, as it is essential to communicate in an effective manner the information to patients. Moreover, concerning novel molecular techniques new challenges appear, such as the need to counsel in relation to findings with low penetrance or variable expression, sometimes including the diagnosis of a progenitor who did not know before that had a genetic disorder. Studies to evaluate genetic counseling and to identify relevant factors influencing patients' perceptions and decision making are valuable to improve patients care.

Epigenetics in congenital malformations

Several studies have shown the relevant role of epigenetics in various diseases such as cancer, psychiatric disorders or diabetes. DNA methylation and histone modification are two of the main processes influencing gene expression; therefore, epigenetics might also play an important role in

development and diseases. Data reported in the study of methylation alterations in CHD proved this fact, as altered methylation profiles were detected in genes related to heart development.

Two alterations in methylation patterns in genes previously described in relation to heart development and heart defects were found among the 22 samples studied. A hypermethylation of a broad region in MSX1 in a fetus with double-outlet right ventricle seems promising, as the animal model with alterations in MSX1 and MSX2 suffers the same CHD (130). An altered methylation pattern in GATA4 was identified in all patients when compared to control. Interestingly, it was also altered in samples of fetuses with Down syndrome without CHD. This finding could have two different explanations depending on the disease. In Down syndrome patients, considering that the alteration was detected independently of the presence or absence of CHD, hypermethayltion of GATA4 could be a result of trisomy 21 and probably related to the high incidence of CHD among patients with Down syndrome. Instead, in fetuses with CHD and normal karyotype the methylation pattern in GATA4 might be a consequence of a misregulation of essential pathways in heart development due to the cause underlying the CHD. Other genes were identified, some of them with alterations in more than one sample.

In methylation studies it is valuable to analyze the tissue affected in order to detect more promising alterations. Otherwise, the detection of methylation aberrations in a tissue different from the one affected may not always provide relevant data, considering that methylation pattern might be, in some cases, tissue specific as it is related to gene expression.

To properly define the involvement of methylation alterations in CHD it is necessary to linked epigenetic alterations to expression levels of the genes affected.

Exome sequencing in congenital malformations

In last years, exome sequencing has provided considerable amounts of data concerning mendelian diseases and several genes have been identified in relation to genetic disorders (37). This powerful technique is also valuable in the study of other diseases, more heterogeneous ones, like mental retardation or schizophrenia. In the light of the results obtained by exome sequencing in both studies, data provided by this technique is crucial in order to improve the knowledge of the mechanisms underlying congenital malformations. Some *de novo* mutations in promising genes were identified in LHH fetuses. Several hits in genes that had been already linked to congenital

malformations in humans were detected, as mutations in *PRICKLE1*, *DACT1*, *GLDC1*, *DTNA* or *ACVR2B* (89, 90, 139-142). In addition, mutations in genes related to developmental pathways or described in animal models as causing the same malformation were identified in LHH as well as in NTD patients. Further studies with larger cohorts are required in order to establish the precise role of those genes in development and in congenital malformations.

Two *de novo* mutations were found in two patient-parent trios from the cohort of fetuses with LHH, *USP32* and *NCAPD3*. A second case with a mutation in *USP32* was also found but the inheritance pattern could not be determined because no parental samples were available. As it was mentioned in the chapter of exome sequencing in LHH, *USP32* is a gene from the ubiquitin-proteosome system, a pathway potentially related to heart development (143). *NCAPD3* encodes for a non-SMC subunit of condensin complex II (144) and it is included in the candidate region for CHD in Jacobsen syndrome (145, 146). Interestingly, a significant proportion of patients that suffers this syndrome present LHH. The combination of *de novo* occurrence, mutation severity considering the high conservation of the nucleotide and potential gene function make both genes *USP32* and *NCAPD3* new candidate genes for LHH. None de novo mutations were found in the two patient-parent trios analyzed from the cohort of NTD.

In both studies, variants in genes already described in association to LHH or NTD were identified. Among the samples with LHH, mutations in *DTNA* and *ACVR2B* were detected. *DTNA* was described in a family with six members affected by non-compactation of the left ventricle; two of them also had LHH (140). The variant identified in *ACVR2B*, a gene related to heterotaxy disorder, could be a SNP regarding the data available now in database (141, 142). Regarding NTD, two mutations in *HSPG2* were described in a fetus with encephalocele; this gene was described as the causative gene for Schwartz-Jampel syndrome and dyssegmental dysplasia Silverman-Handmaker type, some of these patients presenting with encephalocele (147, 148). Moreover, variants in *GLDC1*, *DACT1*, *CELSR1* (87-89, 139)and *PRICKLE1* were also identified among the patients with NTD. Mutations in all those genes have been described in cohorts of patients with NTD. Some of the candidate genes mutated in NTDs are involved in the PCP pathway, which plays a relevant role in neural tube closure leading to a progressive narrowing of the folding neural plate. In total, 7 novel variants in genes of the PCP pathway were identified among the samples with NTD, while only 1 variant in PCP genes was identified in

the cohort of fetuses with CHD. Therefore, there was a clear enrichment of variants in the genes of the PCP pathway in fetuses with NTD, suggesting a likely pathogenic implication and reinforcing the essential role of this pathway in neural tube closure.

In both cohorts, in the overlap analysis some genes were identified as being mutated in three samples with the same malformation. The fact that three samples with the same malformation carry mutations not described previously in the same gene is quite promising. In LHH, three genes (*ATP13A5, SCAF11* and *RB1CC1*) were found to be altered in three samples; all of them were missense mutations. Moreover, one of these genes, *RB1CC1*, is a causative gene for heart hypoplasia in animal models (149); therefore, this gene emerged as a candidate gene for LHH. In the cohort of fetuses with NTD, three missense mutations were found in *PMR1*, a gene expressed in spinal cord; all the variants were missense mutations.

Besides *de novo* mutations, variants in genes previously related to congenital malformations and overlap analysis, other promising findings were done, like variants in *PROX1* or *FBXO25* in LHH (150-152). Specifically in *PROX1*, the same mutation was identified in two samples and the animal model presents exactly the same phenotype. Furthermore, a patient with LHH and a balanced rearrangement truncating the regulating region of *PROX1* was reported. Nevertheless, only parental samples of one of the fetuses were available and the mutation was inherited from the mother, who has Wolf-Parkinson-White syndrome. Wolf-Parkinson-White syndrome is a conductive disorder of the heart and it is caused by the presence of an abnormal accessory electrical conduction pathway between atria and ventricles. No data is available nowadays linking this syndrome to LHH.

Among the variants identified in the cohort of fetuses with NTD, a possible digenic mechanism was found. A stop codon in *FZD1* was detected in a fetus inherited from the healthy mother, while a missense mutation in *DACT1* was identified inherited from the father. Both genes had been described previously related to neural tube defects in animal models; *DACT1* had also been described in humans related to the same disease (88). Interestingly, in mice model, NTD were reported in animals heterozygous for the knockout of *Fzd1* and an additional mutation in *Vangl2* (90), a gene forming a protein complex with *Dact1* (153). As in other cases discussed, a second case with the same combination of mutations would definitely prove the implication of those mutations in the ethiopathology of the disease.

The results in the exome sequencing approach clearly indicate a relevant role of point mutations in the genetic mechanisms underlying congenital malformations, probably as the only causative factor in some cases and as part of a digenic or oligogenic model in others. For both hypotheses, the study of larger cohorts of samples with the same malformation would be worthwhile.

Concluding remarks

Congenital malformations are truly heterogeneous disorders due to the amount of genes and processes related to development. As many genes are related to heart development, then mutations in multiple genes can cause CHD. As an example, for prioritizing the analysis of high throughput sequencing projects, the list of candidate genes that we used for NTD included 390 gens and 363 genes for LHH. This fact hinders the finding of a second case with the same genetic alteration in order to prove the causality of the genetic defect in the malformation. Considering the heterogeneity of the disease is essential to have well phenotyped cohorts with detailed necropsy reports in fetal cases and a complete and extensive clinical history in patients. The availability of this information would allow the researchers to study more homogeneous cohorts trying to minimize the existing causative heterogeneity. Availability of parental samples for analyses is also crucial to confirm or discard causality of the findings, checking for *de novo* occurrence of alterations or inheritance patterns compatible with recessive, additive or oligogenic models. Moreover, extended and detailed family history is also very helpful to find potentially affected individuals in the family with milder or related phenotypes, in order to better define causality of the findings and provide appropriate genetic counseling.

Nevertheless, the results of this thesis highlight the genetic component underlying congenital malformations and the implication of CNVs, methylation alterations as well as point mutations in the etiopathology of the disease. Although the role of those alterations seems relevant, it is difficult to prove the causality for each individual alteration. As an example, detection of a high frequency of deletion-type CNV in fetuses with LHH and CNS malformations pointed out deletions as an important factor in the etiology of congenital malformations, but causality was only clearly defined for two of them. Similar difficulties occur with the detection of novel point mutations not previously described, even if they hit candidate genes. Most of them were proved to be inherited when parental samples were available, making

difficult to prove the relation between the variant and the malformation. The study of larger cohorts of patients with the same type of malformation might prove the role of those genes in the pathophysiology of the disease.

In summary, we have proven the clinical utility of prenatal CMA and the implications for genetic counseling. In addition, using different genetic, genomic and epigenomic approaches several new candidate genes for congenital malformations have been found, mainly for LHH and NTD. Further studies including homogeneous and larger cohorts will provide additional data to definitively establish the role of those genes in development and in congenital malformations.

CONCLUSIONS

1. A higher frequency of deletion-type CNVs was detected by CMA in fetuses with isolated central nervous system malformations and isolated heart hypoplasia compared to a control cohort. The vast majority of the alterations tested were inherited from an apparently healthy progenitor denoting that the rearrangements were not the only cause of the disease. Nevertheless, given the rarity and the gene content of those alterations, some of them might be related to the disease as part of a multifactorial model with other genetic or environmental factors involved.

2. Single causative rearrangements were found in 2,1% of fetuses with congenital malformations, a *de novo* deletion and a deletion inherited from the healthy mother. The first one encompasses the *FOX* gene cluster, a region previously described in relation to congenital heart defects. The second one was a deletion in 15q13.3 reported as a microdeletion syndrome with variable clinical features and incomplete penetrance.

3. In each of two fetuses with central nervous system malformations, we identified two different hits potentially related to the phenotype. In one fetus with holoprosencephaly, two deletions encompassing genes in the phosphatydylinositol pathway were identified, one inherited from the mother and the other one from the father. In a fetus with hydrocephalus, two duplication-type CNVs were detected both harboring candidate genes for brain development, *DMBX1* and *GPRIN2*. Therefore, a double hit mechanism causing cases has been hypothesized.

4. We have proven a higher diagnostic yield for CMA compared to MLPA, QF-PCR and standard karyotype in the prenatal diagnosis of at-risk pregnancies. In terms of technical performance, sensitivity, and specificity, CMA proved to be the most reliable technology after invasive prenatal sampling in clinical setting. The number of detected VOUS increased in CMA compared to other techniques (~3 fold) although most alterations could be classified as likely benign after establishing the inheritance pattern.

5. There was a high acceptance rate (94%) of novel techniques for prenatal diagnosis after appropriate genetic counseling among pregnant women with at-risk gestations. Education level was identified as the main factor influencing anxiety and reasons to extend prenatal testing. Higher education level correlated with a decrease in anxiety and with a greater sensitization of the participants towards scientific research and progress. Anxiety of pregnant women was also influenced by the previous reproductive history and the

perception of risk depending on the indication for invasive prenatal sampling.

6. Country of origin and the indication for invasive prenatal sampling were the main aspects related to reject the extent of prenatal testing, as half of the women were from South-America and half of the women had an indication for prenatal sampling due to ultrasound findings. In those cases, there was a likely denial of the possible unfavorable evolution of the pregnancy, considering that complications had already been detected.

7. In the light of the results obtained, it is recommendable to assess the level of understanding of the patients to provide a counseling session according to it. It is also relevant to allow them to set the pace of the session depending on their comprehension and emotional situation. In addition, it is worthwhile to consider the potential cultural differences to try to minimize their effects in women's comprehension and decision-making.

8. We found methylation abnormalities in heart tissue DNA from fetuses with congenital heart defects affecting genes related to heart development and heart malformations, such as *GATA4* and *MSX1*. Other genes potentially related to cell proliferation also affected by methyl epimutations, some in more than one sample, include *FOXN3/CHES1*, *KLF17*, *FAM107A* or *TCL1B*.

9. Pathways related to cell growth and cell death were enriched pathways among the epimutated genes in heart tissue from fetuses with iCHD. Cell proliferation and apoptosis are both processes whose regulation is crucial in heart development. Therefore, the genes identified in those pathways emerged as strong candidate genes for heart malformations.

10. Point mutations in multiple genes play a relevant role in the mechanisms underlying congenital malformations. Thus, exome sequencing is a valuable tool to explore it. *De novo* and some inherited events, affecting genes previously described in human diseases or candidate genes based on functional information or animal models, were identified in >50% of cases from two cohorts of 14 fetuses with isolated and severe congenital anomalies: LHH and NTD plus Arnold-Chiari malformation.

11. Two *de novo* mutations in fetuses with LHH were identified in *USP32* and *NCAPD3*, both potential candidates for heart development and heart disease considering available data. In two samples, mutations in genes previously described in association to heart diseases in humans were identified, *DTNA*

and *ACVR2B*. In addition, several hits in genes potentially related to LHH were detected, like mutations in *PROX1*, *RB1CC1* or *FBXO25*.

12. Planar cell polarity pathway emerged as one of the most relevant pathways affected in NTD, given the high mutation load of genes in this pathway among the studied cohort of fetuses with NTD. Interestingly, digenic inheritance in planar cell polarity genes was identified in a fetus with myelomeningocele. Additional genes, such as *MTHFD1, PAMR1*, *TSC2*, *GLDC* or *CECR2* were also identified as potential candidates to explain the malformation in the neural tube.

13. Congenital malformations are highly heterogeneous disorders due to the huge amount of genes regulating development, with genetic and epigenetic mechanisms involved in their etiopathogenesis. The heterogeneity and the oligogenic/multifactorial pattern underlying an important proportion of cases hinder to prove causality in individual cases. Nevertheless, promising findings and relevant data have been described in this work including several strong candidate genes that may explain the occurrence of the disease in some cases.

REFERENCES

(1) Martínez-Frías ML, Bermejo E, Frías JL. Pathogenetic classification of a series of 27,145 consecutive infants with congenital defects. Am J Med Genet. 2000 Jan 31;90(3):246-9.

(2) Martínez-Frías ML, Frías JL, Opitz JM. Errors of morphogenesis and developmental field theory. Am J Med Genet. 1998 Apr 1;76(4):291-6.

(3) Pierpont ME, Basson CT, Benson DW, Jr., Gelb BD, et al. (2007). Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation 115(23):3015-38.

(4) Thienpont B, Mertens L, de Ravel T, Eyskens B, Boshoff D, et al. (2007). Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. Eur Heart J 28(22):2778-84.

(5) Kalter H, Warkany J. (1983). Congenital malformations: etiologic factors and their role in prevention (first of two parts). N Engl J Med 308(8):424-31.

(6) Kalter H, Warkany J. (1983). Congenital malformations (second of two parts). N Engl J Med 308(9):491-7.

(7) De Galan-Roosen AE, Kuijpers JC, Meershoek AP, van Velzen D. (1998). Contribution of congenital malformations to perinatal mortality. A 10 years prospective regional study in The Netherlands. Eur J Obstet Gynecol Reprod Biol 80(1):55-61.

(8) Botto LD, Correa A. (2003). Decreasing the burden of congenital heart disease: an epidemiologic evaluation of risk factors and survival. Prog Pediatr Cardiol 18:111–21.

(9) Peral, García Minaur. (1998). Registros poblacionales de defectos congénitos en Espana. An Esp Pediatr.

(10) Burd L, Deal E, Rios R, Adickes E, Wynne J, Klug MG. (2007). Congenital heart defects and fetal alcohol spectrum disorders. Congenit Heart Dis 2(4):250-5.

(11) Autti-Rämö I, Fagerlund A, Ervalahti N, Loimu L, Korkman M, Hoyme HE. (2006). Fetal alcohol spectrum disorders in Finland: clinical delineation of 77 older children and adolescents. Am J Med Genet A 15;140(2):137-43.

(12) Hoyme HE, Jones KL, Dixon SD, Jewett T, Hanson JW, Robinson LK, Msall ME, Allanson JE. (1990). Prenatal cocaine exposure and fetal vascular disruption. Pediatrics 85(5):743-7.

(13) Richardson GA, Goldschmidt L, Willford J. (2007). The effects of prenatal cocaine use on infant development. Neurotoxicol Teratol 30(2):96-106.

(14) Chávez GF, Mulinare J, Cordero JF. (1989). Maternal cocaine use during early pregnancy as a risk factor for congenital urogenital anomalies. JAMA 11;262(6):795-8.

(15) Little BB, Snell LM, Palmore MK, Gilstrap LC. (1988). Cocaine use in pregnant women in a large public hospital. Am J Perinatol 5(3):206-7.

(16) Lipshultz SE, Frassica JJ, Orav EJ. (1991). Cardiovascular abnormalities in infants prenatally exposed to cocaine. J Pediatr 118(1):44-51.

(17) Zuckerman B, Frank D, Brown E. (1995). Overview of the effects of abuse and drugs on pregnancy and offspring. NIDA Res Monogr 149:16-38.

(18) Bracken MB, Holford TR. (1981). Exposure to prescribed drugs in pregnancy and association with congenital malformations. Obstetrics and gynecology 58 (3): 336–44.

(19) King CR. (1986). Genetic counseling for teratogen exposure. Obstet Gynecol 67(6):843-6.

(20) Jentink J, Loane MA, Dolk H, Barisic I, Garne E, Morris JK, de Jong-van den Berg LT; EUROCAT Antiepileptic Study Working Group. (2010). Valproic acid monotherapy in pregnancy and major congenital malformations. N Engl J Med. 10;362(23):2185-93.

(21) Balsells M, García-Patterson A, Gich I, Corcoy R. (2012). Major congenital malformations in women with gestational diabetes mellitus: a systematic review and meta-analysis. Diabetes Metab Res Rev 28(3):252-7.

(22) Balsells M, García-Patterson A, Gich I, Corcoy R. (2009). Maternal and fetal outcome in women with type 2 versus type 1 diabetes mellitus: a systematic review and metaanalysis. J Clin Endocrinol Metab 94(11):4284-91.

(23) Bánhidy F, Puhó EH, Czeizel AE. (2011). Possible association between hyperthyroidism in pregnant women and obstructive congenital abnormalities of urinary tract in their offspring--a population-based case-control study. J Matern Fetal Neonatal Med 24(2):305-12.

(24) Li Y, Shan Z, Teng W, Yu X, Li Y, Fan C, Teng X, Guo R, Wang H, Li J, Chen Y, Wang W, Chawinga M, Zhang L, Yang L, Zhao Y, Hua T. (2010). Abnormalities of maternal thyroid function during pregnancy affect neuropsychological development of their children at 25-30 months. Clin Endocrinol 72(6):825-9.

(25) Henrichs J, Bongers-Schokking JJ, Schenk JJ, Ghassabian A, Schmidt HG, Visser TJ, Hooijkaas H, de Muinck Keizer-Schrama SM, Hofman A, Jaddoe VV, Visser W, Steegers EA, Verhulst FC, de Rijke YB, Tiemeier H. (2010). Maternal thyroid function during early pregnancy and cognitive functioning in early childhood: the generation R study. J Clin Endocrinol Metab 95(9):4227-34.

(26) Prick BW, Hop WC, Duvekot JJ. (2012). Maternal phenylketonuria and hyperphenylalaninemia in pregnancy: pregnancy complications and neonatal sequelae in untreated and treated pregnancies. Am J Clin Nutr 95(2):374-82.

(27) Lee PJ, Ridout D, Walter JH, Cockburn F. (2005). Maternal phenylketonuria: report from the United Kingdom Registry 1978-97. Arch Dis Child 90(2):143-6.

(28) Walker SP, Permezel M, Berkovic SF. (2009). The management of epilepsy in pregnancy. BJOG 116(6):758-67.

(29) Tomson T, Perucca E, Battino D. (2004). Navigating toward fetal and maternal health: the challenge of treating epilepsy in pregnancy. Epilepsia 45(10):1171-5.

(30) Goldenberg RL, Culhane JF, Johnson DC. (2005). Maternal infection and adverse fetal and neonatal outcomes. Clin Perinatol 32(3):523-59. Review.

(31) Lazzarotto T, Guerra B, Gabrielli L, Lanari M, Landini MP. (2011). Update on the prevention, diagnosis and management of cytomegalovirus infection during pregnancy.Clin Microbiol Infect 17(9):1285-93.

(32) Feldman DM, Timms D, Borgida AF. (2010). Toxoplasmosis, parvovirus, and cytomegalovirus in pregnancy. Clin Lab Med 30(3):709-20.

(33) Mattson SN, Jones KL, Gramling LJ, Schonfeld AM, Riley EP, Harris JA, Chambers CD. (2003). Neurodevelopmental follow-up of children of women infected with varicella during pregnancy: a prospective study. Pediatr Infect Dis J 22(9):819-23.

(34) Vissers LE, de Vries BB, Osoegawa K, Janssen IM, Feuth T, et al. (2003). Array-based comparative genomic hybridization for the genome wide detection of submicroscopic chromosomal abnormalities. Am J Hum Genet 73(6):1261-70.

(35) Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, et al. (2004). Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. J Med Genet 41(4):241-8.

(36) Zweier, C., Sticht, H., Aydin-Yaylagul, I., Campbell, C. E., Rauch, A. (2007). Human TBX1 missense mutations cause gain of function resulting in the same phenotype as 22q11.2 deletions. Am J Hum Genet 80: 510-517.

(37) Hoischen, A., van Bon, B. W. M., Rodriguez-Santiago, B., Gilissen, C., Vissers, L. E. L. M., de Vries, P., Janssen, I., van Lier, B., Hastings, R., Smithson, S. F., Newbury-Ecob, R., Kjaergaard, S, Goodship J, McGowan R, Bartholdi D, Rauch A, Peippo M, Cobben JM, Wieczorek D, Gillessen-Kaesbach G, Veltman JA, Brunner HG, de Vries BB. (2011). De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. Nature Genet 43: 729-731.

(38) Gill HK, Splitt M, Sharland GK, Simpson JM. (2003). Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. J Am Coll Cardiol 3;42(5):923-9.

(39) Holmes-Siedle M, Lindenbaum RH, Galliard A. (1992). Recurrence of neural tube defect in a group of at risk women: a 10 year study of Pregnavite Forte F. J Med Genet 29(2):134-5.

(40) Schwaderer AL, Bates CM, McHugh KM, McBride KL. (2007). Renal anomalies in family members of infants with bilateral renal agenesis/adysplasia. Pediatr Nephrol 22(1):52-6.

(41) Bristow J. (1995). The search for genetic mechanisms of congenital heart disease. Cell Mol Biol Res 41(4):307-19.

(42) Barber RC, Lammer EJ, Shaw GM, Greer KA, Finnell RH. (1999). The role of folate transport and metabolism in neural tube defect risk. Mol Genet Metab 66(1):1-9.

(43) Cobourne MT. (2004). The complex genetics of cleft lip and palate. Eur J Orthod 26(1):7-16.

(44) McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, Lam W, Leal SM, Kaplan N, Schliekelman P, Towbin JA, Belmont JW. (2005). Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: Segregation, multiplex relative risk, and heritability. Am J Med Genet A 134A(2):180-6.

(45) Popp S, Schulze B, Granzow M, Keller M, Holtgreve-Grez H, et al. (2002). Study of 30 patients with unexplained developmental delay and dysmorphic features or congenital abnormalities using conventional cytogenetics and multiplex FISH telomere (M-TEL) integrity assay. Hum Genet 111(1):31-9.

(46) Rooms L, Reyniers E, van Luijk R, Scheers S, Wauters J, et al. (2004). Subtelomeric deletions detected in patients with idiopathic mental retardation using multiplex ligation-dependent probe amplification (MLPA). Hum Mutat 23(1):17-21.

(47) Bendavid C, Dubourg C, Pasquier L, Gicquel I, Le Gallou S, et al. (2007). MLPA screening reveals novel subtelomeric rearrangements in holoprosencephaly. Hum Mutat 28(12):1189-97.

(48) Schaeffer AJ, Chung J, Heretis K, Wong A, Ledbetter DH, et al. (2004). Comparative genomic hybridization-array analysis enhances the detection of aneuploidies and submicroscopic imbalances in spontaneous miscarriages. Am J Hum Genet 74(6):1168-74.

(49) Le Caignec C, Boceno M, Saugier-Veber P, Jacquemont S, Joubert M, et al. (2005). Detection of genomic imbalances by array based comparative genomic hybridisation in fetuses with multiple malformations. J Med Genet 42(2):121-8.

(50) Vialard F, Molina Gomes D, Leroy B, Quarello E, Escalona A, et al. (2009). Array comparative genomic hybridization in prenatal diagnosis: another experience. Fetal Diagn Ther 25(2):277-84.

(51) Momma K. (2010). Cardiovascular anomalies associated with chromosome 22q11.2 deletion syndrome. Am J Cardiol 105(11):1617-24.

(52) Webber SA, Hatchwell E, Barber JC, Daubeney PE, Crolla JA, et al. (1996). Importance of microdeletions of chromosomal region 22q11 as a cause of selected malformations of the ventricular outflow tracts and aortic arch: a three-year prospective study. J Pediatr 129(1):26-32.

(53) Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, et al. (2008). Recurrent reciprocal 1q21.1 deletions and duplications associated

with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat Genet 40(12):1466-71.

(54) Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, et al. (2009). De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. Nat Genet 41(8):931-5.

(55) Erdogan F, Larsen LA, Zhang L, Tümer Z, Tommerup N, et al. (2008). High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. J Med Genet 45(11):704-9.

(56) Magenis RE, Toth-Fejel S, Allen LJ, Black M, Brown MG, Budden S, Cohen R, Friedman JM, Kalousek D, Zonana J, et al. (1990). Comparison of the 15q deletions in Prader-Willi and Angelman syndromes: specific regions, extent of deletions, parental origin, and clinical consequences. Am J Med Genet 35(3):333-49.

(57) Locke AE, Dooley KJ, Tinker SW, Cheong SY, Feingold E, Allen EG, Freeman SB, Torfs CP, Cua CL, Epstein MP, Wu MC, Lin X, Capone G, Sherman SL, Bean LJ. (2010). Variation in folate pathway genes contributes to risk of congenital heart defects among individuals with Down syndrome. Genet Epidemiol 34(6):613-23.

(58) Nie Y, Gu H, Gong J, Wang J, Gong D, Cong X, Chen X, Hu S. (2011). Methylenetetrahydrofolate reductase C677T polymorphism and congenital heart disease: a meta-analysis. Clin Chem Lab Med.

(59) Waterham, H. R., Wijburg, F. A., Hennekam, R. C. M., Vreken, P., Poll-The, B. T., Dorland, L., Duran, M., Jira, P. E., Smeitink, J. A. M., Wevers, R. A., Wanders, R. J. A. (1998). Smith-Lemli-Opitz syndrome is caused by mutations in the 7-dehydrocholesterol reductase gene. Am J Hum Genet 63: 329-338.

(60) Stankiewicz, P., Sen, P., Bhatt, S. S., Storer, M., Xia, Z., Bejjani, B. A., Ou, Z., Wiszniewska, J., Driscoll, D. J., Maisenbacher, M. K., Bolivar, J., Bauer, M., and 32 others. (2009). Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. Am J Hum Genet 84: 780-791.

(61) Krantz, I. D., McCallum, J., DeScipio, C., Kaur, M., Gillis, L. A., Yaeger, D., Jukofsky, L., Wasserman, N., Bottani, A., Morris, C. A., Nowaczyk, M. J. M., Toriello, H., and 9 others. (2004). Cornelia de Lange syndrome is caused

by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nature Genet 36: 631-635.

(62) Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., Rothrock, C. R., Eapen, R. S., Hirayama-Yamada, K., Joo, K., Matsuoka, R., Cohen, J. C., Srivastava, D. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature 424: 443-447.

(63) Watanabe, Y., Benson, D. W., Yano, S., Akagi, T., Yoshino, M., Murray, J. C. (2002). Two novel frameshift mutations in NKX2.5 result in novel features including visceral inversus and sinus venosus type ASD. J Med Genet 39: 807-811.

(64) Gandelman, K.-Y., Gibson, L., Meyn, M. S., Yang-Feng, T. L. (1992). Molecular definition of the smallest region of deletion overlap in the Wolf-Hirschhorn syndrome. Am J Hum Genet 51: 571-578.

(65) Kelley, R. I., Zackai, E. H., Emanuel, B. S., Kistenmacher, M., Greenberg, F., Punnett, H. H. (1982). The association of the DiGeorge anomalad with partial monosomy of chromosome 22. J Pediat 101: 197-200.

(66) Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., Beck, A. E., Tabor, H. K., Cooper, G. M., Mefford, H. C., Lee, C., Turner, E. H., and 9 others. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nature Genet 42: 790-793.

(67) Schorry, E. K., Keddache, M., Lanphear, N., Rubinstein, J. H., Srodulski, S., Fletcher, D., Blough-Pfau, R. I., Grabowski, G. A. (2008). Genotype-phenotype correlations in Rubinstein-Taybi syndrome. Am. J. Med. Genet. 146A: 2512-2519.

(68) Mitchell SC, Sellmann AH, Westphal MC, Park J. (1971). Etiologic correlates in a study of congenital heart disease in 56,109 births. Am J Cardiol 28(6):653-7.

(69) Richards AA, Santos LJ, Nichols HA, Crider BP, Elder FF, et al. (2008). Cryptic chromosomal abnormalities identified in children with congenital heart disease. Pediatr Res 64(4):358-63.

(70) Hoffman JI, Kaplan S. (2002). The incidence of congenital heart disease. J Am Coll Cardiol 19;39(12):1890-900.

(71) Nelle M, Raio L, Pavlovic M, Carrel T, Surbek D, Meyer-Wittkopf M. (2009). Prenatal diagnosis and treatment planning of congenital heart defects-possibilities and limits. World J Pediatr 5(1):18-22.

(72) Jaeggi ET, Sholler GF, Jones OD, Cooper SG. (2001). Comparative analysis of pattern, management and outcome of pre- versus postnatally diagnosed major congenital heart disease: a population-based study. Ultrasound Obstet Gynecol 17(5):380-5.

(73) Hickey EJ, Caldarone CA, McCrindle BW. (2012). Left ventricular hypoplasia: a spectrum of disease involving the left ventricular outflow tract, aortic valve, and aorta. J Am Coll Cardiol 3;59(1 Suppl):S43-54.

(74) McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, Lam W, Leal SM, Kaplan N, Schliekelman P, Towbin JA, Belmont JW. (2005). Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: Segregation, multiplex relative risk, and heritability. Am J Med Genet A 15;134A(2):180-6.

(75) Rose V, Keith JD. (1966). The prevalence of ventricular septal defect in elementary school children in the City of Toronto. Can Med Assoc J 95(22):1132-4.

(76) De Luca A, Sarkozy A, Ferese R, Consoli F, Lepri F, Dentici ML, Vergara P, De Zorzi A, Versacci P, Digilio MC, Marino B, Dallapiccola B. (2011). New mutations in ZFPM2/FOG2 gene in tetralogy of Fallot and double outlet right ventricle. Clin Genet 80(2):184-90.

(77) Dubourg C, Bendavid C, Pasquier L, Henry C, Odent S, David V. (2007). Holoprosencephaly. Orphanet J Rare Dis 2;2:8.

(78) Roessler E, Muenke M. (2010). The molecular genetics of holoprosencephaly. Am J Med Genet C Semin Med Genet 154C(1):52-61.

(79) Mercier S, Dubourg C, Garcelon N, Campillo-Gimenez B, Gicquel I, Belleguic M, Ratié L, Pasquier L, Loget P, Bendavid C, Jaillard S, Rochard L, Quélin C, Dupé V, David V, Odent S. (2011). New findings for phenotype-genotype correlations in a large European series of holoprosencephaly cases. J Med Genet 48(11):752-60.

(80) Weichert J, Hartge D, Krapp M, Germer U, Gembruch U, Axt-Fliedner R. (2010). Prevalence, characteristics and perinatal outcome of fetal ventriculomegaly in 29,000 pregnancies followed at a single institution. Fetal Diagn Ther 27(3):142-8.

(81) Cohen FL. (1987). Neural tube defects: epidemiology, detection, and prevention. J Obstet Gynecol Neonatal Nurs 16(2):105-15.

(82) Northrup H, Volcik KA. (2000). Spina bifida and other neural tube defects. Curr Probl Pediatr 30(10):313-32.

(83) Sandler AD. (2010). Children with spina bifida: key clinical issues. Pediatr Clin North Am 57(4):879-92.

(84) Eubanks JD, Cheruvu VK. (2009). Prevalence of sacral spina bifida occulta and its relationship to age, sex, race, and the sacral table angle: an anatomic, osteologic study of three thousand one hundred specimens. Spine (Phila Pa 1976) 1;34(15):1539-43.

(85) Piatt JH Jr. (2010). Treatment of myelomeningocele: a review of outcomes and continuing neurosurgical considerations among adults. J Neurosurg Pediatr 6(6):515-25.

(86) Ueno N, Greene ND. (2003). Planar cell polarity genes and neural tube closure. Birth Defects Res C Embryo Today 69(4):318-24.

(87) Robinson A, Escuin S, Doudney K, Vekemans M, Stevenson RE, Greene ND, Copp AJ, Stanier P. (2012). Mutations in the planar cell polarity genes CELSR1 and SCRIB are associated with the severe neural tube defect craniorachischisis. Hum Mutat 33(2):440-7.

(88) Shi Y, Ding Y, Lei YP, Yang XY, Xie GM, Wen J, Cai CQ, Li H, Chen Y, Zhang T, Wu BL, Jin L, Chen YG, Wang HY. (2012). Identification of novel rare mutations of DACT1 in human neural tube defects. Hum Mutat.

(89) Bosoi CM, Capra V, Allache R, Trinh VQ, De Marco P, Merello E, Drapeau P, Bassuk AG, Kibar Z. (2011). Identification and characterization of novel rare mutations in the planar cell polarity gene PRICKLE1 in human neural tube defects. Hum Mutat 32(12):1371-5.

(90) Yu H, Smallwood PM, Wang Y, Vidaltamayo R, Reed R, Nathans J. (2010). Frizzled 1 and frizzled 2 genes function in palate, ventricular septum and neural tube closure: general implications for tissue fusion processes. Development 137(21):3707-17.

(91) Blom, H.J., Shaw, G.M., Den Heijer, M. and Finnell, R.H. (2006). Neural tube defects and folate: case far from closed. Nat Rev Neurosci 7, 724–731.

(92) Beaudin, A.E. and Stover, P.J. (2009). Insights into metabolic mechanisms underlying folate-responsive neural tube defects: a mini review. Birth Defects Res. A Clin Mol Teratol 85, 274–284.

(93) Song R, Yosypiv IV. (2010). Genetics of congenital anomalies of the kidney and urinary tract. Pediatr Nephrol 26(3):353-64.

(94) Weber S, Moriniere V, Knüppel T, Charbit M, Dusek J, Ghiggeri GM, Jankauskiené A, Mir S, Montini G, Peco-Antic A, Wühl E, Zurowska AM, Mehls O, Antignac C, Schaefer F, Salomon R. (2006). Prevalence of mutations in renal developmental genes in children with renal hypodysplasia: results of the ESCAPE study. J Am Soc Nephrol 17(10):2864-70.

(95) Skinner MA, Safford SD, Reeves JG, Jackson ME, Freemerman AJ. (2008). Renal aplasia in humans is associated with RET mutations. Am J Hum Genet 82(2):344-51.

(96) Nakayama M, Nozu K, Goto Y, Kamei K, Ito S, Sato H, Emi M, Nakanishi K, Tsuchiya S, Iijima K. (2010). HNF1B alterations associated with congenital anomalies of the kidney and urinary tract. Pediatr Nephrol 25(6):1073-9.

(97) Weber S, Taylor JC, Winyard P, Baker KF, Sullivan-Brown J, Schild R, Knüppel T, Zurowska AM, Caldas-Alfonso A, Litwin M, Emre S, Ghiggeri GM, Bakkaloglu A, Mehls O, Antignac C, Network E, Schaefer F, Burdine RD. (2008). SIX2 and BMP4 mutations associate with anomalous kidney development. J Am Soc Nephrol 19(5):891-903.

(98) Harewood L, Liu M, Keeling J, Howatson A, Whiteford M, Branney P, Evans M, Fantes J, Fitzpatrick DR. (2010). Bilateral renal agenesis/hypoplasia/dysplasia (BRAHD): postmortem analysis of 45 cases with breakpoint mapping of two de novo translocations. PLoS One 5(8):e12375.

(99) Klinger G, Merlob P, Aloni D, Maayan A, Sirota L. (1997). Normal pulmonary function in a monoamniotic twin discordant for bilateral renal agenesis: report and review. Am J Med Genet 73(1):76-9.

(100) Mégarbané A, Ravel A, Mircher C, Sturtz F, Grattau Y, Rethoré MO, Delabar JM, Mobley WC. (2009). The 50th anniversary of the discovery of trisomy 21: the past, present, and future of research and treatment of Down syndrome. Genet Med 11(9):611-6.

(101) Antonarakis SE, Lyle R, Dermitzakis ET, Reymond A, Deutsch S. (2004). Chromosome 21 and down syndrome: from genomics to pathophysiology. Nat Rev Genet 5(10):725-38.

(102) Weijerman ME, van Furth AM, van der Mooren MD, van Weissenbruch MM, Rammeloo L, Broers CJ, Gemke RJ. (2010). Prevalence of congenital heart defects and persistent pulmonary hypertension of the neonate with Down syndrome. Eur J Pediatr 169(10):1195-9.

(103) de Rubens Figueroa J, del Pozzo Magaña B, Pablos Hach JL, Calderón Jiménez C, Castrejón Urbina R. (2003). Heart malformations in children with Down syndrome. Rev Esp Cardiol 56(9):894-9.

(104) Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. (2002). Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 30:e57.

(105) Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. (1998). High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20:207-211.

(106) Slavotinek AM. (2008). Novel microdeletion syndromes detected by chromosome microarrays. Hum Genet 124:1-17 83.

(107) Koolen DA, Pfundt R, de Leeuw N, Hehir-Kwa JY, Nillesen WM, Neefs I, Scheltinga I, Sistermans E, Smeets D, Brunner HG, van Kessel AG, Veltman JA, de Vries BB. (2009). Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. Hum Mutat 30:283- 292.

(108) Tucker T, Montpetit A, Chai D, Chan S, Chénier S, Coe BP, Delaney A, Eydoux P, Lam WL, Langlois S, Lemyre E, Marra M, Qian H, Rouleau GA, Vincent D, Michaud JL, Friedman JM. (2011). Comparison of genome-wide array genomic hybridization platforms for the detection of copy number variants in idiopathic mental retardation. BMC Med Genomics 4:25.

(109) Fouse SD, Nagarajan RO, Costello JF. (2010). Genome-scale DNA methylation analysis. Epigenomics 2(1):105-17.

(110) Nygren AO, Ameziane N, Duarte HM, Vijzelaar RN, Waisfisz Q, Hess CJ, Schouten JP, Errami A. (2005). Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res 33(14):e128.

(111) Izzi B, Decallonne B, Devriendt K, Bouillon R, Vanderschueren D, Levtchenko E, de Zegher F, Van den Bruel A, Lambrechts D, Van Geet C, Freson K. (2010). A new approach to imprinting mutation detection in GNAS by Sequenom EpiTYPER system. Clin Chim Acta 411(23-24):2033-9.

(112) Gilissen C, Hoischen A, Brunner HG, Veltman JA. (2012). Disease gene identification strategies for exome sequencing. Eur J Hum Genet 20(5):490-7.

(113) Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12(11):745-55.

(114) Mardis ER. (2008). Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 9:387-402.

(115) Mardis ER. (2008). The impact of next-generation sequencing technology on genetics. Trends Genet 24:133-141.

(116) Wilhelm J, Pingoud A. (2003). Real-time polymerase chain reaction. Chembiochem 4:1120-1128.

(117) Bustin SA, Mueller R. (2005). Real-time reverse transcription PCR (qRTPCR) and its potential use in clinical diagnosis. Clin Sci 109:365-379.

(118) Gates EA. Communicating risk in prenatal genetic testing. (2004). J Midwifery Womens Health 49 (3): 220-7.

(119) Dahl K, Kesmodel U, Hvidman L, Olesen F. (2006). Informed consent: providing information about prenatal examinations. Acta Obstet Gynecol Scand 85 (12): 1420-5.

(120) Bernhardt BA, Biesecker BB, Mastromarino CL. (2000). Goals, Benefits, and Outcomes of Genetic Counseling: Client and Genetic Counselor Assessment. Am J Med Genet 94 :189–197.

(121) Santalahti P, Hemminki E, Latikka AM, Ryynänen M. (1998). Women's decision-making in prenatal screening. Soc Sci Med 46 (8): 1067-76.

(122) van den Berg M, Timmermans DR, Ten Kate LP, van Vugt JM, van der Wal G. (2005). Are pregnant women making informed choices about prenatal screening? Genet Med 7 (5): 332-8.

(123) Agnieszka S, OElêzak R, Pesz K, Gil J, Ssiadek MM. (2007). Prenatal diagnosis - principles of diagnostic procedures and genetic counseling. Folia Histochemica et cytobiologica 45 (1): 11-16.

(124) Vreeburg M, Heitink MV, Damstra RJ, Moog U, van Geel M, et al. (2008). Lymphedema-distichiasis syndrome: a distinct type of primary lymphedema caused by mutations in the FOXC2 gene. Int J Dermatol 47 Suppl 1:52-5.

(125) van Bon BW, Mefford HC, Menten B, Koolen DA, Sharp AJ, et al. (2009). Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from non-pathogenic to a severe outcome. J Med Genet 46(8):511-23.

(126) Lavallee G, Andelfinger G, Nadeau M, Lefebvre C, Nemer G, et al. (2006). The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. Embo J 25(21):5201-13.

(127) Valduga M, Philippe C, Bach Segura P, Thiebaugeorges O, Miton A, et al. (2010). A retrospective study by oligonucleotide array-CGH analysis in 50 fetuses with multiple malformations. Prenat Diagn 30(4):333-41.

(128) Ohtoshi A, Behringer RR. (2004). Neonatal lethality, dwarfism, and abnormal brain development in Dmbx1 mutant mice. Molec Cell Biol 24(17):7548-58.

(129) Chen LT, Gliman AG, Kozasa T. (1999). A candidate target for G protein action in brain. J Biol Chem 274(38):26931-8.

(130) Chen YH, Ishii M, Sun J, Sucov HM, Maxson RE Jr. (2007). Msx1 and Msx2 regulate survival of secondary heart field precursors and post-migratory proliferation of cardiac neural crest in the outflow tract. Dev Biol 308(2):421-37.

(131) Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. (2010). A de novo paradigm for mental retardation. Nat Genet 42(12):1109-12.

(132) Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, Gogos JA, Karayiorgou M. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat Genet 7;43(9):864-8.

(133) Tan TY, Collins A, James PA, McGillivray G, Stark Z, Gordon CT, Leventer RJ, Pope K, Forbes R, Crolla JA, Ganesamoorthy D, Burgess T, Bruno DL, Slater HR, Farlie PG, Amor DJ. (2011). Phenotypic variability of distal 22q11.2 copy number abnormalities. Am J Med Genet A 155A(7):1623-33.

(134) Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, Lalani SR, Graham B, Lee B, Shinawi M, Shen J, Kang SH, Pursley A, Lotze T, Kennedy G, Lansky-Shafer S, Weaver C, Roeder ER, Grebe TA, Arnold GL, Hutchison T, Reimschisel T, Amato S, Geragthy MT, Innis JW, Obersztyn E, Nowakowska B, Rosengren SS, Bader PI, Grange DK, Naqvi S, Garnica AD, Bernes SM, Fong CT, Summers A, Walters WD, Lupski JR, Stankiewicz P, Cheung SW, Patel A. (2008). Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat Genet. 40(12):1466-71.

(135) Li DK, Karlberg K, Wi S, Norem C. Factors influencing women's acceptance of prenatal screening tests. Prenat Diagn 2008: 28 (12): 1136-43.

(136) Darilek S, Ward P, Pursley A et al. Pre- and postnatal genetic testing by array-comparative genomic hybridization: genetic counseling perspectives. Genet Med 2008: 10 (1): 13-18.

(137) Santalahti P, Hemminki E, Latikka AM, Ryynänen M. Women's decision-making in prenatal screening. Soc Sci Med 1998: 46 (8): 1067-76.

(138) van den Berg M, Timmermans DR, Ten Kate LP, van Vugt JM, van der Wal G. Are pregnant women making informed choices about prenatal screening? Genet Med 2005: 7 (5): 332-8.

(139) Narisawa A, Komatsuzaki S, Kikuchi A, Niihori T, Aoki Y, Fujiwara K, Tanemura M, Hata A, Suzuki Y, Relton CL, Grinham J, Leung KY, Partridge D, Robinson A, Stone V, Gustavsson P, Stanier P, Copp AJ, Greene ND, Tominaga T, Matsubara Y, Kure S. (2012). Mutations in genes encoding the glycine cleavage system predispose to neural tube defects in mice and humans. Hum Mol Genet 21(7):1496-503.

(140) Ichida F, Tsubata S, Bowles KR, Haneda N, Uese K, Miyawaki T, Dreyer WJ, Messina J, Li H, Bowles NE, Towbin JA. (2001). Novel gene mutations in patients with left ventricular noncompaction or Barth syndrome. Circulation 103: 1256-1263.

(141) Kosaki, R., Gebbia, M., Kosaki, K., Lewin, M., Bowers, P., Towbin, J. A., Casey, B. (1999). Left-right axis malformations associated with mutations in ACVR2B, the gene for human activin receptor type IIB. Am J Med Genet 82: 70-76.

(142) Ma L, Selamet Tierney ES, Lee T, Lanzano P, Chung WK. (2012). Mutations in ZIC3 and ACVR2B are a common cause of heterotaxy and associated cardiovascular anomalies. Cardiol Young 22(2):194-201.

(143) Akhavantabasi S, Akman HB, Sapmaz A, Keller J, Petty EM, Erson AE. (2010). USP32 is an active, membrane-bound ubiquitin protease overexpressed in breast cancers. Mamm Genome 21(7-8):388-97.

(144) Ono T, Losada A, Hirano M, Myers MP, Neuwald AF, Hirano T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell 115: 109-121.

(145) Grossfeld PD, Mattina T, Lai Z, Favier R, Jones KL, Cotter F, Jones C. (2004). The 11q terminal deletion disorder: a prospective study of 110 cases. Am J Med Genet 129A:51-61.

(146) Ye M, Hamzeh R, Geddis A, Varki N, Perryman MB, Grossfeld P. (2009). Deletion of JAM-C, a candidate gene for heart defects in Jacobsen syndrome, results in a normal cardiac phenotype in mice. Am J Med Genet A 149A(7):1438-43.

(147) Nicole, S., Davoine, C.-S., Topaloglu, H., Cattolico, L., Barral, D., Beighton, P., Ben Hamida, C., Hammouda, H., Cruaud, C., White, P. S., Samson, D., Urtizberea, J. A., Lehmann-Horn, F., Weissenbach, J., Hentati, F., Fontaine, B. (2000). Perlecan, the major proteoglycan of basement membranes, is altered in patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia). Nature Genet 26: 480-483.

(148) Arikawa-Hirasawa, E., Wilcox, W. R., Le, A. H., Silverman, N., Govindraj, P., Hassell, J. R., Yamada, Y. (2001). Dyssegmental dysplasia, Silverman-Handmaker type, is caused by functional null mutations of the perlecan gene. Nature Genet 27: 431-434.

(149) Gan B, Peng X, Nagy T, Alcaraz A, Gu H, Guan JL. (2006). Role of FIP200 in cardiac and liver development and its regulation of TNF-alpha and TSC-mTOR signaling pathways. J Cell Biol 175: 121-133.

(150) Risebro CA, Searles RG, Melville AA, Ehler E, Jina N, Shah S, Pallas J, Hubank M, Dillard M, Harvey NL, Schwartz RJ, Chien KR, Oliver G, Riley PR. (2009). Prox1 maintains muscle structure and growth in the developing heart. Development 136(3):495-505.

(151) Gill HK, Parsons SR, Spalluto C, Davies AF, Knorz VJ, Burlinson CE, Ng BL, Carter NP, Ogilvie CM, Wilson DI, Roberts RG. (2009). Separation of the PROX1 gene from upstream conserved elements in a complex

inversion/translocation patient with hypoplastic left heart. Eur J Hum Genet 17(11):1423-31.

(152) Jang JW, Lee WY, Lee JH, Moon SH, Kim CH, Chung HM. (2011). A novel Fbxo25 acts as an E3 ligase for destructing cardiac specific transcription factors. Biochem Biophys Res Commun 410(2):183-8.

(153) Suriben R, Kivimäe S, Fisher DA, Moon RT, Cheyette BN. (2009). Posterior malformations in Dact1 mutant mice arise through misregulated Vangl2 at the primitive streak. Nat Genet 41(9):977-85.

LIST OF ACRONYMS

aCGH	array Comparative Genomic Hybridization
ASD	auricular septal defects
BAC	Bacterial Artificial Chromosome
CHD	Congenital Heart Defect
CMA	Chromosomal Microarray Analysis
CNS	Central Nervous System
CNV	Copy Number Variation
dbSNP	Database of Single Nucleotide Polymorphisms
DNA	Deoxyribonucleic Acid
EVS	Exome Variant Server
FISH	fluorescence in Situ Hybridization
HLH/LHH	Hypoplastic Left Heart / Left Heart Hypoplasia
iCHD	Isolated congenital heart defect
Indel	Insertion / Deletion
LHH/HLH	Left Heart Hypoplasia / Hypoplastic Left Heart
LOH	Lost of Heterozigosity
MAF	Minor Allele Frequency
MALDI-TOFF	Matrix-assisted Laser Desorption/Ionization-Time of Flight
Mb	Megabase
MCA / MR	Multiple Congenital Anomalies / Mental Retardation
MLPA	Multiplex Ligation-dependent Probe Amplification
MS-MLPA	Methylation Sspecific Multiplex Ligation-dependent Probe Amplification
NGS	Next Generation Sequencing
NTD	Neural Tube Defect
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
UPD	Uniparental Disomy
VSD	Ventricular Septal Defect

GLOSSARY

Aortic valve stenosis: the aortic valve (guards the opening between the left ventricle and the aorta) becomes partially obstructed, leading to significant heart problems. When somebody develops aortic stenosis, the aortic valve no longer opens completely, so the heart must work harder to eject blood.

Aqueductal stenosis: the cerebral aqueduct (of Sylvius) is a narrow channel that connects two of the ventricles and passes through the midbrain. Normal cerebrospinal fluid dynamics require an open aqueduct so that fluid can flow freely through this area. If the aqueduct is blocked, this is called stenosis and can lead to the symptoms of hydrocephalus. Aqueductal stenosis is one of the known causes of hydrocephalus and the most common cause of congenital (present at birth) hydrocephalus.

Bicuspid aortic valve: congenital condition of the aortic valve where two of the aortic valvular leaflets fuse during development resulting in a valve that is bicuspid instead of the normal tricuspid configuration.

Cerebellar tonsils: rounded lobule on the undersurface of each cerebellar hemisphere, continuous medially with the uvula of the vermis. Elongation of the cerebellar tonsils can, due to pressure, lead to this portion of the cerebellum to slip or be pushed through the foramen magnum of the skull resulting in a condition known as Chiari malformation.

Cleft palate: is a birth defect that affects the roof of the mouth. It occurs when the tissues that form the roof of the mouth do not join before birth.

CNV (copy number variation): DNA segment that is 1kb or longer and is present in a variable copy number in comparison with a reference genome.

Coarctation of the aorta: congenital condition whereby the aorta narrows. The coarctation may be discrete, or a long segment of the aorta may be narrowed.

Corpus callosum: thick band of nerve fibers that divides the cerebrum into left and right hemispheres. It connects the left and right sides of the brain allowing the communication between both hemispheres. The corpus callosum transfers motor, sensory, and cognitive information between the brain hemispheres

CpG island: short region of DNA in which the frequency of the CG sequence is higher than in other regions. CpG islands are often located

around the promoters of housekeeping genes or other genes frequently expressed in a cell.

Craniosynostosis: defection of the skull caused by early fusion of one or more cranial sutures. The shape alteration of the cranial vault varies, depending on the fused sutures, so that compensatory growth occurs in dimensions not restricted by sutures.

Cyclopia: congenital abnormality in which there is only one eye centrally placed in the area normally occupied by the root of the nose. There is a missing nose or a nose in the form of a proboscis (a tubular appendage) located above the eye. Cyclopia and milder forms of the same developmental disorder result from holoprosencephaly.

Developmental delay: it is an ongoing major or minor delay in the process of development. It is noticed because the child does not reach the developmental milestones at the expected times. Delay can occur in one or many areas—for example, gross or fine motor, language, social, or thinking skills.

Dilated cardiomyopathy: progressive disease of heart muscle that is characterized by ventricular chamber enlargement and contractile dysfunction with normal left ventricular wall thickness. Dilated cardiomyopathy is the third most common cause of heart failure and the most frequent reason for heart transplantation.

Duodenal atresia: condition in which the first part of the small bowel (the duodenum) has not developed properly. It is not open and cannot allow the passage of stomach contents.

Gel electrophoresis: method used in biochemistry and molecular biology to separate a mixed population of DNA and RNA fragments by length, to estimate the size of DNA and RNA fragments or to separate proteins by charge. Nucleic acid molecules are separated by applying an electric field to move the negatively charged molecules through an agarose matrix. Shorter molecules move faster and migrate farther than longer ones because shorter molecules migrate more easily through the pores of the gel.

Epicanthus: vertical fold of skin on either side of the nose, sometimes covering the inner canthus. Some infants with Down syndrome have marked epicanthal folds.

Foramen magnum: hole in the bottom of the skull through which the spinal cord passes in order to connect to the brain by merging with the brain's lowermost portion, the medulla oblongata. Several other nerves, blood vessels and tendons also pass through the foramen magnum.

Forebrain: the largest part of the brain, most of which is made up of the cerebrum. Other important structures found in the forebrain include the thalamus, the hypothalamus and the limbic system.

Heritability: proportion of observable differences between individuals that is due to genetic differences.

Hirschsprung disease: blockage of the large intestine due to improper muscle movement in the bowel due to the absence of nerves from a part of the bowel. Areas without such nerves cannot push material through.

Hydranencephaly: rare condition in which the brain's cerebral hemispheres are absent and replaced by sacs filled with cerebrospinal fluid.

Hydrocephaly/Hydrocephalus: buildup of too much cerebrospinal fluid in the brain which puts harmful pressure on the brain.

Hypospadias: birth defect in which the opening of the urethra is on the underside, rather than at the end, of the penis.

Hypotelorism: abnormal decrease in the intraorbital distance.

Hypothalamic dysfunction: disorder of the region of the brain called the hypothalamus, which helps control the pituitary gland and regulate many body functions.

Hypotonia: decreased muscle tone. Infants with hypotonia seem floppy, with their elbows and knees loosely extended. They may have poor or no head control.

Lipomeningocele: form of occult spinal dysraphism in which there is intact skin covering the defect. The spinal cord remains within the spinal canal, with the junction between the spinal cord and the lipoma also residing within the canal. In general, individuals with lipomeningoceles are normal at birth and neurologic findings are first noted during the second year of life, with most patients exhibiting some neurologic deficits by early childhood.

Long Qt syndrome: disorder of the heart's electrical activity. It can cause sudden, uncontrollable, dangerous arrhythmias in response to exercise or stress. The term "long QT" refers to an abnormal pattern seen on an electrocardiogram.

MALDI-TOFF (Matrix-assisted laser desorption/ionisation-time of flight mass spectrometry): combination of laser ionization to fragment the molecules analyze (MALDI) with a mass detector based on time of flight (TOF) in order to analyze biomolecules, such as DNA, proteins, peptides and sugars.

Non-disjunction in meiosis: failure of two members of a chromosome pair to separate from one another during meiosis, causing both chromosomes to go to a single daughter cell. Meiotic non-disjunction can be responsible for the extra chromosome 21 in trisomy 21 and for extra and missing chromosomes that cause other birth defects and many miscarriages.

Microcephaly: condition in which a person's head is significantly smaller than normal for their age and sex, based on standardized charts. Head size is measured as the distance around the top of the head. Microcephaly most often occurs because the brain fails to grow at a normal rate.

Polydactyly: condition in which a person has more than five fingers per hand or five toes per foot. Having an abnormal number of digits (6 or more) can occur on its own, without any other symptoms or disease.

Potter sequence: constellation of findings demonstrated postnatally as a consequence of severe, prolonged oligohydramnios in utero. It consists of pulmonary hypoplasia (often severe and incompatible with life), growth restriction, abnormal facies (including low set ears, flattened nose, wrinkled skin and micrognathia) and limb abnormalities (including club feet and contractures).

Proboscis: rare anomaly where an anterior tubular appendage-like structure is seen projecting from about the midline fetal face / forehead. The presence of a proboscis can be associated with several anomalies which include holoprosencephaly.

Prosencephalon: part of the brain developed from the anterior of the three primary brain vesicles, comprising the diencephalon and telencephalon. The segment of the adult brain that develops from the embryonic forebrain includes the brain, thalamus, and hypothalamus.

Pulmonic stensosis: dynamic or fixed anatomic obstruction to flow from the right ventricle to the pulmonary arterial vasculature.

Recurrent genomic disorder: an increasing number of human diseases are recognized to result from recurrent DNA rearrangements involving unstable genomic regions. These are termed genomic disorders, in which the clinical phenotype is a consequence of abnormal dosage of gene(s) located within the rearranged genomic fragments. Both inter- and intrachromosomal rearrangements are facilitated by the presence of region-specific low-copy repeats (LCRs) and result from nonallelic homologous recombination (NAHR) between paralogous genomic segments. LCRs usually span ~10–400 kb of genomic DNA, share ≥ 97% sequence identity, and provide the substrates for homologous recombination, thus predisposing the region to rearrangements. Moreover, it has been suggested that higher order genomic architecture involving LCRs plays a significant role in karyotypic evolution accompanying primate speciation.

SNP (Single nucleotide polymorphisms): DNA sequence variation occurring when a single nucleotide in the genome differs between members of the same species. The nucleotide variation has to be present, at least, in 1% of the population.

Tethered cord: disorder in which a child's spinal cord is pulled down and stuck, or fixed, to the spinal canal. This can permanently damage the spinal nerves.

Wolff-Parkinson-White syndrome: heart condition in which there is an extra electrical pathway in the heart. The condition can lead to episodes of rapid heart rate (tachycardia). Wolff-Parkinson-White is one of the most common causes of fast heart rate disorders in infants and children.