

From the Department of Molecular Medicine and Surgery,
Karolinska Institutet, Stockholm, Sweden

GENETIC STUDIES OF NEURODEVELOPMENTAL DISORDERS

Josephine Wincent



**Karolinska
Institutet**

Stockholm 2012

All previously published papers were reproduced with permission from the publisher.

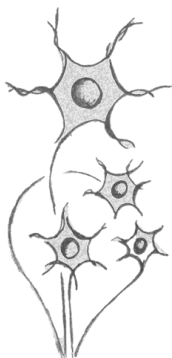
Published by Karolinska Institutet. Printed by Larseric Digital Print AB

© Josephine Wincent, 2012

ISBN 978-91-7457-715-0

*”Om du inte var en sån där liten rar
och ful bleknosning med skeva ben,
så var du ju inte min Skorpa,
den som jag tycker om.”*

*Jonatan till Skorpan i Bröderna Lejonhjärta
av Astrid Lindgren.*



ABSTRACT

Neurodevelopmental disorders (NDDs) constitute a heterogeneous group of disorders that adversely impacts a child's behavioural and learning processes. Developmental delay (DD) and mental retardation are included among the NDDs and are frequently associated with a wide range of accompanying disabilities such as multiple congenital anomalies and dysmorphic features. Despite extensive clinical and laboratory investigation, the cause of the patient's symptoms remains unknown in approximately half of the cases. For the children's families this is often frustrating since an aetiological diagnosis not only gives an explanation of why the child has symptoms but may also provide better prognosis evaluation, adequate genetic counselling and enable prenatal diagnosis. In approximately 20% of patients, a clear genetic cause can be found, including both single-gene disorders and chromosomal disorders.

In **paper I** a *NIPBL* and *SMC1L1* mutation screening by direct sequencing and MLPA was performed in a group of nine index patients diagnosed with Cornelia de Lange syndrome (CdLS), which is characterized by severe mental and growth retardation and distinctive dysmorphic facial features. We identified seven *NIPBL* mutations and showed that a splice-site mutation lead to skipping of an exon. A clear genotype-phenotype correlation was not found.

In **paper II** sequencing and MLPA analysis revealed 18 *CHD7* mutations in 28 index patients with CHARGE syndrome. In addition, inherited variants were identified and clinical interpretation of these are discussed. Our results indicate that hypoplastic semicircular canals is not obligatory for a *CHD7* mutation, although we agree that it is the most frequent and specific sign of CHARGE syndrome. A *CHD7* mutation was found in a patient not fulfilling clinical criteria showing that also atypical patients benefit from testing.

Paper I and II confirm that *NIPBL* and *CHD7* are the main causative genes for CdLS and CHARGE syndrome respectively. However, in >30% of our patients no causal mutation could be detected. Whole genome-/exome sequencing might find new causative genes and/or mutations in non-coding sequences of known genes.

The patient described in **paper III** had an 18.2 Mb *de novo* deletion of chromosome 11q13.4-q14.3. By comparing his phenotype to the few previously described patients, we show that a common phenotype for patients with deletions in this region might be emerging, comprising mild-moderate DD, a sociable personality and dysmorphic facial features.

The implementation of high-resolution array-CGH over the last decade has enabled the genome-wide identification of submicroscopic copy number variations (CNVs) in patients with NDDs. In **study IV** we wanted to evaluate array-CGH as a diagnostic tool in our clinical laboratory. In the 160 investigated patients, 21 (13,1%) causal CNVs and 15 (9,4%) CNVs of unclear clinical significance were detected. Standard karyotyping had in seven cases failed to detect causal CNVs ≥ 5 Mb, five of which were ≥ 10 Mb, emphasizing that more reliable methods were needed to exclude CNVs in these patients. Array-CGH proved to be very useful and became recommended as the first step investigation for patients with idiopathic DD. However, increasing the resolution of a whole genome screen in the diagnostic setting has its drawback of detecting an increased number of CNVs of unclear clinical significance.

In **paper V** we report on the clinical and molecular characterization of 16 individuals with distal 22q11.2 duplications. The patients displayed a variable phenotype, and many of the duplications were inherited (83%). The possible pathogenicity of these duplications is discussed and we conclude that it is likely that distal 22q11.2 duplications represent a susceptibility/risk locus for NDDs rather than being causal variants. Additional genetic, epigenetic or environmental factors are likely required to cause a phenotype. Five patients had additional CNVs of unclear clinical significance making a 2-hit event plausible.

Paper IV and V illustrate that the identification of CNVs of uncertain clinical significance puts new demands on genetic counselling and continuous research and submission of cases to databases are still important.

Future challenges include how to deal with the interpretation of multiple rare variants in one individual and to find ways to estimate how great a risk factor certain CNVs, such as distal 22q11.2 duplications, actually are for a phenotypic effect.

LIST OF PUBLICATIONS

- I. Schoumans J, **Wincent J**, Barbaro M, Djureinovic T, Maguire P, Forsberg L, Staaf J, Thuresson AC, Borg A, Nordgren A, Malm G and Anderlid BM.
Comprehensive mutational analysis of a cohort of Swedish Cornelia de Lange syndrome patients.
European Journal of Human Genetics 2007, 15:143-149.
- II. **Wincent J**, Holmberg E, Stromland K, Soller M, Mirzaei L, Djureinovic T, Robinson K, Anderlid BM, and Schoumans J.
CHD7 mutation spectrum in 28 Swedish patients diagnosed with CHARGE syndrome.
Clinical Genetics 2008, 74:31-38.
- III. **Wincent J**, Schoumans J, and Anderlid BM.
De novo deletion of chromosome 11q13.4-q14.3 in a boy with microcephaly, ptosis and developmental delay.
European Journal of Medical Genetics 2010, 53:50-53.
- IV. **Wincent J**, Anderlid BM, Lagerberg M, Nordenskjold M, and Schoumans J.
High-resolution molecular karyotyping in patients with developmental delay and/or multiple congenital anomalies in a clinical setting.
Clinical Genetics 2010, 79:147-157.
- V. **Wincent J**, Bruno DL, van Bon BW, Bremer A, Stewart H, Bongers EM, Ockeloen CW, Willemsen MH, Keays DD, Baird G, Newbury DF, Kleefstra T, Marcelis C, Kini U, Stark Z, Savarirayan R, Sheffield LJ, Zuffardi O, Slater HR, de Vries BB, Knight SJ, Anderlid BM, and Schoumans J.
Sixteen New Cases Contributing to the Characterization of Patients with Distal 22q11.2 Microduplications.
Molecular Syndromology 2011, 1:246-254.

RELATED PUBLICATIONS

- I. **Wincent J**, Schulze A, and Schoumans J.
Detection of CHD7 deletions by MLPA in CHARGE syndrome patients with a less typical phenotype.
European Journal of Medical Genetics 2009, 52:271-272.
- II. Anderlid BM, Blennow E, Giacobini M, Nordgren A, **Wincent J**, Schoumans J, and Nordenskjold M.
Gene dosage array can even discover small chromosome changes. More children with developmental deviations may be offered an etiological diagnosis.
Läkartidningen 2010, 107:1144-1149.



CONTENTS

INTRODUCTION	1
Neurodevelopmental disorders	1
Mental retardation, intellectual disability and developmental delay	1
Syndromic neurodevelopmental disorders	2
Non-genetic causes of mental retardation	3
Studying the human genome	4
From the double helix to array-CGH	4
Reverse phenotypics	6
The 22q11.2- region and low copy repeats	6
Genetic causes of mental retardation	8
Single gene disorders	8
Chromosome aberrations	8
AIM	10
MATERIAL AND METHODS	11
Patients	11
DNA sequencing	12
Multiplex Ligation-dependent Probe Amplification	13
Array-CGH	14
RESULTS AND DISCUSSION	16
Single gene alterations (papers I and II)	16
Cornelia de Lange syndrome	16
CHARGE syndrome	17
Conclusions	18
Copy number variations (papers III, IV, V)	19
Deletion of chromosome 11q13.4-q14.3	19
Array-CGH in a clinical setting	20
Distal 22q11.2 duplications	22
Conclusions	23
Genotype-phenotype correlations (all papers)	24
Cornelia de Lange syndrome	24
CHARGE syndrome	25
Deletion of chromosome 11q13.4-q14.3	26
Array-CGH in a clinical setting	28
Distal 22q11.2 duplications	30
CONCLUDING REMARKS AND FUTURE PERSPECTIVES	31
SAMMANFATTNING PÅ SVENSKA	33
ACKNOWLEDGEMENTS	36
REFERENCES	38

LIST OF ABBREVIATIONS

BAC	Bacterial artificial chromosomes
DbGAP	Database of genotypes and phenotypes
bp	Base pair
BCRL	Breakpoint cluster region-like
CHARGE	Coloboma, Heart defect, Atresia choanae, Retarded growth and/or development, Genital hypoplasia and Ear anomalies/deafness
CGH	Comparative genome hybridization
CNV	Copy number variation
CdLS	Cornelia de Lange syndrome
DECIPHER	DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources
DGV	Database of genomic variants
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DD	Developmental delay
DSM-IV	Diagnostic and statistical manual of mental disorders
ddNTP	Dideoxyribonucleotide triphosphate
DGS/VCFS	DiGeorge syndrome/Velocardiofacial syndrome
ECARUCA	European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations
FISH	Fluorescent in situ hybridization
IQ	Intelligence quotient
ICD-10	International Classification of Diseases
Kb	Kilobase
LCR	Low copy repeat
Mb	Megabase
MR	Mental retardation
MCA	Multiple congenital anomalies
MLPA	Multiplex Ligation-dependent Probe Amplification
NDD	Neurodevelopmental disorder
NAHR	Non-allelic homologous recombination
PCR	Polymerase chain reaction
RT-PCR	Reverse transcriptase PCR
RNA	Ribonucleic acid
SNP	Single nucleotide polymorphism

INTRODUCTION

NEURODEVELOPMENTAL DISORDERS

Neurodevelopmental disorders (NDD) comprise disorders of brain structure, chemistry or physiology that adversely impacts the normal-, behavioural, emotional, physical and learning processes that unfolds with maturity in living species¹. Included among the neurodevelopmental disorders are intellectual developmental disorders, communication disorders, learning disorders, motor disorders, autism spectrum disorders and attention deficit/hyperactivity disorders².

The majority of the patients described in this thesis had an intellectual developmental disorder. Many terms have been used to describe this condition, including mental retardation and intellectual disability, which in this thesis will be used synonymously. Some patients had their main problems in one of the other diagnose groups (mainly communication disorders) and many patients had symptoms from several subgroups. However, the main focus of this thesis has been on intellectual developmental disorders.

Mental retardation, intellectual disability and developmental delay

Mental retardation (MR) is a disability that affects cognitive as well as non-cognitive functions and is defined in many different ways. In the diagnostic and statistical manual of mental disorders (DSM-IV), the diagnostic criteria include significant sub-average intellectual functioning, significant limitations in at least two areas of adaptive behaviour (i.e. the ability to function at age level in an ordinary environment) and the onset before the age of 18 years³. In DSM-V, that will be available shortly, the term MR will be replaced by intellectual disability and is proposed to be defined as a deficit in general mental abilities, impaired function in comparison to a person's age and cultural group by limiting and restricting participation and performance in one or more aspects of daily life activities with an onset during the developmental period².

The international Classification of Diseases (ICD-10) defines MR as a condition of arrested or incomplete development of the mind, which is especially characterized by impairment of skills manifested during the development period; skills which contribute to the overall level of intelligence, ie, cognitive, language, motor and social abilities⁴.

With regard to the intellectual criterion for the diagnosis, MR is generally defined by an IQ-test score of approximately 70 and below. In DSM-IV and ICD-10, MR is sub-grouped into mild MR (IQ ~50-70), moderate MR (IQ ~35-50), severe MR (IQ~20-35) and profound MR (IQ ~ < 20). Severe and profound MR are often grouped together as is sometimes mild and moderate MR. The distribution goes from a high proportion of mild MR (85%) to a low proportion of profound MR (1-2%)³.

In young children (approximately under age 5), standardized IQ-testing is not reliable and the term developmental delay (DD) is often used instead. DD is defined as a significant delay in two or more of the following areas; gross or fine motor development, speech/language, cognition, social/personal development and activities of daily living, and is thought to predict the future manifestation of MR⁵.

The commonly used definitions described above try to capture the limitations in different aspects that individuals with MR/DD are affected of, emphasizing the vast impact on everyday functioning this NDD has on the patients and their families.

The prevalence of MR varies in different studies. In a review by Leonard and Wen, the prevalence of severe MR (IQ<50) was 3-4/1000 children, with a range of 1-7/1000 while the prevalence of mild MR (IQ 50-70) was approximately 33/1000 but showed even

more variation with a range of 2-35/1000⁶. Much of the wide range is likely due to differences in definition, classification and methods of investigation rather than true differences in prevalence, although socio-economic differences between populations may also exist. The incidence of MR has decreased in Sweden during the last decades. This is probably related to both medical and social progress and most strikingly has the number of cases due to pre- and perinatal infections and traumatic deliveries decreased. Different studies in our country give a prevalence of around 7/1000 children⁷.

Syndromic neurodevelopmental disorders

NDDs, and particularly MR, are frequently associated with a wide range of accompanying symptoms such as multiple congenital anomalies (MCA), dysmorphic features, pre- and postnatal growth retardation, epilepsy and sensory (vision and/or hearing) impairment^{8,9}. Congenital malformations, affecting for example the limbs, heart and brain, result from an intrinsically abnormal developmental process. Brain malformations such as agenesis of the corpus callosum, polymicrogyria or holoprosencephaly may be directly related to the NDD. Dysmorphic features are visible deviations of outward body form, for example epicanthus, low set ears and clinodactyly. The NDD can thus sometimes be one of the symptoms of a syndrome, ie, a particular set of clinical characteristics occurring together in a recognizable pattern that is known or assumed to have a mutual aetiology. Examples of such syndromes are Cornelia de Lange Syndrome (CdLS) and CHARGE syndrome.

CdLS is characterized by severe mental and growth retardation and distinctive dysmorphic facial features including low anterior hairline, long eyelashes, arched eyebrows, synophrys, anteverted nares, maxillary prognathism, long philtrum and thin lips. Other important clinical features are microcephaly, hirsutism, upper limb- and gastrointestinal malformations. CdLS has a variable phenotype, with the mild phenotype characterized by lesser mental and growth retardation and milder limb anomalies^{10,11}.

CHARGE syndrome is an autosomal dominant disorder with an incidence that might be as high as 1 in 8,500 births¹². The original diagnostic criteria required the presence of four out of the six CHARGE characteristics; **C**oloboma, **H**ear defect, **A**tresia choanae, **R**etarded growth/development, **G**enital hypoplasia and **E**ar anomalies/deafness. At least one of these characteristics had to be either coloboma or choanal atresia¹³.

Blake 1998	Major criteria	Minor criteria
<u>Classical CHARGE:</u> 4 major <i>or</i> 3 major + 3 minor	1. Coloboma 2. Choanal atresia 3. Characteristic external ear anomaly 4. Cranial nerve dysfunction	1. Cardiovascular malformations 2. Tracheo-oesophageal defects 3. Genital hypoplasia or delayed pubertal development 4. Cleft lip and/or palate 5. Developmental delay 6. Growth retardation 7. Characteristic face
Verloes 2005	Major criteria	Minor criteria
<u>Typical CHARGE:</u> 3 major <i>or</i> 2 major + 2 minor	1. Ocular coloboma 2. Choanal atresia 3. Hypoplastic semicircular canals	1. Heart or oesophagus malformation 2. Malformation of the middle or external ear 3. Rhombencephalic dysfunction including sensorineural deafness 4. Hypothalamo-hypophyseal dysfunction 5. Mental retardation
<u>Partial CHARGE:</u> 2 major + 1 minor		
<u>Atypical CHARGE:</u> 2 major <i>or</i> 1 major + 3 minor		

Table 1. Clinical diagnostic criteria for CHARGE syndrome according to Blake and Verloes^{14,15}.

In 1998 Blake et al. defined major and minor criteria of CHARGE syndrome and proposed that the major characteristics often occur in CHARGE syndrome but are less common in other conditions¹⁴. Verloes proposed diagnostic criteria for CHARGE syndrome in 2005 that reinforce the embryological defects and avoid secondary anomalies and sex-dependent criteria¹⁵ (table 1).

Non-genetic causes of mental retardation

The aetiology of MR and the diagnostic yield appears to be highly variable in different studies¹⁶⁻¹⁸. In a study by Stevenson et al. from 2003 including 10.997 individuals with MR drawn from a service delivery population, an aetiological diagnosis was made in 44%¹⁹. So, despite the extensive clinical and laboratory investigation that these individuals undergo, the cause for their symptoms can only be determined in less than half of the cases. For the children's families this is frustrating since they will have no information of the prognosis for the child or the recurrence risk in a new pregnancy. The lack of an aetiological diagnosis in a great number of cases also hampers the development of specific therapy and preventive measures.

Although having its limitations, the study by Stevenson from 2003 (ie before the introduction of some of the methods described in this thesis as well as before the identification of the causative genes for CdLS and CHARGE syndrome) gives a general idea of the main causative categories (Figure 1). Approximately 16% of MR cases were ascribed to environmental factors¹⁹. Examples of environmental factors are excess maternal alcohol consumption or drug abuse during pregnancy and maternal infections such as rubella, toxoplasmosis and cytomegalovirus. Furthermore complications of prematurity or delivery, postnatal emotional deprivation, malnutrition and infectious diseases, such as meningitis and encephalitis, may cause MR.

In around 8% a multifactorial cause was likely. Multifactorial disorders result from the action of one or multiple genes in combination with environmental factors. Examples include congenital deformities of the central nervous system leading to NDDs, such as neural tube defects, hydrocephaly and agenesis of the corpus callosum.

Approximately 20% was attributed to genetic causes. This figure is likely slightly underestimated since conditions with early lethality would have been missed. Nevertheless, there is an increasing body of evidences indicating that many of the patients with hitherto unexplained MR have a genetic cause²⁰. Trisomy 21 and fragile-X mental retardation, the two most frequent causes of MR¹⁹, have genetic aetiologies. Furthermore, the co-occurrence of NDDs with congenital malformations or dysmorphic features indicates a constitutional, possibly genetic background. In addition, cases with MR, without known diagnosis, often have several affected close relatives, suggesting a common genetic background.

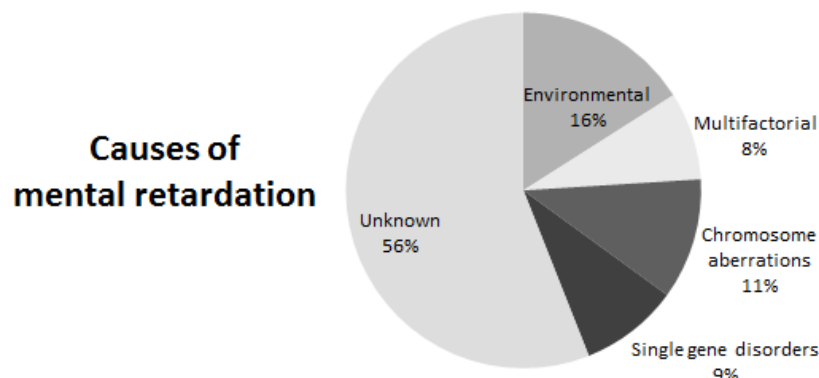


Figure 1. Causes of mental retardation. Adapted from Stevenson *et al.*, 2003¹⁹.

STUDYING THE HUMAN GENOME

From the double helix to array-CGH

The word genetics means the studies of inherited elements, and ever since methods of how to study the human genome started to develop, new genetic causes of NDDs have continuously been identified. Gregor Mendel was the first to describe inherited characteristics in the 1860's in his experiments with peas and a few decades later in 1882 Walther Flemming identified the chromosomes. The DNA double helix was identified by Watson and Crick in 1953, whereupon it was recognized that the genes, ie a coding nucleotide sequence carrying the basic elements of hereditary traits in living organisms, were located within the DNA molecule.

A chromosome consists of one single DNA molecule that is tightly packed by histones and other proteins (Figure 2). The DNA molecule itself has a linear backbone of sugar and phosphate residues and attached to each sugar residue is a nitrogen base (adenine, cytosine, guanine or thymine). A nucleotide is the sugar-phosphate residue with its nitrogen base and it constitutes the basic repeating unit of the DNA. The DNA double helix is bound together by hydrogen bonds between complementary bases (ie T-A and C-G). In the genes, a set of three nucleotides, ie a codon, encodes an amino acid, the basic repeating unit of the proteins. The central dogma of molecular biology with a unidirectional flow of information from DNA to RNA to protein was introduced in a paper by Crick in 1957 and the structure of eukaryotic genes with exons and introns and the process of transcription, splicing and translation has in the years that followed been described. The haploid human genome consists of approximately 3 billion DNA base pairs and approximately 21.000 genes coding for proteins and functional RNAs.

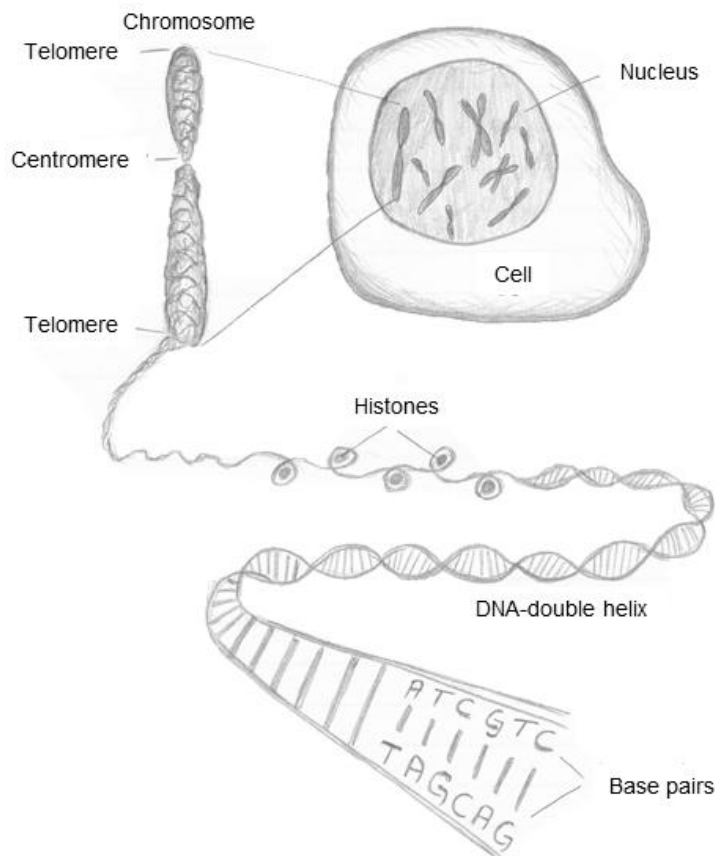


Figure 2. Illustration of DNA packed into a chromosome

In the fifties orcein staining of chromosomes was introduced and in 1956 Tjio and Levan determined the diploid human chromosome number to 46, which paved the way for identification of numerical chromosome aberrations in patients with various symptoms including NDDs. In 1959 Jérôme le Lejeune showed that Down's syndrome was caused by an extra chromosome 21 and Charles Ford detected the 45,X karyotype in Turner syndrome. This was followed by the identification of 47,XXY in Klinefelter syndrome, trisomy 18 in Edwards's syndrome and trisomy 13 in Patau syndrome. During the 1960s and 70s the quality of chromosome analysis improved with the introduction of banding techniques that gave each chromosome a characteristic banding pattern. With this came the possibility to detect losses or gains of parts of chromosomes as for example loss of material from the long arm of chromosome 18 in 18q-deletion syndrome or loss of the short arm of chromosome 5 in Cri-du-chat syndrome. Thenceforward, deletions, duplications, translocations and inversions have continuously been reported and categorized.

In the eighties molecular techniques were rapidly developed. PCR and Sanger sequencing enabled robust and easy DNA analysis down to the single base pair although the technique is labour-intensive and in the beginning was limited by the fact that the complete human genome sequence was not yet known. The development of these methods became crucial steps for the identification genes and mutations involved in monogenic diseases.

After the introduction of fluorescence in situ hybridization (FISH) in the nineties, submicroscopic deletions were detected in several syndromic forms of NDDs such as deletion of 22q11.2 in DiGeorge-/Velocardiofacial syndrome (DGS/VCFS) and deletion of 7q11.2 in Williams-Beuren syndrome. Although FISH enabled the detection of submicroscopic genomic imbalances, the technique is targeted and the clinician needs to have a prior idea of which chromosomal regions is of interest and should be investigated. Several quantitative PCR based techniques such as Quantitative Fluorescent-PCR and Multiplex Ligation-dependent Probe Amplification (MLPA) have also been developed for the identification of submicroscopic chromosome aberrations. However, these techniques are also targeted only allowing investigation of a limited number of loci in a single experiment. With approaches such as multiprobe FISH and spectral karyotyping simultaneous visualization of all chromosomes with fluorescent probes became possible. Nevertheless, these techniques are labour intensive and have a limited resolution²¹.

In the nineties chromosome based comparative genomic hybridization (CGH) was developed. The technique is based on hybridization of equal amounts of patient and reference DNA, which are labelled with different fluorophors, to normal human metaphase chromosomes²². Although, the detection of small, cryptic aberrations still was limited, this technique paved way for the development of array-CGH.

The microarray technology (Array-CGH/molecular karyotyping, in which the patient and reference DNA is hybridized to DNA-probes on a glass slide instead of metaphase spreads), enabled high resolution high-throughput genome-wide detection of submicroscopic deletions and duplications reducing the gap between cytogenetic techniques and molecular genetics^{23,24}.

Initially, the probes on the arrays were BAC-clones and the technology was mainly available to researchers with dedicated microarray facilities. However, gradually the BAC-arrays were replaced by commercially available oligonucleotide-arrays that could more easily be implemented in clinical diagnostic laboratories. The oligonucleotide-arrays generally provide higher resolution and better genome coverage.

Before the onset of the work described in this thesis, our group investigated 41 children with MR using a 1Mb BAC-array²⁵. Although it gave encouraging results with a diagnostic yield of 10%, it also showed that we could not rely on single BAC-clones for detection of genomic imbalances and needed denser arrays to avoid false positives. Therefore, the resolution was gradually improved by increasing the number of BACs on the array²⁶ and subsequently commercially available platforms were validated before implementing the technology in our diagnostic setting²⁷.

Reverse phenotypics

The implementation of array-CGH over the last decade has enabled the identification of submicroscopic genetic aberrations in patients with NDDs and related symptoms. Although most imbalances are non-recurrent and spread across the genome, several overlapping aberrations have also been identified. When investigating the clinical features of patients with overlapping aberrations it has sometimes been possible to determine common clinical features in retrospect, leading to the delineation of new clinical syndromes. This “genotype first” or “reverse phenotypics” approach^{28,29} by which patients are identified by a similar genomic aberration before a common clinical presentation is defined, has proven to be successful in many cases, as for examples the 17q21.31 deletion syndrome³⁰.

For other imbalances, for example 22q11.2 duplications, it has not been easy to define a common clinical presentation³¹. With increased use of array-CGH, the 22q11.2 region, that has long been recognized as a hotspot for genomic rearrangement and related disorders, such as 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome (DGS/VCFS)), has been further explored and new recurrent imbalances has been investigated. These include duplications reciprocal to the deletions commonly seen in DGS/VCFS region and deletions and duplications located distally to the DGS/VCFS - region.

22q11.2 duplications reciprocal to the DGS/VCFS region (ie proximal 22q11.2 duplications) have been reported in approximately 50 index cases³¹. Distal 22q11.21–q11.23 duplications are also rare and only 22 cases have previously been described^{32–35}. The paucity of reported proximal and distal 22q11.2 micro-duplications may, in part, be explained by the absence of a defined phenotype and the wide range of sometimes mild symptoms³⁶.

The phenotypes of the patients with both distal and proximal 22q11.2 duplications are diverse, with symptoms ranging from mild DD and mild dysmorphic facial features to severe MR and multiple congenital malformations with no clearly definable collection of phenotypic features shared among the patients. Many of the duplications are inherited from mildly affected or asymptomatic parents^{31,33,34}.

The 22q11.2- region and low copy repeats

For many of the imbalances identified with array-CGH there are no common breakpoints, but in some cases, as with the 22q11.2-region, the genomic architecture predisposes the genomic region to rearrangements. The 22q11.2 region is characterized by the presence of several segmental duplications or low copy repeats (LCR) that function as mediators of non-allelic homologous recombination (NAHR) and the breakpoints of the recurrent rearrangements in this region cluster around these LCRs. LCRs are defined a segment of DNA, >1Kb in size, that occurs in two or more copies per haploid genome with the

different copies sharing >90% sequence identity³⁷. Due to the high level of similarity between LCRs, they may mediate NAHR. NAHR is based on alignment and subsequent crossing over of non-allelic homologous LCRs during meiosis, resulting in duplication, deletion or inversion³⁸.

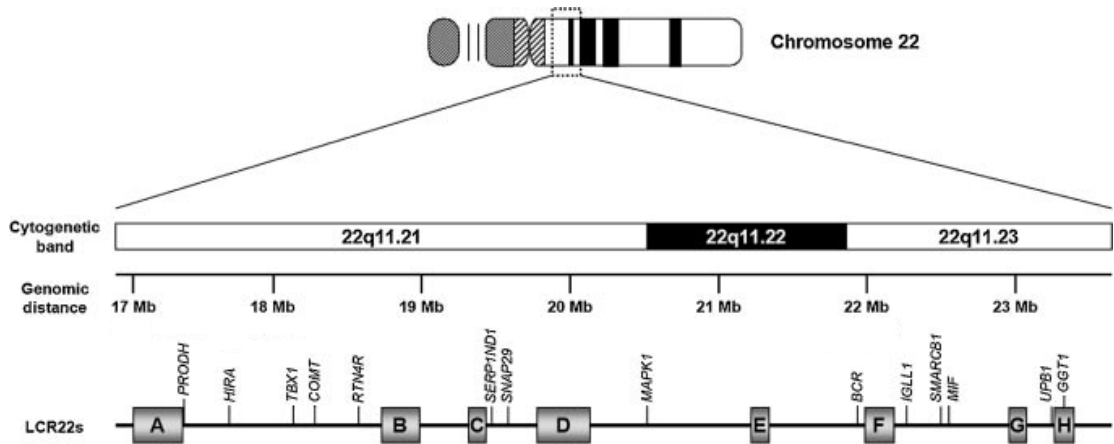


Figure 3. The eight LCR-cluster in the 22q11.2-region. Modified from Descartes *et al.*, 2008³².

Eight LCR clusters (LCR22A–H) have been identified in the 22q11.2-region (Figure 3)^{32,39}. The modules that build these LCR show significant (97–98%) sequence identity to each other, although the LCR22s differ between each other in content and organization of the modules⁴⁰. Most (>85%) individuals with proximal (involving LCR22A–D) 22q11 deletions (i.e. DGS/VCFS) have a 3 Mb deletion with breakpoints in LCR22s A and D, the largest and most complex of the LCR22s⁴¹. Deletions mediated by distal LCR22s (LCR22E–H) have also been described, although these deletions are found less frequently than the common proximal 22q11 deletions⁴². This may be due to differences in the rates of genomic rearrangement mediated by the various LCR clusters (due to underlying sequence identity/motif organization differences)³⁹ or the wider phenotypic spectrum associated with distal deletions.

GENETIC CAUSES OF MENTAL RETARDATION

As mentioned earlier, in the study by Stevenson et al, 20% of MR had genetic causes. Generally, genetic disorders can be divided into multifactorial disorders, single-gene disorders and chromosomal disorders. According to the study by Stevenson et al approximately 9% had a single gene cause, and in 11% a light-microscope visible chromosome aberration was detected. (It should however be noted that in the study by Stevenson micro-deletions associated with Prader–Willi syndrome, Angelman syndrome, Williams syndrome, and DGS/VCFS were included in the single gene category).

Single gene disorders

A gene can be disrupted in several ways. Point mutations exchange a single nucleotide for another and may result in silent mutations (no amino acid change), missense mutations (amino acid changes), nonsense mutations (introduction of a premature stop-codon) or splice-site mutations (disrupts a splice-site). Insertions and deletions add or remove one or a few nucleotides. If in frame, the result is an insertion or deletion of one or more amino acids. But the result may also be a frameshift mutation in which the reading frame is disrupted resulting in a completely different translation from the original and often a stop codon will eventually be introduced. The most common NDD-associated single gene disorder is fragile-X syndrome which is caused by mutations in *FMRI*¹⁹.

In 2004 two studies reported mutations in the *NIPBL* gene to cause CdLS^{43,44}. *NIPBL* is located on chromosome 5p13, consists of 47 exons and encodes delangin, a 2,804 amino acid protein that is important for sister chromatid cohesion. Heterozygous mutations in *NIPBL* have been found in approximately 60% of patients and in another 5% mutations are found in the cohesin structural components *SMC1A* and *SMC3*¹⁰.

Using array-CGH and sequencing of candidate genes, the gene encoding the Chromodomain helicase DNA binding protein 7 (*CHD7*) was in 2004 identified as a causative gene of CHARGE syndrome⁴⁵. *CHD7* is located at chromosome 8q12.1 and consists of 38 exons. It encodes a 2997 amino acid protein belonging to the chromatin organization modifier family. These proteins form part of a complex that is involved in modifying chromatin organization and gene expression and play an important role during embryonic development⁴⁵.

The DNA sequence is not static, and besides pathogenic mutations there are other small-scale changes including point mutations and deletions or insertions of one or few bases that may be benign. If common in the population, many of these benign changes are known and reported in different databases as single nucleotide polymorphisms (SNPs). However, rare or population specific variants are often not reported and sometimes complicates the interpretation of mutations found.

Chromosome aberrations

Chromosome aberration visible in a light microscope by cytogenetic analysis can further be divided into numerical aberrations and structural rearrangements. Numerical aberrations comprise changes in overall copynumber such as aneuploidy, (eg trisomy or monosomy), and ploidy changes, (eg triploidy). Structural rearrangements, (generally defined as genomic alterations larger than 1 kb in size³⁷), affect the structure of one or

several chromosomes, and may include translocations, insertions and inversions, but also changes in copynumber over specific regions (segmental aneuploidies) such as deletions and duplications. Chromosomal aberrations are a major cause of NDDs. Using routine karyotyping (resolution ~5-10Mb), an unbalanced karyotype can be found in 10-16% of cases^{18,46}. With an estimated frequency of 1/800 births trisomy 21 is the most common NDD-associated chromosome abnormality.

Not all chromosome abnormalities are visible in the light microscope and submicroscopic subtelomeric rearrangements have been identified in 2.5-6% of individuals with idiopathic MR^{21,47}. The 1p36 micro-deletion syndrome is the most frequently observed subtelomeric deletion and deletion of 22q11.2 is the most common interstitial submicroscopic aberrations readily identified by FISH¹⁸. Submicroscopic genomic variants that alter chromosome structure are also referred to as structural variation. Copy number variation (CNV) is a subgroup of structural variation defined as a segment of DNA that is 1kb or larger and is present at a variable copy-number in comparison with a reference genome³⁷.

At the onset of the work described in this thesis there had been reports of array-CGH identifying clinically relevant CNVs in approximately 10% of patients with idiopathic MR. The exact clinical interpretation of the CNVs observed, however, was, and still is, often challenging⁴⁸. One of the major difficulties is that CNV is much more common in control cohorts than what was previously thought. More than 12% of the reference genome likely involves CNV and it is considered that CNV contributes significantly to genetic variation between humans.^{38,49-52} Even monozygotic twins and different tissues from the same individual may differ in CNV status, showing that on-going somatic mutations may occur also during the lifetime of an individual^{53,54}. In addition, the *de novo* CNV rate in controls is estimated to be at least 1.2×10^{-2} CNVs per genome per transmission⁵⁵. This often makes it challenging to evaluate the clinical relevance of an imbalance when using whole genome array-CGH.

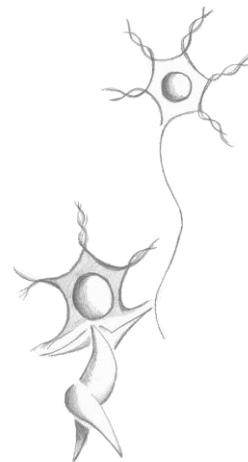
AIM

At the onset of the work described in this thesis, the etiological diagnosis for patients with MR was unknown in more than 50% of the patients. Karyotyping was the main tool for investigation in the clinical setting, sometimes followed by subtelomere screening with FISH. In addition, well-defined clinical syndromes were routinely investigated using FISH and/or PCR-based techniques. Causative genes of Cornelia de Lange syndrome and CHARGE syndrome had recently been identified but analysis was not yet clinically available. Furthermore, the first studies using array-CGH showed promising results and subsequently, molecular karyotyping found its way into the clinical workup of individuals with NDDs.

An aetiological diagnosis is of major importance for the patients and their families as it not only gives an explanation of the symptoms of the child but may provide more accurate prognostic information, adequate genetic counselling including recurrence risk estimations and enables prenatal diagnosis. When a genetic cause has been identified it is important to perform genotype-phenotype correlation studies in order to further understand the consequences of the genetic alteration.

The general aim of this thesis was to obtain a better understanding of the genetic basis of neurodevelopmental disorders, and mental retardation in particular, by aiming at the following objectives:

- Investigate the mutation frequencies in Swedish cohorts of patients with neurodevelopmental syndromes in which causative genes had recently been identified (Paper I,II).
- Characterize known aberrations with array-CGH (paper III, V).
- Evaluate the use of array-CGH in the clinical setting for patients with hitherto unexplained MR (Paper IV).
- Investigate the clinical features in the patients in order to enable genotype-phenotype correlations (all papers).



MATERIAL AND METHODS

PATIENTS

In study I we performed a *NIPBL* mutation screening by direct sequencing in a group consisting of eleven patients diagnosed with CdLS, including nine sporadic and one familial case (brother and sister). All patients had been referred to one of the clinical genetics departments in Sweden and were diagnosed by experienced Swedish paediatricians or clinical geneticists.

Thirty patients diagnosed with CHARGE syndrome were included in study II. The patients comprised 26 sporadic cases and two familial cases. One patient was diagnosed in Australia and the remaining patients were diagnosed by Swedish paediatricians or clinical geneticists. Twenty-three patients fulfilled Pagon's criteria and seven additional patients were included because it was strongly suspected that their less specific phenotypes were variants of CHARGE syndrome.

The patient described in study III was referred to the clinical genetics department at Karolinska University Hospital because of DD. In the clinical setting metaphase slides were prepared from lymphocyte cultures of peripheral blood and were examined with routine chromosome analysis. At the time of this study, array-CGH was not available in the clinical and therefore further investigation was performed as a research project.

Included in study IV were the first 160 patients with idiopathic DD/MCA that were referred for clinical array-CGH testing at the Department of Clinical Genetics at the Karolinska University Hospital (86 females and 74 males, age range 1 week - 46 years, average age 6.3 years, median age 4 years). Clinical data were reviewed for all patients, particularly inquiring degree of DD.

In study V we describe 16 patients with distal 22q11.2 duplications that were identified among 11,463 patients with idiopathic MR, brain malformations, autism spectrum disorders, and/or speech delay that were referred to different European and Australian clinical genetics centres for investigation with array-CGH analysis. Six patients were recruited from Nijmegen (the Netherlands), 6 patients from Melbourne (Australia), 2 patients from Oxford (England), 1 patient from Pavia (Italy) and 1 patient from Stockholm (Sweden). Two of the patients had previously been published elsewhere^{56,57}. Phenotypic data on patients and parents were collected from the referring physicians.

DNA SEQUENCING

DNA sequencing was used to search for single base substitutions, or deletions or insertions of one or a few bases, in *NIPBL* and *SMC1L1* in patients with Cornelia de Lange syndrome and in *CHD7* in patients with CHARGE syndrome.

Direct sequencing analysis is a method that accurately and specifically detects DNA base substitutions and small insertions or deletions. Dideoxynucleotides (ddNTPs) labelled with four different fluorescent colours, one for each nucleotide type (A, T, G, C), are mixed with deoxynucleotides (dNTPs). A doubled stranded PCR product is denatured and hybridized with a target primer and the sequencing enzyme polymerise the addition of nucleotides. Each time a ddNTP is incorporated, the chemical properties of the ddNTP (a hydrogen group on the 3' carbon instead of a hydroxyl group), disallow further incorporation of nucleotides. The end product of the reaction is composed of DNA strands of different lengths, all with a labelled ddNTP at the 3' end. These DNA strands are size separated by electrophoresis and the fluorescence is detected in an automatic DNA-sequencer. The differently labelled nucleotides are presented as peaks of different colours in generated chromatograms and can be compared to a reference sequence. Heterozygous mutations are seen as overlapping peaks of different colours (figure 4).

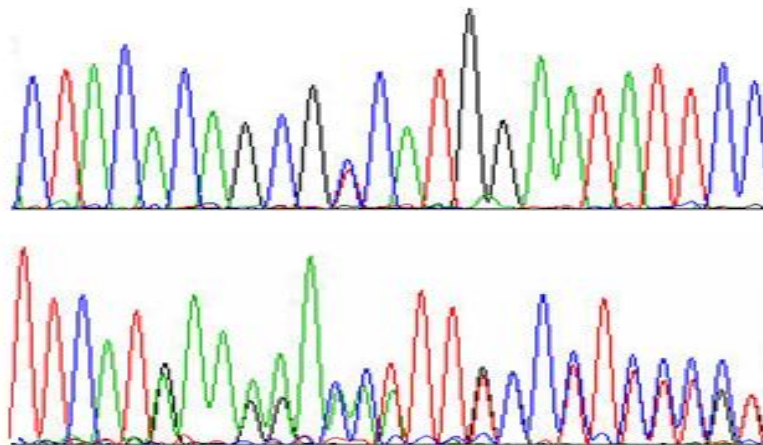


Figure 4. Chromatogram showing a nonsense mutation (top) and a frameshift mutation (bottom) in two patients from study II.

MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION

Multiplex Ligation-dependent Probe Amplification (MLPA) was used in study I and II in order to search for exon deletions or duplications in the *CHD7* and *NIPBL*-genes. In study IV MLPA was used for confirmation of array-CGH results for small duplications and also in a few other cases when no cell suspension was available. In both study IV and V MLPA was used for investigation of parental samples.

MLPA is a robust PCR-based method that detects copy number changes of genomic DNA simultaneously in several different loci. Two oligonucleotide “half-probes” are designed to bind adjacently to each other in each target sequence. The half-probes are hybridized to the test DNA and a ligase joins the two half probes into a complete probe. The probes are then amplified in a single reaction, using fluorescently labelled primers complementary to flanking sequences present in all probes. The probes are designed in such a way that the length of each amplification product has a unique size and can thus be separated and quantified by capillary electrophoresis in an automatic DNA-sequencer. Comparison of the relative peak area of each amplification product to a normal control reflects the relative copy number of the target sequence (figure 5).

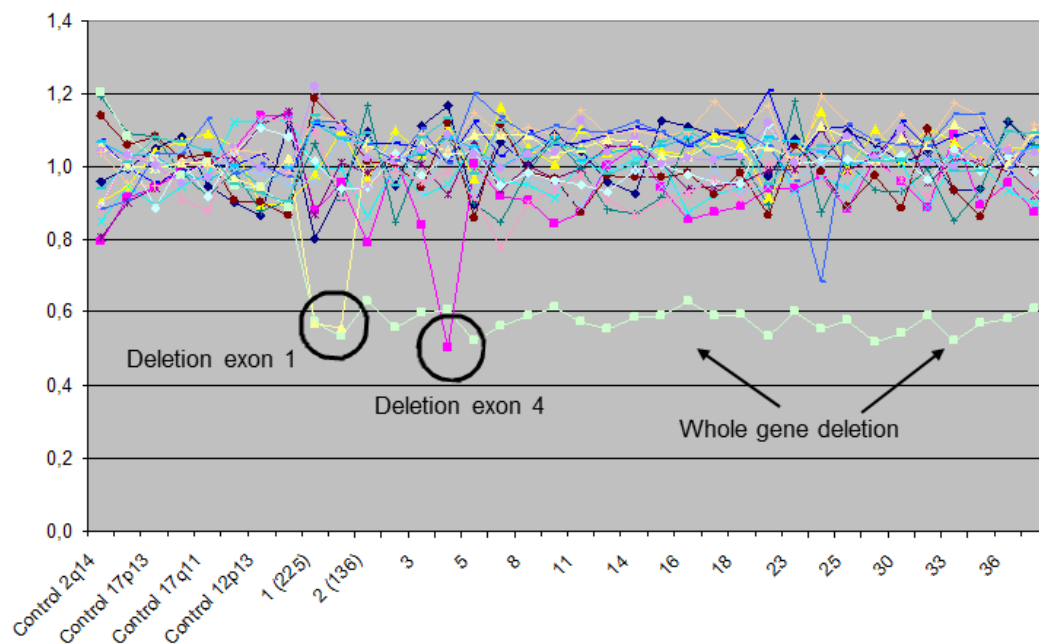


Figure 5. Result of MLPA analysis after calculation in Microsoft Excel for patients in study II. Deleted probes have a value of approximately 0.6.

ARRAY-CGH

Array-CGH offers genome-wide analysis of gain and loss of genomic material at high resolution. The method is based on hybridization of differently labelled test- and reference DNA, which are competitively hybridized to complementary DNA probes on a glass surface.

A DNA-array is composed of a glass slide on which genomic target sequences (probes) are attached, forming individual spots (Figure 6). For the arrays used in the work described in this thesis, these probes have mainly been bacterial artificial chromosomes (BAC-clones, size between 75-200 Kb), or synthetic oligonucleotides (with a size of 60 base pairs). Arrays can also be constructed using polymorphic oligonucleotide probes (CGH+SNP or SNP-arrays) that provide simultaneously genotyping information, which enables the identification of loss of heterozygosity without copy number changes, so called copy neutral loss of heterozygosity. Large stretches of copy neutral loss of heterozygosity indicate the presence of isodisomy due to uniparental disomy, which can also cause NDDs.

The number of probes on the slide varies between different designs and the resolution depends on the size and density of the probes. However, the resolution is also affected by the genomic spacing and the hybridization sensitivity of the probes as well as the quality of the experiment.

In principle, patient and control DNA are labelled with differently coloured fluorophors and are then mixed and hybridized together to the array. Hybridization of repetitive sequences is blocked by the addition of Cot-1 DNA. The arrays are scanned and the ratio of the test versus reference fluorescence signal intensity is determined (Figure 6). Because of the competitive nature of the binding, regions of the test-DNA with an increased copy number are identified by fluorescence as an increase in signal intensity of the test-DNA compared to the reference-DNA. Likewise, regions with genomic loss of the test-DNA are identified by an increase in signal intensity of the reference-DNA compared to the test-DNA (Figure 7).

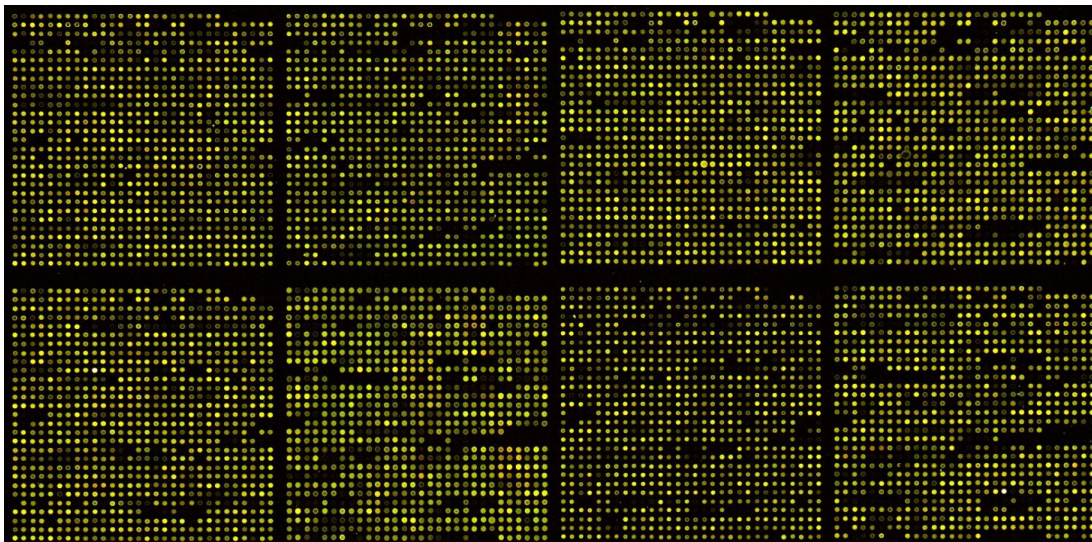


Figure 6. 1/6 of the scanned image of the 38K BAC-array from a patient in study IV.

In study IV DNA from the patients were investigated by either a 33/38K BAC-array or a 244K oligonucleotide array. 62 patients were investigated with a tiling path BAC-array with complete genome coverage containing 33,370 or 38,370 clones (33K for three and 38K for 59 patients) produced by the Swegene DNA Microarray Resource Center, Lund University. For array analysis Bio Array Software Environment (BASE)⁵⁸ was used. A

threshold of at least three consecutive aberrant clones was applied resulting in an effective average resolution of approximately 300kb. Ninety-eight samples were investigated with a 244K oligonucleotide-array with complete genome coverage produced by Agilent Technologies. Analysis was performed with Feature Extraction Software v. 9.1 and CGH-Analytics 3.4 (Agilent Technologies)⁵⁹. A threshold of at least six consecutive aberrant probes was applied resulting in an effective average resolution of approximately 50kb.

Paper V is a result of a collaboration between five clinical genetics centres why different array-platforms were initially used for investigation of the patients. The arrays used encompassed 38K BAC, 180K Agilent, 244K Agilent, Illumina-12-300K and Affymetrix 250K Nsp SNP. The 38K BAC array and the 250K SNP array have slightly lower resolution compared to the other platforms used. This is due to the large probes and the uneven distribution of the polymorphic probes across the genome respectively. When samples were available, the patients initially analysed with the 38K BAC array or the 250K SNP array were reanalysed with the 244K/180K Agilent array in order to refine and get more comparable breakpoints.

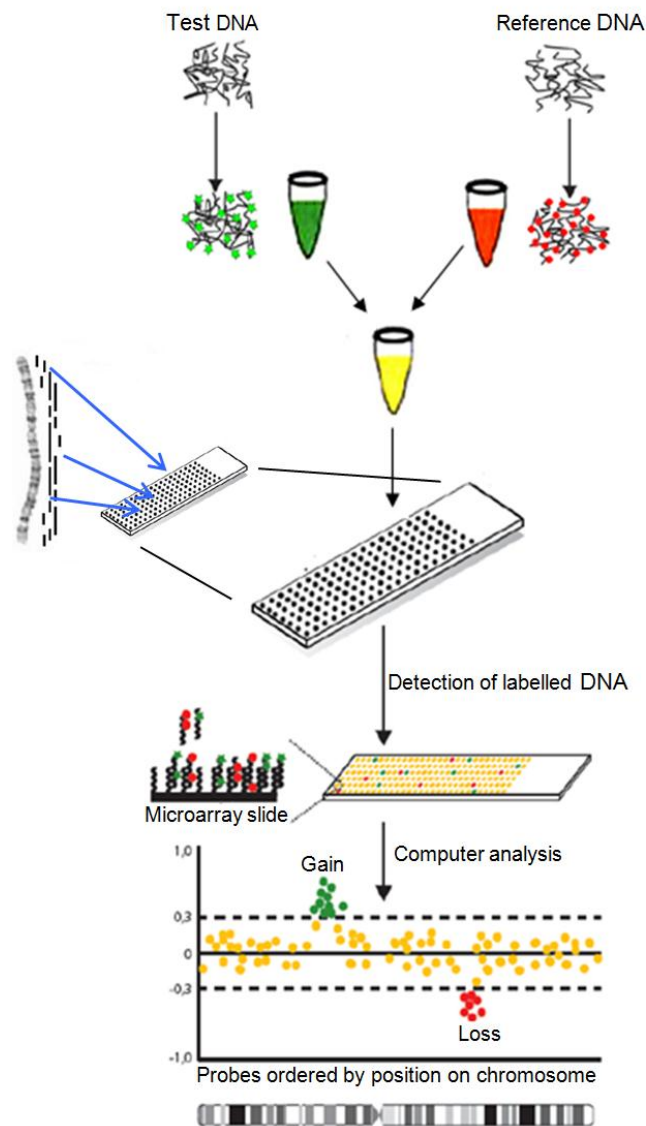


Figure 7. Schematic illustration of the principle of Array-CGH. Modified from Koolen, D.A, 2008⁶⁰.

RESULTS AND DISCUSSION

SINGLE GENE ALTERATIONS (PAPERS I AND II)

Cornelia de Lange syndrome

In paper I all 47 exons of the *NIPBL* gene were screened for mutations in eleven patients with Cornelia de Lange syndrome. The patient cohort comprised nine sporadic cases and one familial case consisting of a brother and a sister. Previous studies had identified *NIPBL* mutations in 26-56% of CdLS cases^{44,61-64}. We identified seven heterozygous mutations in our cohort including 3 nonsense mutations, 2 missense mutations, 1 splice mutation and 1 small deletion. All mutations were novel except for a nonsense mutation in exon 10 (p.R832X), which was previously reported in one case⁶². The two missense mutations (p.T2146P and p.A2436T) altered residues that were highly conserved across species and were not detected in 150 control subjects.

For five patients, samples from both parents were available and in all these cases the mutations occurred *de novo*. However, in two cases parental samples were not available and the inheritance is unknown. In case one, sample from the father was not available, but the identified mutation was a nonsense mutation and had previously been reported which strengthens the pathogenicity of this mutation⁶². In case 5, the in-frame deletion of 6-bp was predicted to result in a deletion of 2 amino acids that were highly conserved across species indicating that the mutation is pathogenic.

RT-PCR was performed in case 2 (splice site mutation affecting exon 19) and in case 5 (in-frame 6-bp deletion at the 5' end of exon 36) in order to investigate disruption of splice sites. In case 2 the analysis revealed an aberrant band sized 254 bp and a normal band sized 335 bp, demonstrating that the splice site mutation results in skipping of exon 19. In case 5 the splicing was unaffected since only one normal band with a size of 403 bp was detected (figure 8).

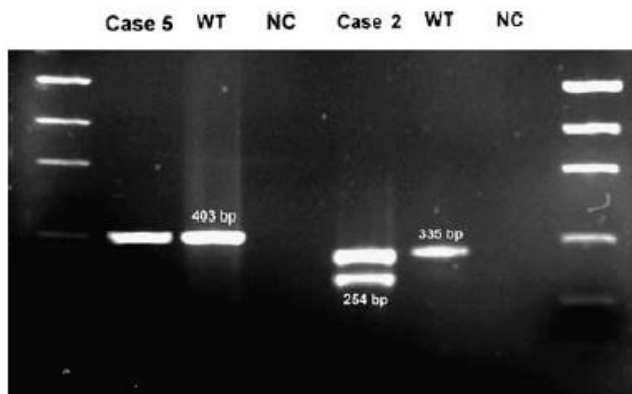


Figure 8. Picture of agarose gel electrophoresis of RT-PCR product from cases 2 and 5 in study I. An aberrant band is shown in case 2 indicating that exon 19 has been skipped during splicing. Case 5 shows normal splicing of exon 36. WT=wild-type, NC=negative control. Schoumans *et al*, 2007⁶⁵.

In four patients (case 8, 9 10a and 10b) no *NIPBL* mutations were detected by direct sequencing. These patients were analysed by MLPA for detection of *NIPBL* whole exon deletions or duplications and *SMC1L1* mutation screening was performed in the two boys, but no aberrations were found. These four patients were also investigated by tiling resolution array-CGH (33K BAC) for detection of cryptic chromosome imbalances. In case 8, a 0.6 Mb *de novo* duplication of chromosome 9p24.3 was identified. At the time of this study the clinical significance of this duplication was unknown. However, there have now been many reports of duplications of this region in the database of genomic variants (DGV) and the duplication is likely a normal variant.

CHARGE syndrome

In paper II, a series of 28 index patients (26 sporadic cases, one familial case consisting of a brother-sister case and one case consisting of monozygotic twins) were examined by direct sequencing of the 37 coding exons of the *CHD7* gene. Patients negative for *CHD7* point mutations or with missense mutations were further investigated by MLPA. In previous studies, *CHD7* mutations were identified in 58-71% of individuals with CHARGE syndrome^{45,66-68}. In our study we identified mutations in 18 of 28 cases (64%) that are most likely causal for the CHARGE phenotype. The mutations were *de novo* in all cases for which parental samples were available (15/18). The mutations comprised 15 point mutations (six nonsense (33%), six frameshift (33%) and three missense mutations (17%)), two exon deletions and one whole gene deletion (17%). The mutations were scattered throughout the gene (figure 9).

The twelve nonsense and frameshift mutations were truncating and therefore very likely to be causal for the phenotype. Two of the missense mutations were located in functional domains of *CHD7* and could affect the respective functions of the domains. The third missense mutation, p.V1742D, was not located in a functional domain. The mutation could on the other hand affect splicing, however, *in silico* testing did not support this and RNA was not available for *in vivo* testing. Nonetheless, mutations outside the functional domains have previously been reported as pathogenic^{66,68}. These three missense mutations were *de novo*, affecting amino acids that are conserved across species and were not detected in 90 control subjects. It seems likely that these mutations are pathogenic.

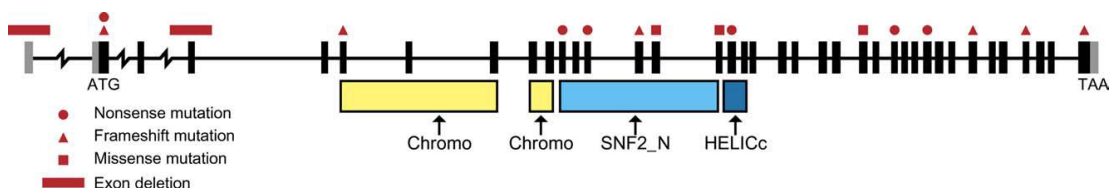


Figure 9. Summary of *CHD7* mutations detected in study II. Wincent *et al.*, 2008⁶⁹.

Inherited *CHD7* variants

Inherited missense variants were detected in four patients (1, 4, 9 and 19). Case 1 had inherited a missense variant (p.G117D) in exon 2 from his apparently healthy father. The affected amino acid was semi-conserved across different species and was not found in 180 control subjects. In this patient a *de novo* deletion of the 5' untranslated region (5' UTR) was also detected. Thus, it seems more likely that the *de novo* deletion of the 5'UTR caused CHARGE syndrome in this patient, and that the paternally inherited change likely is a rare variant without clinical significance. At the time of this study further investigation of expression of *CHD7* in this patient could not be performed due to lack of RNA. However, we were later able to collect RNA from this patient and expression analysis is ongoing. In addition we have investigated 150 control subjects with MLPA for presence of the deletion and none have been found. The breakpoints of the deletion have also been fine-mapped by custom array-CGH analysis with a ultra high dense coverage of *CHD7* and the flanking regions (unpublished data).

An inherited missense variant (p.S103T) was identified in case 9 and a small inherited duplication (p.K684_A685dup) was found in case 4. The mother of case 4, who carried the same duplication, was born with cleft lip and palate. Although the amino acids affected in both cases were conserved among different species, these

changes are unlikely to be pathogenic since they were found in control subjects and additional *de novo* mutations were found in the two patients (a previously reported causal missense mutation⁶⁶ in case 9 and a deletion of exon 4 in case 4).

In case 19, only a maternally inherited missense variant (p.R1592W) was found. The mother, who is very well functioning with normal hearing, has short stature and congenital hip dislocation but no signs of CHARGE syndrome. The affected amino acid was conserved among different species and the change was not found in 180 control subjects. Mildly affected carriers transmitting mutations to their children have been reported. In one family both affected children had severe expression of CHARGE syndrome but the father, who also carried the mutation, only had asymmetric anomaly of the pinnae^{66,70}. The mother of case 19 could thus have a very mild phenotype, not recognizable as CHARGE syndrome, or she could be mosaic for the variant. However, the clinical significance of this variant is uncertain.

The clinical importance of inherited variants may be difficult to interpret and it cannot be excluded that they contribute to the phenotypes of the patients.

Conclusions

Both of these studies confirm that *NIPBL* and *CHD7* are the main causative genes for CdLS and CHARGE syndrome respectively. This research project has contributed to the implementation of *NIPBL* and *CHD7* mutation analysis in the diagnostic setting at the Clinical genetics department at the Karolinska University Hospital.

However, in more than 30% of our CdLS- and CHARGE syndrome patients no causal mutation could be detected. This might be due to alterations not detectable by the approaches used so far, such as intragenic rearrangements or mutations in the intronic- or promoter regions of the genes. However, the identification of *SMC1A* and *SMC3* mutations in patients diagnosed with CdLS implies that locus heterogeneity is present for CdLS and this could also be the case for CHARGE syndrome. Furthermore, at the International Congress of Human Genetics (ICHG 2011) additional candidate genes for CdLS were presented.

COPY NUMBER VARIATIONS (PAPERS III, IV, V)

Deletion of chromosome 11q13.4-q14.3

The patient described in paper III had been investigated with standard chromosome analysis in the clinical setting and an interstitial deletion of chromosome 11q was detected. However, it was not possible to determine whether the deletion comprised band 11q14 or 11q22 due to the symmetrical band pattern. We performed a 38K BAC array-CGH analysis that showed an 18.2 Mb deletion at 11q13.4-q14.3 comprising approximately 100 genes (Figure 10). Both parents showed normal karyotypes, thus the deletion was *de novo*. At least 30 of the deleted genes are expressed in the brain. Six of the genes are reported to be disease-causing if disrupted. Four of these cause autosomal recessive disorders with clinical signs not observed in our patient. Defects in two genes, *KCNE3* and *FZD4*, are associated with autosomal dominant disorders.

KCNE3 encodes a potassium voltage-gated channel and a missense mutation in this gene has been associated with hypokalemic periodic paralysis, although other studies have subsequently shown that this variant likely is a rare polymorphism. Missense mutations in *KCNE3* have also been found in a family with Brugada syndrome (a condition characterized by an increased risk of cardiac arrhythmia). However, functional studies indicate that the missense mutation in Brugada syndrome causes a gain-of-function of *KCNE3*, which the deletion in our patient will not do.

FZD4 is a member of the frizzled gene family that encodes receptors for the Wnt type MMTV integration site family of signalling proteins. Mutations in *FZD4* leading to loss of activity⁷¹ cause autosomal dominant exudative vitreoretinopathy 1 (EVR). EVR is characterized by avascularity of the peripheral retina and exhibits a variable phenotype, with the most serious form resulting in blindness. It is likely that our patient had some clinical features of EVR, since the penetrance is regarded to be 100%. However, the clinical expression is variable and he probably has a very mild form since he had a normal ophthalmological examination.

A possible candidate gene for the patient's DD is *ARRB1* that is expressed in the central nervous system and is a member of the arrestin/beta-arrestin protein family, which is thought to cause specific dampening of cellular responses to stimuli such as hormones, neurotransmitters, or sensory signals. However, pointing out specific candidate genes is difficult because of the many genes in the deleted region. The phenotype seen in our patient is likely a result of the haploinsufficiency of a number of genes in the region.

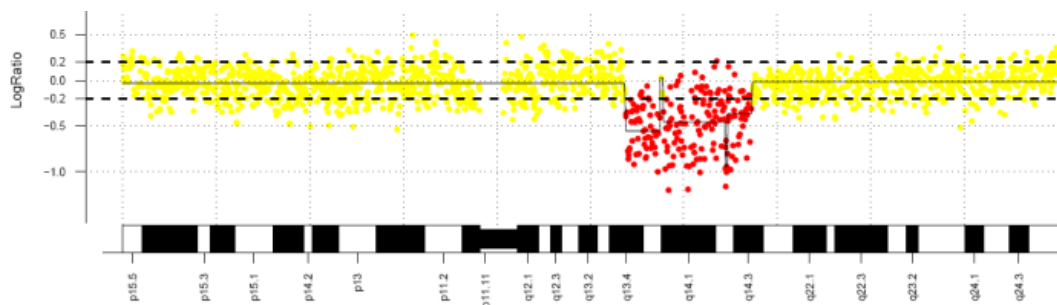


Figure 10. Deletion of chromosome 11q13.4-q14.3 detected by a 38K BAC-array in the patient in study III.

Array-CGH in a clinical setting

At the onset of the work described in this thesis, investigation of patients with MR, MCA and/or dysmorphic features with array-CGH had already revealed that CNVs are an important cause of otherwise unexplained causes of MR. Furthermore, array-CGH had started to be used for clinical investigation of patients with NDDs, mainly MR and autism, and often in combination with congenital malformations, pre- and postnatal growth retardation and/or dysmorphism. At the time of the planning of study IV, the array platforms we selected did not yet contain SNP-probes (as is the case in current CGH+SNP arrays) but were chosen because they were found to be the most suitable for our purpose at that time, due to its higher sensitivity for small CNVs in mosaic and the option for flexible design²⁷. However, as a consequence uniparental disomy has not been investigated.

The aim of study IV was to evaluate the usefulness of high-resolution arrays as a diagnostic tool in our clinical laboratory, to investigate the diagnostic yield in patients clinically referred for investigation of DD/MCA, and to inquire the level of severity of DD in the patients tested and compare the diagnostic yield in the different subgroups.

Our study was conducted on patients referred to our clinical medical genetics service between 2007 and 2008 and at the start of the study the wide range of affordable high-resolution and high-density array platforms with flexible selection of probe coverage, was not yet commercially available. Thus, at first a “home brewed” 33K or 38K BAC-array was introduced into the clinical setting, but during the course of the study, the 244K Agilent oligonucleotide-array became available for our diagnostic service why this array was gradually introduced for routine array analysis. The first patients that were analysed were mainly “unsolved cases”, ie patients likely to have a chromosomal abnormality due to their clinical presentation and who had been thoroughly investigated with available methods. As experience and confidence increased in the detection and interpretation of CNVs, array testing was increasingly used, and as it turned out to be more cost efficient to perform array as a first tier genetic analysis, some of the patients were not investigated with chromosome analysis before array-CGH testing. In total, 80% of patients had previously been investigated by conventional karyotyping and 62% had undergone at least one type of additional testing. Most common were molecular testing for Fragile-X syndrome (28%), subtelomere-FISH/MLPA (13%) and exclusion of 22q11.2-deletion (14%).

Imbalances not overlapping with previously reported CNVs in DGV and which included at least one gene were confirmed by MLPA or FISH and parental samples were simultaneously examined to investigate inheritance. The pathogenicity of the CNVs were assessed using the guidelines described by Lee *et al.*⁴⁹. Briefly, an imbalance was considered likely causal if it arose *de novo*, contained genes, overlapped with a known genomic syndrome or was previously reported to cause a specific phenotype in the DECIPHER or ECARUCA databases and was not a CNV reported in DGV. The criteria were not exclusively applied and the gene-content of the CNVs and their function was also taken into account.

Diagnostic yield

Of the 160 investigated patients, CNVs not previously reported in DGV and including at least one gene were detected in 36 (22.5%) cases. Twenty-one (13,1%) aberrations were considered causal to the phenotype the patient was referred for, corresponding well to previous studies in which causal copy number alterations have been identified in circa 10% of patients with idiopathic DD⁷². Of the 21 causal findings, 13 overlapped a well-characterized syndrome (8.1% of all cases studied, 61.9% of cases with causal

array-CGH findings). Most common was 1p36-microdeletion syndrome, Wolf-Hirschhorn syndrome and Potocki-Lupski syndrome (identified in two patients each).

Eight causal CNVs were found with the BAC-array (12.9%) and 13 causal CNVs were found with the oligonucleotide-array (13.3%). None of the causal alterations were <1 Mb in size and ten were >5 Mb, of the latter, all but three had previously been investigated with standard karyotyping but escaped detection. Thus, standard karyotyping had in seven cases failed to detect CNVs ≥ 5 Mb, five of which were ≥ 10 Mb, emphasizing that more accurate and reliable methods were needed to exclude chromosome imbalances in DD/MCA patients. Therefore we subsequently started to use array-CGH instead of conventional karyotyping as the first step genetic investigation for patients with idiopathic DD/MCA.

BAC- and oligonucleotide-array comparison

The BAC-array detected three (1.7%) possibly causal aberrations that could not be confirmed with FISH/MLPA and therefore were considered as false positives. No false-positives were detected with the oligonucleotide-array, which was a great advantage compared to the BAC-array. This can be explained by the higher hybridization specificity of the oligonucleotide array compared to the large insert clones from the BAC-array (average size 100 kb on BAC-array compared to 60bp on oligo-array). Moreover, the densely covered oligonucleotide-array allowed a higher probe cut-off (6 versus 3 probes). From a counselling perspective, however, follow-up FISH in the patient and the parents is nevertheless valuable for the estimation of the recurrence risk, as it reliably detects possible balanced rearrangements (translocations or inversions) in healthy carriers.

The diagnostic yield of causal aberrations was equal for the BAC- and oligonucleotide-array. However, for the 15 (9.4%) CNVs of unclear clinical significance, the BAC-array detected three (4.8%) and the oligonucleotide-array detected twelve (12.2%). Two of these 15 CNVs were >1Mb. The BAC-array detected two CNVs <1 Mb while the Agilent-array detected eleven. Although the effective resolution of the 244K Agilent-array is higher than the 38K BAC-array, the diagnostic yield of both platform was approximately equal and no causal aberrations <300 kb were detected in this study. Increasing the resolution of a whole genome screen in the diagnostic setting further will likely identify a few clearly causal CNVs <300 kb but it will have its drawback of detecting a much higher number of CNVs of unclear clinical significance. Custom designed arrays containing only known syndrome regions and/or coding regions of known causative genes, such as exon targeting, have the advantage of high sensitivity for detecting causal CNVs, together with a low detection rate of CNVs of uncertain clinical significance.

Slightly reducing the density of the array or reducing the practical resolution by increased detection thresholds in the software also limits the pick-up rate of CNVs of uncertain clinical significance, while still enabling the detection of novel microdeletion/microduplication syndromes with high accuracy and allowing reverse phenotypics to increase our understanding of the clinical implication of CNVs of uncertain clinical significance.

Nevertheless, array-CGH has revolutionized the clinical investigation of patients with DD/MCA and enables us to provide more patients with an aetiology-based diagnosis. Since 2009 we have applied array-CGH as the first step analysis in the genetic evaluation of patients with NDDs in our clinical genetic department. This is now common practice and quality guideline for diagnostic laboratories have been established in Europe and USA^{73,74}.

Distal 22q11.2 duplications

In study V we report on the clinical and molecular characterization of 16 individuals with distal duplications of chromosome 22q11.2 (involving LCR22D–H and located distal to the region typically deleted in DGS/VCFS). We identified these 16 duplications among 11,463 patients with a variety of NDDs resulting in an estimated frequency of approximately 0.1%, which is slightly higher compared to the study by Coppinger et al., who identified 18 distal duplications among 22,096 patients tested³³.

Detailed molecular analysis of distal 22q11.2 deletion breakpoints has previously shown that they mapped to a BCRL (breakpoint cluster region-like) module in the LCRs, suggesting that this module may represent a rearrangement hotspot (Shaikh, et al., 2007). In addition, it has been proposed that modules within the LCRs that have a direct orientation with respect to one another are likely to mediate rearrangements⁷⁵ and thus the orientation of these BCRL modules may predict between which LCR22s it is likely that NAHR will occur. Many of the duplications in our cohort were flanked by LCRs E–H, F–H or E–F (Figure 11). The BCRLs in LCR E, F, and H are in the same orientation³⁹ and thus our findings of E–H-, E–F and F–H-mediated duplications support the hypothesis that BCRL motifs in the same orientation facilitate NAHR³⁹. More surprising was that 3 of our distal breakpoints were not flanked by known LCR22s, but all resided in a region between LCR22D and E. This could indicate an additional locus that predisposes to NAHR.

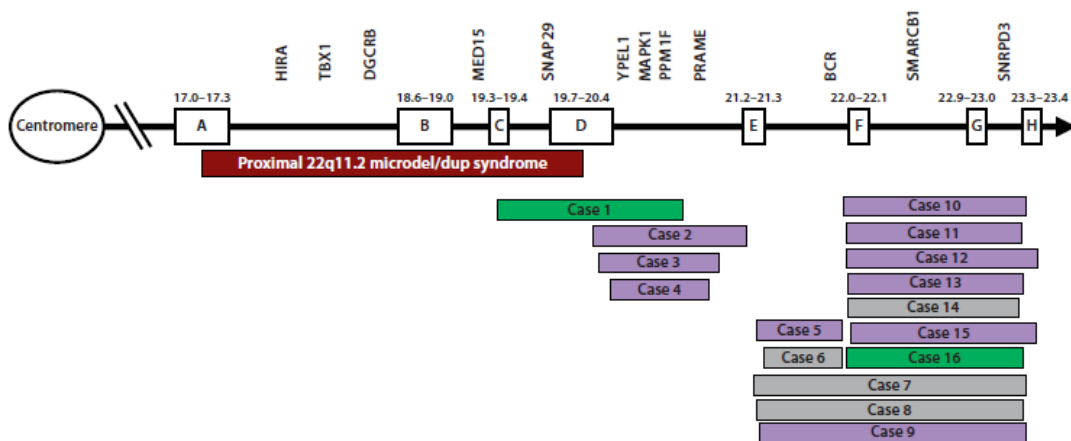


Figure 11. Schematic representation of the duplications identified in study V. Green bar = *de novo* duplication, purple bar = inherited duplication, grey bar = duplication of unknown inheritance. Modified from Vincent *et al.*, 2011⁷⁶.

Pathogenicity of distal 22q11.2 duplications

In concordance with previous studies of both proximal and distal 22q11.2 duplications, we found a high degree of inherited duplications in our study (83% of cases with available parental samples). Parents of a child with an inherited CNV may sometimes show mild variations of the child's phenotype, which for example have been reported for the DGS/VCFS that predominantly has a *de novo* occurrence⁷⁷. Unfortunately, we only had phenotypic data available on 7 of the parents from whom duplications were inherited. Six of these parents were apparently healthy and one parent was affected.

The varied phenotypic expression and incomplete penetrance observed for distal 22q11.2 duplications makes it exceedingly difficult to ascribe pathogenicity for these duplications. Although the fact that all probands reported so far display a clinical phenotype might be due to ascertainment bias, distal 22q11.2 duplications are rarely reported as normal variants. In a recent study by Cooper et al.⁷⁸, no distal 22q11.2 duplications were identified among approximately 8,000 healthy controls. In addition,

the duplications reported in control samples in the DGV overlapping the duplications identified in our patients are smaller and do not cover all the genes.

Given that distal 22q11.2 duplications, along with a growing number of recurrent genomic deletions and duplications⁷⁹, appear to be enriched in individuals with neurodevelopmental and neurobehavioral phenotypes compared to control samples, it is likely that distal 22q11.2 duplications represent a susceptibility/risk locus for NDDs rather than being causal variants. These copy number changes are insufficient to cause the observed phenotypic abnormality, and additional genetic, epigenetic or environmental factors may be required.

It is noteworthy that a digenic/multigenic model has recently been demonstrated for 16p12.1 deletions⁸⁰, which are inherited in the majority of cases and show considerable variability in expression. These deletions have been shown to co-occur with secondary pathogenic or ‘uncertain significance’ copy number change in approximately 24% of cases. The second hit could potentially be another copy number variant, a disruptive single-base-pair mutation in a functionally related gene, or an environmental event that influences the phenotype. There are also reports showing that patients with NDDs have an increased CNV-burden compared to controls⁸¹.

We identified additional copy number changes of unclear clinical significance in 5 (31%) of our cases making a 2-hit event plausible. A paternally inherited 170-kb duplication of chromosome 6p22.3 was found in case 2. The duplication comprised 2 exons of *JARID2*, encoding an ortholog of the mouse jumonji gene, which encodes a nuclear protein essential for mouse embryogenesis. Case 3 showed a paternally inherited 1.57-Mb deletion in 16q24.1–q24.2 involving approximately 20 genes. Case 8 had a 77-kb deletion of chromosome 4q12 comprising three genes. In case 14, a 250-kb deletion of 16p13.2 of unknown origin involving *GRIN2A*, that encodes an NMDA receptor subunit, was identified. Submicroscopic deletions, point mutations and translocation encompassing *GRIN2A* have recently been associated with neurodevelopmental phenotypes^{82,83}. Case 15 had a paternally inherited 1.94-Mb duplication of chromosome 4q35.2 comprising 7 genes.

In case 15, the additional copy number change was inherited from the parent carrying the 22q11 duplication. However, in case 2 and 3, the additional change was inherited from the non-22q11.2 duplication carrier parent making these additional copy number changes good candidates for the ‘second hit’.

Conclusions

Studies III and IV both show that more accurate and reliable methods besides karyotyping were needed to detect chromosome imbalances in DD/MCA patients. Through the development and implementation of array-CGH, the technical gap between molecular genetics and cytogenetic testing has been bridged. Besides the higher resolution molecular karyotyping provides other advantages over conventional karyotyping, including that the method is not dependent of the ability of cells to grow in order to generate good metaphase spreads and that it is amenable to automatization. A drawback is that balanced rearrangements and ploidy variation escape detection.

Nevertheless, the studies confirm that array-CGH is a highly effective technique in the diagnostics of individuals with MR/MCA. Also, as shown in study IV and V, the identification of submicroscopic structural variation of uncertain clinical significance puts new demands on genetic counselling and requires more research. Therefore, submission of cases to databases such as DECIPHER, ECARUCA and DbGAP are still important in order to make information about unpublished patients available to the diagnostic community as an aid in the interpretation of CNVs of uncertain clinical significance.

GENOTYPE-PHENOTYPE CORRELATIONS (ALL PAPERS)

Cornelia de Lange syndrome

Six of the seven patients that had a mutation in *NIPBL* (cases 1-6) demonstrated the classical CdLS phenotype including characteristic facial features (figure 12), severe growth- and mental retardation and four of had limb reduction. One patient with a missense mutation in exon 43 (case 7) had a milder phenotype with mild MR, growth retardation, distinctive facial features and no limb deficiencies. Two of the patients that had no detectable *NIPBL* mutation (case 8 and 9) showed some features overlapping with the CdLS phenotype (such as limb reduction and characteristic facial features) but demonstrated a clearly milder growth and mental delay compared to classical CdLS patients. The siblings (case 10a and 10b) however demonstrated severe CdLS phenotypic features except for limb reduction, but no mutations were detected.



Figure 12. Photograph patients in study I showing characteristic CdLS facial features. (a) case 1, (b) case 2, (c) case 3, (d) case 6, (e) case 10a and (f) case 10b. Schoumans et al. 2007⁶⁵.

A previous report on genotype-phenotype correlation suggested that missense mutations showed a trend towards a milder phenotype compared to other types of mutations and genotype-phenotype correlation in mutation positive and mutation negative individuals was observed⁶². However, a clear correlation between genotype and phenotype could not be confirmed in the study performed by Bhuiyan *et al*⁶⁴, nor by our study. We found a missense mutation in a patient with a severe phenotype (case 6) but also in a patient with a mild phenotype (case 7) and two patients demonstrated a severe CdLS phenotype while no *NIPBL* mutation was detected (case 10a and 10b). However, the limited sample size of our study makes it difficult to perform genotype-phenotype correlation. Nevertheless, *NIPBL* mutations are detected in the majority of individuals demonstrating the classical CdLS phenotype.

CHARGE syndrome

Of the 20 patients (18 cases) that had a *CHD7* mutation, there were five patients who had all four of Blake's as well as all three of Verloes' major criteria. This was not the case for any of the patients without a mutation. However, one patient (case 23) in the mutation negative group had two of Verloes' major criteria as well as three of Blake's major criteria, and several minor criteria, thus having typical/classical CHARGE syndrome. The other patients in the mutation negative group at most had two of Blake's major criteria and one of Verloes' major criteria (although information on temporal bone malformation was missing in several cases). Thus most cases without a detectable *CHD7* mutation did not have a classical/typical CHARGE syndrome phenotype.

Jongmans et al. reported that vestibular abnormalities were present in all investigated patients in their cohort of *CHD7* mutation positive patients⁶⁸. In our study, investigation of temporal bone malformations had been performed in 9 of the patients with *CHD7* mutation, and a temporal bone malformation was present in 8 of these cases (89%). Temporal bone malformation was thus also in our study an important clinical feature for the diagnosis of CHARGE syndrome. Nevertheless, the only patient in our study known to be negative for temporal bone malformation (case 4) had a likely causal *de novo* deletion of exon 4. Temporal bone malformation does thus not seem to be obligatory for *CHD7* mutation positive CHARGE syndrome. However, the boy died within the first year of life and abnormalities of the temporal bone might have been identified in further scanning at an older age. Temporal bone malformation remains the most frequent and specific sign of CHARGE syndrome.

Inherited CHD7 variants

Although the majority of CHARGE cases are sporadic, familial cases including inherited mutations have been reported^{66,68,70}. The brother and sister with the same nonsense mutation differed somewhat in their clinical presentation. The sister had choanal atresia and heart malformation not present in the brother, whereas the brother had facial nerve palsy and temporal bone malformation. The father tested negative for the mutation but unfortunately no sample was available for testing from the mother. However, it seems likely that the siblings inherited their mutation from one of their parents who might carry a gonadal mosaic mutation, or have a mild phenotype.

The phenotypes of monozygotic twins with *CHD7* anomalies previously reported have been similar but not identical^{66,68,84}. This was also the case for the monozygotic twins in this study who shared the same *de novo* nonsense mutation. Case 13a had a unilateral cleft lip and palate while case 13b had a bilateral cleft lip and palate. Case 13a had growth retardation while 13b had normal growth; however this could be explained by different intrauterine conditions. Furthermore case 13a had necrotizing enterocolitis and a perforated intestine, conditions that were not seen in case 13b. The heart malformation of case 13a comprised complete atrio-ventricular septal defect and a common large atrio-ventricular vault with mild insufficiency. Case 13b had a large atrio-septal defect and almost joint atria and an Epstein malformation of the tricuspid valve.

Unusual phenotypes

Two patients (cases 4 and 6) with immunological abnormalities were included in study II. Deletion of 22q11.2 had previously been excluded in both cases. There had been a previous report that seven of ten fetuses with truncating *CHD7* mutation had thymic hypoplasia⁸⁵ and there was a report of two CHARGE patients with nonsense mutations in *CHD7* who had severe T-cell deficiency⁸⁶. Case 4 in our study died within the first

year of life. His immunological abnormalities included hypoplasia of the thymus gland and severe T-cell deficiency. In addition, he had hydronephrosis, hypoparathyroidism and neonatal hypocalcemia. Case 4 had a *de novo* deletion of exon 4 and he also carried a maternally inherited two amino-acid-duplication in exon 3. Case 6 who had a nonsense mutation (p.Y913X) in exon 10, died at 12 months of age. He had a typical CHARGE syndrome phenotype but also displayed a severe T-cell deficiency, hypoparathyroidism, gastroesophageal reflux and a double aortic arch.

Two atypical patients, cases 8 and 26, with neither coloboma nor choanal atresia (or cleft lip and palate, that in some cases can substitute for choanal atresia since the two defects rarely occur together¹⁴) were included in this study. Thereby they neither had classical or typical CHARGE syndrome nor fulfilled Pagon's original diagnostic criteria. In one of these patients, case 8, a nonsense mutation (p.W1099X) in exon 13 was identified. She had ear abnormality, temporal bone malformation and swallowing difficulty (a sign of cranial nerve abnormality⁸⁷). In addition she had hearing deficit, genital abnormality, retardation of growth, mild DD and mild difficulties in nasal breathing, although no choanal atresia was present. She thus had two major criteria according to Blake and one major criterion according to Verloes. One similar case had previously been reported; a girl who had slightly dysmorphic ears, severe hearing impairment, bilateral agenesis of the semicircular canals and DD. She had a nonsense mutation in exon 29⁶⁸. Both these cases showed that *CHD7* mutations can be found in patients with an atypical phenotype.

Deletion of chromosome 11q13.4-q14.3

The patient described in paper III had an 18.2 Mb *de novo* deletion of chromosome 11q13.4-q14.3, rarely reported in the literature.

This 3½-years-old boy was born after an unremarkable pregnancy. He had moderate DD, microcephaly and dysmorphic facial features including a broad nasal base, epicanthus, thin lips, large ears, brachycephaly, a round face with a short middle face and bilateral ptosis (figure 13). Other symptoms included a submucous cleft palate, an undescended testis, bilateral inguinal hernia and generalized seizures. MRI examination showed no abnormalities of the brain and ophthalmological examination revealed a mild strabismus and refraction error but was otherwise normal. He had a happy disposition in combination with a hyperactive behaviour and sleeping disorder.

Reports of chromosomal imbalances in the 11q13.4-q14.3-region are scarce and in some cases the fine mapping of the aberration is uncertain (figure 14). Joyce *et al.*⁸⁸ found a *de novo* deletion of 11q13.5-14.2 in a boy with a clinical diagnosis of the Williams-Beuren syndrome. Both Joyce's case and our patient had moderate DD and sociable personalities, as well as full cheeks, long philtrum and prominent ear lobules. Our patient had microcephaly and Joyce's case had micrognathia.



Figure 13 Frontal and profile photo of the patient described in paper III at age 1 year 10 months. Wincent *et al.*, 2010⁸⁹.

The position of a deletion reported by Klep-de Pater *et al.*⁹⁰ is uncertain and the patient's mother was reported to take drugs and drink alcohol in unknown quantities during the pregnancy, making a correlation between these two patients vague although the girl, as our patient, had DD, hypotonia, ptosis and a submucous cleft palate. Guc-Scekic *et al.* reported a two months old patient with a deletion of 11q13-q21 that had DD and feeding difficulties in common with our patient⁹¹. However, both the young age of Guc-Scekic's case when described and the difference in size of the deletions makes it difficult to compare the two patients.

In addition, two more cases with deletions within the region were listed in ECARUCA. Case ID 4366 was a 6-year old boy with a 7.5 Mb deletion of 11q14.1-q14.1 and Case ID 3945 was a 2½-year old boy with a deletion of 11q14.1-q14.2. These two cases had smaller deletions compared to our patient and they both had mild DD and had some facial features in common with our patient. Case 4366 had ptosis and epicanthic folds and case 3945 had full cheeks, large ear lobules and thin lips, all of which were displayed by our patient. In DECIPHER three deletions in the 11q13.4-q14.3-region was listed, the largest being 0.37 Mb. However, limited clinical data was available for these patients and it remains unclear whether the deletions are causal to the phenotype of these patients or are rare benign variants.

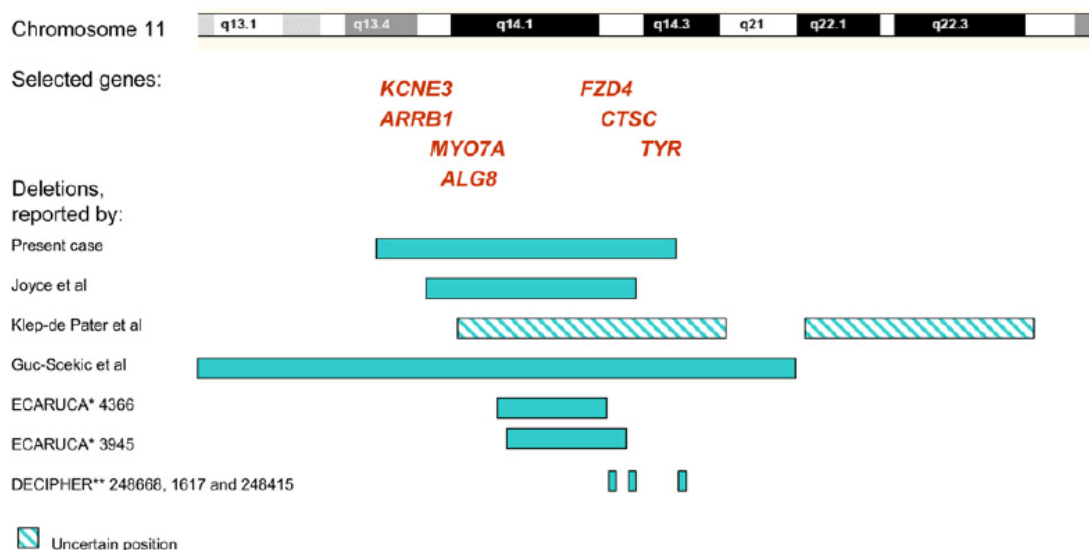


Figure 14. Schematic representation of reported deletions comprising 11q13.4eq14.3 and seven of the genes listed in the region. The breakpoints are in most of the cases uncertain. Wincent *et al.*, 2010⁸⁹.

Overall, genotype-phenotype correlations were difficult to establish due to the paucity of reported cases and lack of adequate mapping data in some of the cases. If all cases are taken into account, there are some overlapping phenotypic features observed including mild-moderate DD, a sociable personality and dysmorphic facial features including full cheeks and prominent ear lobules. Reporting accurate clinical and molecular data of more patients with deletions in the 11q13-q14-region is needed for better genotype-phenotype correlation.

Array-CGH in a clinical setting

Patients with mild, moderate and severe DD had similar diagnostic yield of causal imbalances, 13.8%, 13.3% and 13.6% respectively (figure 15). Although the diagnostic yield among the patients with severe and moderate DD would be expected to be higher than among the patients with mild DD^{16,17,92}, we observed approximately equal diagnostic yields. This could imply that there are no great differences in diagnostic yield between the groups, but it could also reflect an ascertainment effect; patients with severe DD may primarily have undergone other investigations leading to an aetiology-based diagnosis. Another explanation might be an increased detection of duplications and detection of smaller deletions that may cause less severe phenotypes in general. Nonetheless, array-CGH investigation should be offered to all MR-patients, irrespective of the level of DD.

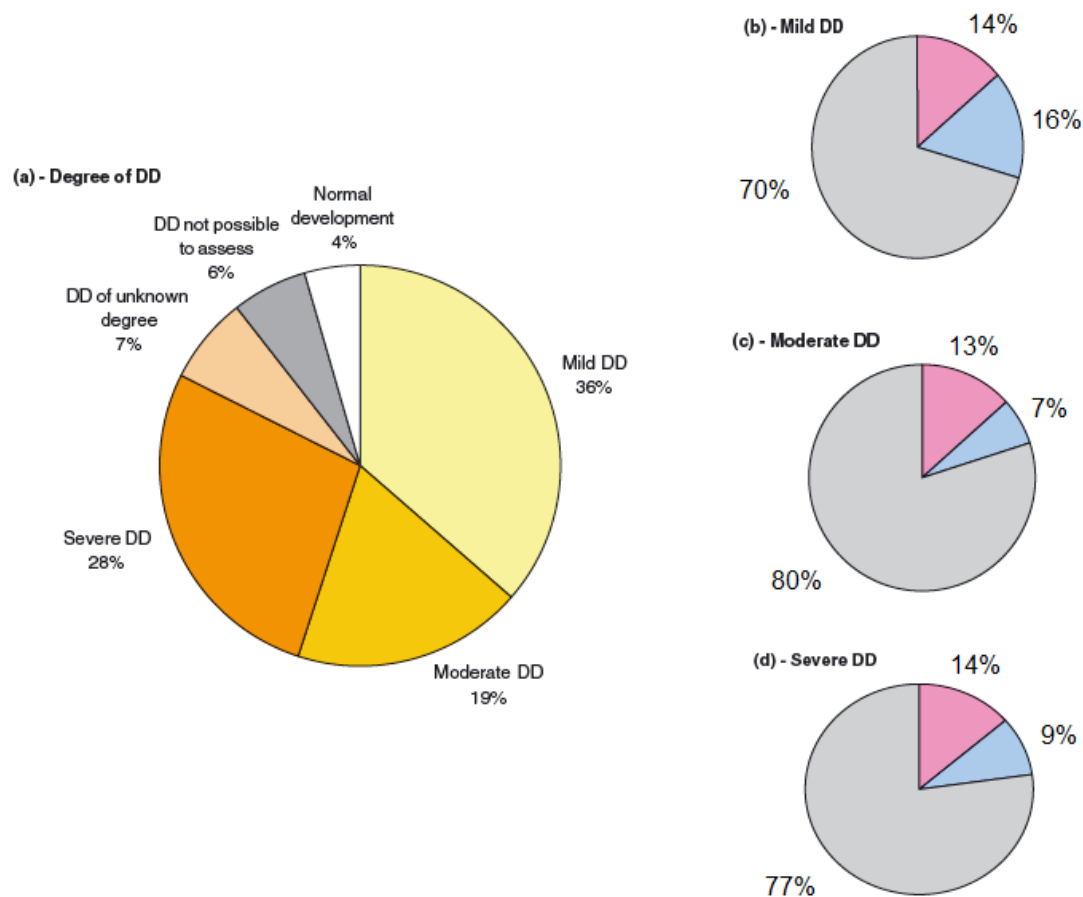


Figure 15. Severity of DD and diagnostic yield. **A:** allotment of DD among the patients in paper IV. **B-D:** percentages of causal CNVs (pink), CNVs of uncertain clinical significance (blue) and not causal CNVs (grey) among the patients with mild (b), moderate (c) and severe (d) DD. Modified from Wincent *et al.*, 2010⁹³.

Many CNVs of unclear significance are very rare and may therefore not be reported in neither healthy nor affected individuals and in addition, many CNVs are population specific⁹⁴. In study IV we identified 16 aberrations of unknown clinical significance. It cannot be excluded that they contribute to the patients phenotypes but in the majority of cases they were inherited, not overlapping with known syndromes and the few genes included were not strong candidate genes to contribute to the phenotypes demonstrated.

In two cases with aberrations inherited from apparently healthy parents the interpretation was somewhat more complicated because of the regions involved. Case 27 had a complex heart malformation, moderate DD, dysmorphic features and seizures with EEG-changes. A 290 Kb deletion of chromosome 7q35 (two exons of *CNTNAP2*) and a 160 Kb duplication of 15q15.1, both inherited from a healthy mother, were identified. At the time of this study *CNTNAP2* had been associated with epilepsy, schizophrenia and autism spectrum disorder^{95,96} and was a plausible candidate gene. However, there are now reports of small deletions in *CNTNAP2* in DGV, making it unlikely that heterozygous deletions in this gene are causal. The 15q15.1-duplication might contribute to the heart malformation since it involves a gene that is highly expressed in the heart.

Case 31 was a patient with mild DD, epilepsy, scoliosis and tall stature. Duplications of 0.8 Mb of chromosome 16p13.11 was identified in the patient, her sister (who had Asperger syndrome) and in their apparently healthy father. At the time of this study, there were reports suggesting that duplications of this region are causal but showing incomplete penetrance^{97,98}. However, Hannes *et al.*⁹⁹ showed that the duplications did not co-segregate with phenotype, and found them in a control-population at a rate that was not significantly different from that in patients and this has later also been shown by Cooper *et al.*⁷⁸. The duplications might have a phenotypic effect with variable expression (controls could have a phenotype that has passed unnoticed) or could work in combination with other predisposing factors to give a phenotype⁹⁹. Nevertheless, it today seems unlikely that duplications in this region are strongly associated with NDDs.

While it is easy to assume that *de novo* alterations result in the observed phenotype, only the recurrent association of imbalances with specific phenotypic features may reinforce this causal relation for a majority of alterations. Case 34 was a boy with mild DD and a severe speech and language disturbance that had a *de novo* 1.2 Mb duplication of chromosome 22q11.23 (see paper V). Despite the fact that the duplication in our case is *de novo* it remained unclear whether it contributed to the patient's phenotype because the few patients with micro-duplications overlapping our patient's duplication so far reported in the literature demonstrated a highly variable phenotype and were in the majority of cases inherited from healthy parents^{32,33}.

Hence, it will be essential to collect genotypic and phenotypic information on a large number of patients with duplications of this region as well as for other aberrations found in patients with DD/MCA. The above-mentioned examples illustrate that although array-CGH is of great value in the clinical setting, interpreting the results can be difficult.

Distal 22q11.2 duplications

The clinical phenotypes of the patients in study V were variable with one of the mildest affected individual displaying mild DD and speech delay while one of the most severely affected individual had severe MR, epilepsy, autism and a brain malformation. However, a majority of cases displayed speech disturbances and various degrees of DD, ranging from mild to severe. Other clinical features present in more than 5 cases included behavioural problems, hypotonia and dysmorphic facial features. Notably, none of the patients in our study had a diagnosed congenital heart defect.

Case 16 in our study, with a *de novo* duplication involving LCRF–H, has a phenotype that is highly concordant with that of patient 14 in the study by Coppinger et al.³³. Both cases had speech impairment and they have a similar facial appearance although neither shows evident facial dysmorphic features. However, case 10, a 35-years old male who also had an F–H duplication had a significantly more severe phenotype. He spoke his first words at 4 years of age and although he was later able to speak, he has had a severe decline in functioning since age 25 and can no longer speak. Furthermore, he had a gastrostomy because of severe difficulties with swallowing, and he is now wheelchair dependent due to progressive spasticity. There is a high suspicion of a mitochondrial disorder, although this could not be confirmed by genetic and mitochondrial tests. Patient 10 likely has additional factors accounting for the aetiology of his phenotype besides the 22q11.2 duplication.

It is noteworthy that 6 of the 10 patients with E/F–H duplications in our study had a speech delay. It may be that distal 22q11.2 duplications are associated with an increased risk for speech delay while the additional more severe phenotypes seen in some of the patients are because of additional not yet identified factors.

Although there are now more than 35 index cases with distal 22q11.2 micro-duplications (including the patients from this study) reported in the literature, extended investigations of families harbouring these duplications are needed to provide insight into the pathogenicity of these duplications. There is an urgent need for ascertainment of risk figures for phenotypic abnormality in individuals with 22q11.2 distal duplications to help alleviate the current interpretational challenges for diagnostic testing and counselling.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The results of the studies included in this thesis show the value of array-CGH when investigating patients with NDDs, both in the clinical- and research setting, by identifying causal aberrations, candidate genes and fine-mapping of aberrations. Furthermore, the work in this thesis has led to the identification of an aetiological genetic diagnosis in more patients with idiopathic NDDs, although it also stresses the difficulty in interpretation of certain aberrations before there is an established link between a specific genotype and phenotype.

Deletions of 11q13.4-q14.3 may soon be a recognizable syndrome as more patients with deletions in this region are reported. After our article was published there has been another report of a patient with a deletion in the same region, showing phenotypic similarities to our patient¹⁰⁰.

Although both CdLS and CHARGE syndrome remain primarily clinical diagnoses the molecular techniques may help with the diagnosis in atypical cases. Furthermore, an aetiological diagnosis improves the genetic counselling in terms of prognosis and recurrence risk estimations and enables prenatal testing. However, in a large proportion of cases the aetiology is still unknown. In general, the identification of autosomal dominant genes for NDDs is complicated by the fact that they manifest themselves mostly in isolated cases. The new sequencing techniques that have been developed in the last years may be the solution for many cases.

Whole-genome/-exome sequencing will likely find genetic causes in many of the patients with NDDs of unknown aetiology. However, there are technical challenges of these sequencing techniques when utilized in a diagnostic setting such as that the technology has been focused on high throughput more than on accuracy. Because of the vast amount of data generated, the data needs to be filtered, but the bioinformatic approaches and tools are not yet standardized. For example, the same raw data can give different end results depending on the bioinformatic processing. Also, still a relative large number of false positives are generated and needs to be investigated by Sanger sequencing before being disregarded, which makes the analysis costly.

Because of the enormous amount of information generated, many studies have initially used exome sequencing rather than whole-genome sequencing. If exome sequencing will not identify new genes causative of CdLS and CHARGE syndrome it seems reasonable to start to explore the already causative genes further, especially looking for mutations in highly conserved intronic regions, as has successfully been done recently for the immune defect disorder familial hemophagocytic lymphohistiocytosis¹⁰¹.

The new concept of molecular karyotyping has significantly changed the field of clinical cytogenetics in the last decade and although large chromosomal aberrations are still easier detected by array-CGH than by the new sequencing methods, the question arises which position array-CGH will have in the future next to whole genome sequencing. Eventually, the new sequencing techniques may replace the microarrays as the main tool in the genetic diagnostic process of NDDs. However, even if Array-CGH will no longer be used in clinical practice we will surely have benefited from the lessons learnt from the array era considering the experience gained on how to interpret the aberrations.

When array-CGH first became available, inheritance was very important for the interpretation. However, it is becoming increasingly evident that caution is recommended when interpreting the causality of CNVs based on inheritance. There are examples of identification of CNVs believed to be causal much because they were *de novo*, but later other clearly causal aberrations were identified¹⁰². Therefore, the continuous identification of patients with overlapping aberrations and thorough description of their clinical presentation will be essential for confirmation of the causal role of aberrations, no matter which method is used.

There are still ethical issues with regards to whole-genome array testing and they will likely not be fewer with the new sequencing techniques. For example, uncertain information may lead to increased anxiety and unnecessary clinical procedures. The methods may also bring the possibility to “screen for imperfections” prenatally. Most public healthcare providers would likely not allow this, but maybe commercial clinics will. However, the question is important for all of society since there is of course a grey-scale between a serious disability and an “imperfection”. Who is to decide where to draw the line? Also, the finding of an aberration may not always be positive but may lead to unwanted labelling and stigmatization. Then there is the potential of finding incidental findings, i.e. aberrations that have a clinical consequence for the patient but does not explain the phenotype the patient was investigated for. However, this may in part be solved by masking certain genes/regions from analysis.

How to deal with the interpretation of multiple rare variants in one individual and to estimate how great a risk factor certain CNVs, such as distal 22q11.2 duplications, actually are, will be challenges for future projects. It may be that for patients with a family history of NDDs, the presence of a 22q11.2 duplication is more likely to result in symptoms, than if there is no family history of NDDs, as the family history may be indicative of other predisposing factors in the family which could represent a “second hit”.

Hopefully, increased insight into genetic aberrations and pathways causing MR may result in the development of specifically targeted therapies. In fact, already some strategies are being developed. For example, a reduction in metabotropic glutamate receptor signalling has been shown to reverse clinical features of fragile-X syndrome in animal models¹⁰³. Discoveries regarding the genetics of tuberous sclerosis enabled the creation of a mouse model for the disease and the testing of potential therapeutic agents, which has shown that these agents may be used for treatment¹⁰⁴. Another example where understanding of the genetic basis of a disease is used for development of therapeutic strategies has recently been shown for Angelman syndrome. Angelman syndrome is caused by loss of the maternal copy of *UBE3A*, but in mice it seems as if the topoisomerase inhibitor topotecan can “unsilence” the epigenetically silenced paternal *UBE3A* allele¹⁰⁵. These encouraging results gives hope that further insight into the genetics of NDDs may pave the way for the development of novel treatment strategies.

In conclusion, the work of this thesis has contributed with clinical and molecular data on patients with Cornelia de Lange syndrome, CHARGE syndrome, duplications of distal 22q11.2 and deletion of chromosome 11q13.4-q14.3. Since these are rare disorder, most clinicians will not have encountered many patients with the exact same aberration and therefore this information has an immediate impact for clinicians and patients in the genetic counselling situation.

SAMMANFATTNING PÅ SVENSKA

Utvecklingsstörning är ett funktionshinder som innebär att en individs kognition inte utvecklas normalt. Barn med utvecklingsstörning har betydligt svårare att ta in och tolka information, uttrycka sina tankar och känslor samt att anpassa sig till nya situationer än andra barn. Barnen är ofta försenade i den motoriska utvecklingen och vänder sig, sitter och går senare än andra barn. Antalet barn med utvecklingsstörning har under de senaste årtiondena, tack vare både medicinska och sociala framsteg, minskat i Sverige och prevalensen ligger nu uppskattningsvis på cirka 1%. Framförallt har utvecklingsstörning på grund av pre- och perinatale infektioner samt traumatiska förlossningar minskat. Många av patienterna har även associerade symptom såsom missbildningar, tillväxthämning, utseendemässiga särdrag och epilepsi.

Trots att barn med utvecklingsstörning genomgår många kliniska och laborietekniska undersökningar kan orsaken till deras symptom endast fastställas i ca hälften av fallen. Detta är ofta frustrerande för barnens familjer eftersom en etiologisk diagnos inte bara förklarar varför barnet är drabbat utan också möjliggör bättre bedömning avseende prognos och en mer tillförlitlig upprepningsrisk vid en ny graviditet.

Det finns mycket som talar för att orsaken till en stor del av utvecklingsstörning hos denna heterogena grupp av patienter har genetisk bakgrund. Trisomi 21 och fragilt X mental retardation, de två vanligaste orsakerna till utvecklingsstörning, har genetisk etiologi och utvecklingsstörning är ett vanligt symptom vid många välkända genetiska sjukdomar. Att många barn med utvecklingsstörning även har kongenitala missbildningar eller utseendemässiga särdrag indikerar också en konstitutionell, möjligen genetisk, bakgrund.

Genetiska orsaker brukar delas in i kromosomavvikelse (t.ex. deletion eller duplikation av en hel kromosom eller delar av en kromosom) och monogena avvikelser (t.ex. punktmutationer som påverkar genens uttryck). Kromosomavvikelser som är så stora att de kan ses när man studerar kromosomerna i ett ljusmikroskop (t.ex. Trisomi 21) påvisas hos ca 11% av patienterna med utvecklingsstörning. Hos ca 9% av patienterna kan symptomen kopplas till en monogen sjukdom.



Figur 16. Orsaker till utvecklingsstörning, baserat på en studie av Stevenson 2003¹⁹.

Det övergripande **syftet** med detta doktorandprojekt var att uppnå en förbättrad diagnostik för patienter med utvecklingsförsening genom att identifiera genetiska avvikelser hos patienter där man tidigare inte kunnat hitta orsaken till utvecklingsstörningen.

I **Arbete I** använde vi oss av DNA-sekvensering och MLPA för leta efter mutationer i genen *NIPBL* hos 11 svenska patienter med Cornelia de Lange Syndrom (CdLS). Detta syndrom kännetecknas bland annat av karakteristiska ansiktsdrag, tillväxthämning och måttlig till svår utvecklingsstörning. År 2004 visade amerikanska och brittiska forskargrupper att mutationer i genen *NIPBL* fanns hos en stor andel av patienterna med CdLS. I vår studie hittade vi 7 mutationer, varav bara en hade rapporterats tidigare. En av mutationerna satt inte i den kodande delen av genen, men borde teoretiskt påverka det färdiga proteinet. Genom analys av RNA kunde vi visa att exon 16 klipptes bort under processandet av RNA i cellkärnan, vilket leder till ett defekt protein.

Denna studie konfirmerade att *NIPBL*-mutationer kan påvisas hos en hög andel av patienter med CdLS, men vi kunde inte påvisa en tydlig genotyp-fenotyp korrelation, det vill säga att vissa typer av mutationer skulle orsaka en svårare, eller mildare, fenotyp. Efter denna studie infördes sekvensering av *NIPBL* som ett kliniskt diagnostiskt test på avdelningen för klinisk genetik på Karolinska Universitetssjukhuset.

I **Arbete II** använde vi oss av DNA-sekvensering och MLPA för att undersöka genen *CHD7* hos 30 svenska patienter med CHARGE syndrom. Karakteristiskt för CHARGE syndrom är framförallt kolobom, koanalatri, annorlunda formade ytteröron och underutveckling av balansorganet i innerörat. En nederländsk forskargrupp fann år 2004 att mutationer i genen *CHD7* orsakar CHARGE syndrom. Vi identifierade patogena mutationer hos 64% av våra patienter. Dessutom hittade vi en nedärvd mutation av oklar signifikans och tre stycken nedärvda mutationer som sannolikt är ovanliga normalvarianter. Vår studie illustrerade att det ibland kan vara svårt att avgöra den kliniska betydelsen av mutationer även då de påvisas i en förväntad gen.

Två av patienterna i studien uppfyllde inte de kliniska kriterierna för CHARGE syndrom men inkluderades för att en erfaren barnläkare ändå bedömde att de hade CHARGE syndrom. Hos en av dessa patienter hittade vi en patogen mutation. Detta vidgar det fenotypiska spektrumet associerat med CHARGE syndrom och visar att även atypiska patienter bör undersökas för förekomst av *CHD7*-mutation. *CHD7*-mutationsdiagnostik finns nu tillgänglig som ett kliniskt diagnostiskt test på avdelningen för klinisk genetik på Karolinska Universitetssjukhuset.

I **Arbete III** rapporterar vi noggrant symptomen hos en patient med en deletion av kromosomregionen 11q13.4-q14.3. Deletioner av denna region är mycket ovanliga och jämförelse med de få deletioner som fanns beskrivna i litteraturen var svår eftersom det ofta saknades information om deletionernas exakta position och storlek. Några gemensamma kliniska drag fanns dock hos de patienter som hade de mest överlappande deletionerna. Dessa symptom var mild till måttlig utvecklingsstörning, en social personlighet och vissa utseendemässiga särdrag. Om fler patienter med deletioner i detta område rapporteras kan det i framtiden leda till etableringen av ett nytt kliniskt igenkännbart syndrom.

I **Arbete IV** utvärderar vi införandet av array-CGH för klinisk undersökning av patienter med utvecklingsstörning. Med den relativt nya tekniken array-CGH kan man med en

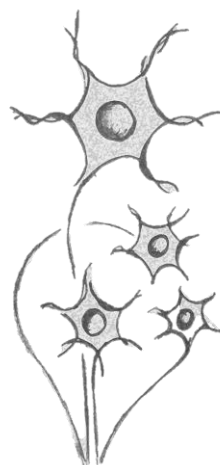
upplösning som är mycket högre än upplösningen för konventionell kromosomanalys leta efter gendosavvikelser i hela genomet. De första 160 patienterna som undersöktes kliniskt med array-CGH på avdelningen för klinisk genetik på Karolinska Universitetssjukhuset inkluderades i studien.

Vi identifierade 21 patogena gendosavvikelser och 15 gendosavvikelser av oklar klinisk signifikans. Sju av de patogena avvikelserna var tillräckligt stora för att kunna ses i ljusmikroskop men hade ändå inte hittats vid tidigare undersökning med kromosomanalys. Detta illustrerar att noggrannare och mer pålitliga metoder, såsom array-CGH, behövdes för att påvisa gendosavvikelser hos patienter med utvecklingsstörning. En nackdel med metoden var att vi hittade många avvikelser av oklar klinisk signifikans.

I **Arbete V** beskriver vi 16 patienter från Australien, Nederländerna, England, Italien och Sverige som har duplikationer av kromosomregionen 22q11.22-q11.23. Det fanns innan vår studie 22 patienter rapporterade i litteraturen med duplikationer av detta område. Dessa patienter hade en mycket varierande fenotyp och duplikationerna var ofta nedärvda från föräldrar med få eller inga symptom.

Även i vår studie hade patienterna symptom av varierande svårighetsgrad och en hög andel av duplikationerna var nedärvda (83%). Att duplikationerna så ofta är nedärvda gör att det är svårt att säga om de är sjukdomsassocierade eller inte, men de är sällan rapporterade som normalvarianter hos friska kontrollindivider. Vi anser att dessa duplikationer är riskfaktorer för utvecklingsstörning och kommunikationssjukdomar, även om andra genetiska och/eller icke genetiska faktorer krävs för att orsaka fenotypen hos patienterna. Hos fem av patienterna hittade vi ytterligare gendosavvikelser som skulle kunna vara en sådan modifierande faktor.

Sammanfattningsvis har detta doktorandprojekt bidragit till att fler patienter fått en etiologisk diagnos och att information om deras symptom och genetiska avvikelser finns tillgängliga för läkare runt om i världen. Detta ger bättre medicinskt underlag för bedömning av patienter med dessa ovanliga genetiska avvikelser. Dessutom har vi illustrerat hur svårt det kan vara att tolka en genetisk avvikelse innan det finns en etablerad länk mellan genotyp och fenotyp. Mer forskning och rapportering av ovanliga avvikelser tillsammans med detaljerade kliniska beskrivningar av patienterna behövs för att etablera denna länk.



ACKNOWLEDGEMENTS

First, I would like to express my gratefulness to all the patients and their families that participated in the studies, and to our collaborators in Sweden and around the globe.

My profound gratitude to my supervisors without whom this work neither would have been started nor finished;

- ✧ *Jacqueline Schoumans*, for accepting me as summer-student and being willing to keep me working on the project. For teaching me different molecular techniques and especially the art of array-CGH. Thank you for sharing your vast knowledge and experience during this process and for your excellent comments that always improves the papers tremendously.
- ✧ *Magnus Nordenskjöld*, for having faith in me from the beginning of this work and for making it possible for me to combine my medical studies with research. For providing me with guidance and support during the whole process. Thank you for thinking that having children is a good thing even if you are a PhD-student, and for always finding solutions to problems.
- ✧ *Britt-Marie Anderlid*, for being inspiring and enthusiastic, it has always been so fun telling you when I managed to find a mutation! For helping me untangling vast amounts of clinical information and pinpointing the important clues of a patient's phenotype. Thank you for all discussions and the encouragement you have given me and most of all for being a great role model.

I also want to direct my appreciation to many persons at CMM and the department of clinical genetics who have in different ways contributed to the work in this thesis:

- ✧ Fellow researcher friends at CMM, old and new, for sharing times of troubles, frustration and laughter. Especially to *Aron Luthman* for your positive attitude and for working so hard on the Rubinstein-Taiby- and epilepsy projects. My warmest appreciation to *Anna Bremer, Johanna Winberg, Marie Meeths, Mårten Winge, Miriam Entesarian, Susanna von Holst, Anna Lindstrand, Michela Barbaro, Josefina Edner, Eva Rudd, Fredrik Lundberg, Mai-Britt Giacobini, Johanna Rantala, Vasilios Zachariadis, Tobias Laurell* and *Malin Kvarnung*. Thank you for always making it fun to come to work!
- ✧ *Agne Liedén* and *Johanna Lundin* for introducing me to clinical work with the arrays and for having patience with my never- ending questions. Also a special thanks to *Anh-Nhi Tran, Kristina Lagerstedt-Robinson* and *Helena Malmgren*, for your great knowledge about genetics and always being helpful in sharing it.
- ✧ *Ann Nordgren, Peter Gustavsson* and *Giedre Grigelioniene* for being great inspirations.
- ✧ *Margareta Lagerberg* and *Ulla Grandell*, my go-to-ladies when it comes to practical array and sequencing issues, for being so kind and for always managing to find time to help me. Also thank you to *Sigrid Sahlén, Anna-Lena Kastman, Christina Nyström, Isabel Neira, Ingela Forsberg, Helén Nee, Irene White, Inger Malmberg*

and *Karin Kindberg* for much help with practical issues and advice on numerous topics. To *Ellika Sahlin* and *Anna Hammarsjö* for being truly mood-brightening!

- ✧ I have had many troubles with computer problems during these years, as *Rudolf Matousek*, *Lennart Helleday* and *Jan-Erik Kaare* are probably too well aware of. Thank you for trying to solve them and for trying to explain to me how I should avoid them. I am also very grateful to the kind staff at the MMK and CMM administrations.

Finally, thank you to all my friends and family, and especially to those that have patiently listened to my joy and my complaints about the research;

- ✧ My mother *Catharina*, my father *Bo* and *Marianne* who have always taken an interest in my research. Also a huge thank you for taking such good care of Sam on many occasions. Thanks also to all other family members and relatives that have supported and believed in me.
- ✧ My friends from medical school, *Stina*, *Jennifer*, *Petter*, *Marie*, *Sebastian*, *Jonathan*, *Robert*, *Hannes* and *Nina*, that have contributed with everything from sharing lecture notes and intensive hours of studying, to Oktoberfests and numerous potlucks.
- ✧ *Jeannette*, my hilarious dissertation-guru during the past months, and of course so much more than that. *Linda* and *Johanna* for sharing the enthusiasm for lab-work and for all the times of sharing thoughts about everything and nothing over a glass of wine.
- ✧ *Ylva*, for somehow always managing to be close by and involved, although physically most often being in another country.
- ✧ *Tom*, for all your encouragement and support, in big and small, practical and philosophical. Thank you for endless discussions, for challenging my way of thinking, and for your amazing ability to make me laugh.



REFERENCES

- 1 Goldstein, S. Personal communication. *Author of "Handbook of Neurodevelopmental and Genetic Disorders in Children", 2011, (2012).*
- 2 DSM-V. Neurodevelopmental Disorders,
<<http://www.dsm5.org/proposedrevision/Pages/NeurodevelopmentalDisorders.aspx>> (2012).
- 3 DSM-IV. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. (*American Psychiatric Association*, 2000).
- 4 ICD-10. International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) Version for 2010, Mental retardation (F70-F79), <<http://apps.who.int/classifications/icd10/browse/2010/en#/F70-F79>> (2010).
- 5 Shevell, M. Global developmental delay and mental retardation or intellectual disability: conceptualization, evaluation, and etiology. *Pediatr Clin North Am* 55, 1071-1084, xi, (2008).
- 6 Leonard, H. & Wen, X. The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment Retard Dev Disabil Res Rev* 8, 117-134, (2002).
- 7 Hagberg, B. & Kyllerman, M. Epidemiology of mental retardation--a Swedish survey. *Brain Dev* 5, 441-449, (1983).
- 8 Shaffer, L. G. American College of Medical Genetics guideline on the cytogenetic evaluation of the individual with developmental delay or mental retardation. *Genet Med* 7, 650-654, (2005).
- 9 McLaren, J. & Bryson, S. E. Review of recent epidemiological studies of mental retardation: prevalence, associated disorders, and etiology. *Am J Ment Retard* 92, 243-254, (1987).
- 10 Liu, J. & Baynam, G. Cornelia de Lange syndrome. *Adv Exp Med Biol* 685, 111-123, (2010).
- 11 Allanson, J. E., Hennekam, R. C. & Ireland, M. De Lange syndrome: subjective and objective comparison of the classical and mild phenotypes. *J Med Genet* 34, 645-650, (1997).
- 12 Issekutz, K. A., Graham, J. M., Jr., Prasad, C. *et al.* An epidemiological analysis of CHARGE syndrome: preliminary results from a Canadian study. *Am J Med Genet A* 133, 309-317, (2005).
- 13 Pagon, R. A., Graham, J. M., Jr., Zonana, J. *et al.* Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. *J Pediatr* 99, 223-227, (1981).
- 14 Blake, K. D., Davenport, S. L., Hall, B. D. *et al.* CHARGE association: an update and review for the primary pediatrician. *Clin Pediatr (Phila)* 37, 159-173, (1998).
- 15 Verloes, A. Updated diagnostic criteria for CHARGE syndrome: a proposal. *Am J Med Genet A* 133A, 306-308, (2005).
- 16 Curry, C. J., Stevenson, R. E., Aughton, D. *et al.* Evaluation of mental retardation: recommendations of a Consensus Conference: American College of Medical Genetics. *Am J Med Genet* 72, 468-477, (1997).
- 17 Moog, U. The outcome of diagnostic studies on the etiology of mental retardation: considerations on the classification of the causes. *Am J Med Genet A* 137, 228-231, (2005).

- 18 Rauch, A., Hoyer, J., Guth, S. *et al.* Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet A* 140, 2063-2074, (2006).
- 19 Stevenson, R. E., Procopio-Allen, A. M., Schroer, R. J. *et al.* Genetic syndromes among individuals with mental retardation. *Am J Med Genet A* 123A, 29-32, (2003).
- 20 Flint, J. & Wilkie, A. O. The genetics of mental retardation. *Br Med Bull* 52, 453-464, (1996).
- 21 Rooms, L., Reyniers, E. & Kooy, R. F. Subtelomeric rearrangements in the mentally retarded: a comparison of detection methods. *Hum Mutat* 25, 513-524, (2005).
- 22 Kallioniemi, A., Kallioniemi, O. P., Sudar, D. *et al.* Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818-821, (1992).
- 23 Solinas-Toldo, S., Lampel, S., Stilgenbauer, S. *et al.* Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 20, 399-407, (1997).
- 24 Pinkel, D., Seagraves, R., Sudar, D. *et al.* High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 20, 207-211, (1998).
- 25 Schoumans, J., Ruivenkamp, C., Holmberg, E. *et al.* Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). *J Med Genet* 42, 699-705, (2005).
- 26 Schoumans, J., Staaf, J., Jonsson, G. *et al.* Detection and delineation of an unusual 17p11.2 deletion by array-CGH and refinement of the Smith-Magenis syndrome minimum deletion to approximately 650 kb. *Eur J Med Genet* 48, 290-300, (2005).
- 27 Zhang, Z. F., Ruivenkamp, C., Staaf, J. *et al.* Detection of submicroscopic constitutional chromosome aberrations in clinical diagnostics: a validation of the practical performance of different array platforms. *Eur J Hum Genet* 16, 786-792, (2008).
- 28 Shaffer, L. G., Bejjani, B. A., Torchia, B. *et al.* The identification of microdeletion syndromes and other chromosome abnormalities: cytogenetic methods of the past, new technologies for the future. *Am J Med Genet C Semin Med Genet* 145C, 335-345, (2007).
- 29 Slavotinek, A. M. Novel microdeletion syndromes detected by chromosome microarrays. *Hum Genet* 124, 1-17, (2008).
- 30 Koolen, D. A., Vissers, L. E., Pfundt, R. *et al.* A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* 38, 999-1001, (2006).
- 31 Portnoi, M. F. Microduplication 22q11.2: a new chromosomal syndrome. *Eur J Med Genet* 52, 88-93, (2009).
- 32 Descartes, M., Franklin, J., Diaz de Stahl, T. *et al.* Distal 22q11.2 microduplication encompassing the BCR gene. *Am J Med Genet A* 146A, 3075-3081, (2008).
- 33 Coppinger, J., McDonald-McGinn, D., Zackai, E. *et al.* Identification of familial and de novo microduplications of 22q11.21-q11.23 distal to the 22q11.21 microdeletion syndrome region. *Hum Mol Genet* 18, 1377-1383, (2009).
- 34 Ou, Z., Berg, J. S., Yonath, H. *et al.* Microduplications of 22q11.2 are frequently inherited and are associated with variable phenotypes. *Genet Med* 10, 267-277, (2008).

- 35 Shimojima, K., Imai, K. & Yamamoto, T. A de novo 22q11.22q11.23
interchromosomal tandem duplication in a boy with developmental delay,
hyperactivity, and epilepsy. *Am J Med Genet A* 152A, 2820-2826, (2010).
- 36 Wentzel, C., Fernstrom, M., Ohrner, Y. *et al.* Clinical variability of the 22q11.2
duplication syndrome. *Eur J Med Genet* 51, 501-510, (2008).
- 37 Feuk, L., Carson, A. R. & Scherer, S. W. Structural variation in the human
genome. *Nat Rev Genet* 7, 85-97, (2006).
- 38 Zhang, F., Gu, W., Hurles, M. E. *et al.* Copy number variation in human health,
disease, and evolution. *Annu Rev Genomics Hum Genet* 10, 451-481, (2009).
- 39 Shaikh, T. H., O'Connor, R. J., Pierpont, M. E. *et al.* Low copy repeats mediate
distal chromosome 22q11.2 deletions: sequence analysis predicts breakpoint
mechanisms. *Genome Res* 17, 482-491, (2007).
- 40 Shaikh, T. H., Kurahashi, H., Saitta, S. C. *et al.* Chromosome 22-specific low
copy repeats and the 22q11.2 deletion syndrome: genomic organization and
deletion endpoint analysis. *Hum Mol Genet* 9, 489-501, (2000).
- 41 Emanuel, B. S. Molecular mechanisms and diagnosis of chromosome 22q11.2
rearrangements. *Dev Disabil Res Rev* 14, 11-18, (2008).
- 42 Ben-Shachar, S., Ou, Z., Shaw, C. A. *et al.* 22q11.2 distal deletion: a recurrent
genomic disorder distinct from DiGeorge syndrome and velocardiofacial
syndrome. *Am J Hum Genet* 82, 214-221, (2008).
- 43 Krantz, I. D., McCallum, J., DeScipio, C. *et al.* Cornelia de Lange syndrome is
caused by mutations in NIPBL, the human homolog of *Drosophila melanogaster*
Nipped-B. *Nat Genet* 36, 631-635, (2004).
- 44 Tonkin, E. T., Wang, T. J., Lisgo, S. *et al.* NIPBL, encoding a homolog of fungal
Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in
Cornelia de Lange syndrome. *Nat Genet* 36, 636-641, (2004).
- 45 Vissers, L. E., van Ravenswaaij, C. M., Admiraal, R. *et al.* Mutations in a new
member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet*
36, 955-957, (2004).
- 46 van Karnebeek, C. D., Jansweijer, M. C., Leenders, A. G. *et al.* Diagnostic
investigations in individuals with mental retardation: a systematic literature review
of their usefulness. *Eur J Hum Genet* 13, 6-25, (2005).
- 47 Ravnan, J. B., Tepperberg, J. H., Papenhausen, P. *et al.* Subtelomere FISH
analysis of 11 688 cases: an evaluation of the frequency and pattern of
subtelomere rearrangements in individuals with developmental disabilities. *J Med*
Genet 43, 478-489, (2006).
- 48 Gijsbers, A. C., Schoumans, J. & Ruivenkamp, C. A. Interpretation of array
comparative genome hybridization data: a major challenge. *Cytogenet Genome*
Res 135, 222-227, (2011).
- 49 Lee, C., Iafrate, A. J. & Brothman, A. R. Copy number variations and clinical
cytogenetic diagnosis of constitutional disorders. *Nat Genet* 39, S48-S54, (2007).
- 50 Iafrate, A. J., Feuk, L., Rivera, M. N. *et al.* Detection of large-scale variation in
the human genome. *Nat Genet* 36, 949-951, (2004).
- 51 Redon, R., Ishikawa, S., Fitch, K. R. *et al.* Global variation in copy number in the
human genome. *Nature* 444, 444-454, (2006).
- 52 Sebat, J., Lakshmi, B., Troge, J. *et al.* Large-scale copy number polymorphism in
the human genome. *Science* 305, 525-528, (2004).
- 53 Piotrowski, A., Bruder, C. E., Andersson, R. *et al.* Somatic mosaicism for copy
number variation in differentiated human tissues. *Hum Mutat* 29, 1118-1124,
(2008).

- 54 Bruder, C. E., Piotrowski, A., Gijsbers, A. A. *et al.* Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Hum Genet* 82, 763-771, (2008).
- 55 Itsara, A., Wu, H., Smith, J. D. *et al.* De novo rates and selection of large copy number variation. *Genome Res* 20, 1469-1481, (2010).
- 56 Wincent, J., Anderlid, B. M., Lagerberg, M. *et al.* High-resolution molecular karyotyping in patients with developmental delay and/or multiple congenital anomalies in a clinical setting. *Clin Genet*.
- 57 Bruno, D. L., Ganesamoorthy, D., Schoumans, J. *et al.* Detection of cryptic pathogenic copy number variations and constitutional loss of heterozygosity using high resolution SNP microarray analysis in 117 patients referred for cytogenetic analysis and impact on clinical practice. *J Med Genet* 46, 123-131, (2009).
- 58 Saal, L. H., Troein, C., Vallon-Christersson, J. *et al.* BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. *Genome Biol* 3, SOFTWARE0003, (2002).
- 59 Fan, Y. S., Jayakar, P., Zhu, H. *et al.* Detection of pathogenic gene copy number variations in patients with mental retardation by genomewide oligonucleotide array comparative genomic hybridization. *Hum Mutat* 28, 1124-1132, (2007).
- 60 Koolen, D. A. *Copy number variation and mental retardation* PhD thesis, Radboud University Nijmegen Medical Center, (2008).
- 61 Borck, G., Redon, R., Sanlaville, D. *et al.* NIPBL mutations and genetic heterogeneity in Cornelia de Lange syndrome. *J Med Genet* 41, e128, (2004).
- 62 Gillis, L. A., McCallum, J., Kaur, M. *et al.* NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. *Am J Hum Genet* 75, 610-623, (2004).
- 63 Miyake, N., Visser, R., Kinoshita, A. *et al.* Four novel NIPBL mutations in Japanese patients with Cornelia de Lange syndrome. *Am J Med Genet A* 135, 103-105, (2005).
- 64 Bhuiyan, Z., Klein, M., Hammond, P. *et al.* Genotype-Phenotype correlations of 39 patients with cornelia de Lange syndrome: the Dutch experience. *J Med Genet*, (2005).
- 65 Schoumans, J., Wincent, J., Barbaro, M. *et al.* Comprehensive mutational analysis of a cohort of Swedish Cornelia de Lange syndrome patients. *Eur J Hum Genet* 15, 143-149, (2007).
- 66 Lalani, S. R., Safiullah, A. M., Fernbach, S. D. *et al.* Spectrum of CHD7 Mutations in 110 Individuals with CHARGE Syndrome and Genotype-Phenotype Correlation. *Am J Hum Genet* 78, 303-314, (2006).
- 67 Aramaki, M., Udaka, T., Kosaki, R. *et al.* Phenotypic spectrum of CHARGE syndrome with CHD7 mutations. *J Pediatr* 148, 410-414, (2006).
- 68 Jongmans, M. C., Admiraal, R. J., van der Donk, K. P. *et al.* CHARGE syndrome: the phenotypic spectrum of mutations in the CHD7 gene. *J Med Genet* 43, 306-314, (2006).
- 69 Wincent, J., Holmberg, E., Stromland, K. *et al.* CHD7 mutation spectrum in 28 Swedish patients diagnosed with CHARGE syndrome. *Clin Genet* 74, 31-38, (2008).
- 70 Delahaye, A., Sznajer, Y., Lyonnet, S. *et al.* Familial CHARGE syndrome because of CHD7 mutation: clinical intra- and interfamilial variability. *Clin Genet* 72, 112-121, (2007).
- 71 Qin, M., Hayashi, H., Oshima, K. *et al.* Complexity of the genotype-phenotype correlation in familial exudative vitreoretinopathy with mutations in the LRP5 and/or FZD4 genes. *Hum Mutat* 26, 104-112, (2005).

- 72 Sagoo, G. S., Butterworth, A. S., Sanderson, S. *et al.* Array CGH in patients with learning disability (mental retardation) and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects. *Genet Med* 11, 139-146, (2009).
- 73 Kearney, H. M., Thorland, E. C., Brown, K. K. *et al.* American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med* 13, 680-685, (2011).
- 74 Vermeesch, J. R., Brady, P. D., Sanlaville, D. *et al.* Genome-wide arrays: Quality criteria and platforms to be used in routine diagnostics. *Hum Mutat*, (2012).
- 75 Shaffer, L. G. & Lupski, J. R. Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Annu Rev Genet* 34, 297-329, (2000).
- 76 Wincent, J., Bruno, D. L., van Bon, B. W. *et al.* Sixteen New Cases Contributing to the Characterization of Patients with Distal 22q11.2 Microduplications. *Mol Syndromol* 1, 246-254, (2011).
- 77 McDonald-McGinn, D. M., Tonnesen, M. K., Laufer-Cahana, A. *et al.* Phenotype of the 22q11.2 deletion in individuals identified through an affected relative: cast a wide FISHing net! *Genet Med* 3, 23-29, (2001).
- 78 Cooper, G. M., Coe, B. P., Girirajan, S. *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* 43, 838-846, (2011).
- 79 Itsara, A., Cooper, G. M., Baker, C. *et al.* Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 84, 148-161, (2009).
- 80 Girirajan, S., Rosenfeld, J. A., Cooper, G. M. *et al.* A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* 42, 203-209, (2010).
- 81 Girirajan, S., Brkanac, Z., Coe, B. P. *et al.* Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genet* 7, e1002334, (2011).
- 82 Endeley, S., Rosenberger, G., Geider, K. *et al.* Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* 42, 1021-1026, (2010).
- 83 Reutlinger, C., Helbig, I., Gawelczyk, B. *et al.* Deletions in 16p13 including GRIN2A in patients with intellectual disability, various dysmorphic features, and seizure disorders of the rolandic region. *Epilepsia* 51, 1870-1873, (2010).
- 84 Johnson, D., Morrison, N., Grant, L. *et al.* Confirmation of CHD7 as a cause of CHARGE association identified by mapping a balanced chromosome translocation in affected monozygotic twins. *J Med Genet* 43, 280-284, (2006).
- 85 Sanlaville, D., Etchevers, H. C., Gonzales, M. *et al.* Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *J Med Genet* 43, 211-217, (2006).
- 86 Writzl, K., Cale, C. M., Pierce, C. M. *et al.* Immunological abnormalities in CHARGE syndrome. *Eur J Med Genet* 50, 338-345, (2007).
- 87 Dobbela, C., Peacocke, S. D., Blake, K. *et al.* Feeding Difficulties in Children with CHARGE Syndrome: Prevalence, Risk Factors, and Prognosis. *Dysphagia*, (2007).
- 88 Joyce, C. A., Zorich, B., Pike, S. J. *et al.* Williams-Beuren syndrome: phenotypic variability and deletions of chromosomes 7, 11, and 22 in a series of 52 patients. *J Med Genet* 33, 986-992, (1996).
- 89 Wincent, J., Schoumans, J. & Anderlid, B. M. De novo deletion of chromosome 11q13.4-q14.3 in a boy with microcephaly, ptosis and developmental delay. *Eur J Med Genet* 53, 50-53, (2010).
- 90 Klep-de Pater, J. M., de France, H. F. & Bijlsma, J. B. Interstitial deletion of the long arm of chromosome 11. *J Med Genet* 22, 224-226, (1985).

- 91 Guc-Scekic, M., Pilic-Radivojevic, G., Mrdjenovic, G. *et al.* Interstitial deletion of 11q. *J Med Genet* 26, 205-206, (1989).
- 92 Knight, S. J., Regan, R., Nicod, A. *et al.* Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* 354, 1676-1681, (1999).
- 93 Wincent, J., Anderlid, B. M., Lagerberg, M. *et al.* High-resolution molecular karyotyping in patients with developmental delay and/or multiple congenital anomalies in a clinical setting. *Clin Genet* 79, 147-157, (2010).
- 94 Jakobsson, M., Scholz, S. W., Scheet, P. *et al.* Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* 451, 998-1003, (2008).
- 95 Friedman, J. I., Vrijenhoek, T., Markx, S. *et al.* CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Mol Psychiatry* 13, 261-266, (2008).
- 96 Poot, M., Beyer, V., Schwaab, I. *et al.* Disruption of CNTNAP2 and additional structural genome changes in a boy with speech delay and autism spectrum disorder. *Neurogenetics* 11, 81-89, (2010).
- 97 Ullmann, R., Turner, G., Kirchhoff, M. *et al.* Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* 28, 674-682, (2007).
- 98 Mefford, H. C., Cooper, G. M., Zerr, T. *et al.* A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* 19, 1579-1585, (2009).
- 99 Hannes, F. D., Sharp, A. J., Mefford, H. C. *et al.* Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *J Med Genet* 46, 223-232, (2009).
- 100 Melis, D., Genesio, R., Cozzolino, M. *et al.* An emerging phenotype of proximal 11q deletions. *Eur J Med Genet* 53, 340-343, (2010).
- 101 Meeths, M., Chiang, S. C., Wood, S. M. *et al.* Familial hemophagocytic lymphohistiocytosis type 3 (FHL3) caused by deep intronic mutation and inversion in UNC13D. *Blood* 118, 5783-5793, (2011).
- 102 Vermeesch, J. R., Balikova, I., Schrandt-Stumpel, C. *et al.* The causality of de novo copy number variants is overestimated. *Eur J Hum Genet* 19, 1112-1113, (2011).
- 103 Hagerman, R., Lauterborn, J., Au, J. *et al.* Fragile X syndrome and targeted treatment trials. *Results Probl Cell Differ* 54, 297-335, (2012).
- 104 Meikle, L., Pollizzi, K., Egnor, A. *et al.* Response of a neuronal model of tuberous sclerosis to mammalian target of rapamycin (mTOR) inhibitors: effects on mTORC1 and Akt signaling lead to improved survival and function. *J Neurosci* 28, 5422-5432, (2008).
- 105 Huang, H. S., Allen, J. A., Mabb, A. M. *et al.* Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. *Nature* 481, 185-189, (2012).