Realia et naturalia

DISSERTATIONES
BIOLOGICAE
UNIVERSITATIS
TARTUENSIS
264

OLGA ŽILINA

Chromosomal microarray analysis as diagnostic tool: Estonian experience

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

264

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
264

OLGA ŽILINA

Chromosomal microarray analysis as diagnostic tool: Estonian experience

UNIVERSITY OF TARTU
PRESS

Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

The dissertation is accepted for the commencement of the degree of Doctor of Philosophy (in gene technology) on 10.09.2014 by the Council of the Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

Supervisors: Prof. Ants Kurg, PhD
Institute of Molecular and Cell Biology,
University of Tartu, Tartu, Estonia

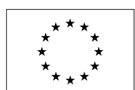
Prof. Katrin Õunap, MD, PhD
Department of Pediatrics, University of Tartu, Tartu, Estonia

Reviewer: Prof. Pärt Peterson, PhD
Department of General and Molecular Pathology, University of Tartu, Tartu, Estonia

Opponent: Hartmut Engels, PhD
Institute of Human Genetics, University of Bonn, Bonn, Germany

Commencement: Room No 105, Institute of Molecular and Cell Biology, University of Tartu; 23b Riia St, Tartu, on November 13th, 2014, at 10:00.

Publication of this dissertation is granted by the Institute of Molecular and Cell Biology, University of Tartu and by the Graduate School in Biomedicine and Biotechnology created under the auspices of European Social Fund.



ISSN 1024-6479
ISBN 978-9949-32-681-5 (print)
ISBN 978-9949-32-682-2 (pdf)

Copyright: Olga Žilina, 2014

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS 7
ABBREVIATIONS 8
INTRODUCTION 10
1. REVIEW OF THE LITERATURE 11
1.1. Chromosomal microarray analysis 11
1.1.1. Human cytogenetics: from microscopes to microarrays 11
1.1.2. Array-based techniques for analysis of DNA copy-number changes 15
1.1.2.1. Comparative genomic hybridization arrays 15
1.1.2.2. Single nucleotide polymorphism arrays 17
1.1.3. Application of chromosomal microarray analysis in research 18
1.1.4. Application of chromosomal microarray analysis in clinical practice 21
1.1.5. Use of chromosomal microarray analysis in prenatal diagnostics 26
1.2. Speech and language disorders associated with 7q31 genomic region and *FOXP2* gene 29
1.3. Reciprocal deletions and duplications of 5q35.2-q35.3 and conveyed phenotypes 31
1.4. Detection of mosaicism and uniparental disomy by single nucleotide polymorphism array analysis 33
1.4.1. Genetic mosaicism 34
1.4.2. Excessive homozygosity and uniparental disomy 36
2. AIMS OF THE STUDY 38
3. MATERIALS AND METHODS 39
3.1. Study subjects 39
3.1.1. Patients studied for chromosomal aberrations by chromosomal microarray analysis 39
3.1.2. Detailed clinical description of patients with 7q31 deletion from Publication II 40
3.1.2.1. Family 1 40
3.1.2.2. Family 2 42
3.1.3. Clinical description of a patient with 5q35.2-q35.3 duplication from Publication III 43
3.1.4. Clinical description of a patient with mosaic maternal UPD15 from Publication IV 43

3.2. Methods ... 44
3.2.1. Analysis of genomic aberrations by SNP-based chromosomal microarray analysis ... 44
3.2.2. CNV confirmation by qPCR ... 45
3.2.3. CNV breakpoints analysis by sequencing in Family 1 and 2 with 7q31 deletion ... 46
3.2.4. Gene expression analysis by RT-qPCR for 5q35.2-q35.3 region ... 47
3.2.5. Methylation-specific MLPA for 15q11 chromosomal region ... 50
4. RESULTS AND DISCUSSION ... 51
4.1. Application of chromosomal microarray analysis for the diagnosis of Estonian patients with developmental delay, multiple congenital anomalies and autism spectrum disorders (Publication I) ... 51
4.1.1. Clinically relevant CMA findings (Publication I) ... 52
4.1.1.1. Deletions in 7q31 encompassing *FOXP2* cause speech and language disorder (Publication II) ... 61
4.1.1.2. Reciprocal deletions and duplications in 5q35.2-q35.3 cause mirror phenotypes related to growth (Publication III and unpublished data) ... 63
4.1.1.3. Mosaic matUPD15 in a patient with partial Prader-Willi syndrome phenotype as an example of the utility of SNP-based arrays in detecting uniparental disomy and mosaicism (Publication IV) ... 65
4.1.2. Findings with uncertain clinical significance (Publication I) ... 68
4.1.3. Estonian experience of chromosomal microarray analysis application in prenatal diagnosis (Publication I) ... 70
5. SUMMARY AND CONCLUSIONS ... 75
REFERENCES ... 78
WEB RESOURCES ... 94
SUMMARY IN ESTONIAN ... 95
ACKNOWLEDGEMENTS ... 98
PUBLICATIONS ... 99
CURRICULUM VITAE ... 134

LIST OF ORIGINAL PUBLICATIONS

The current thesis is based on the following original publications, which will be referred to by the corresponding Roman numerals in the text:

I. **Žilina O.,** Teek R., Tammur P., Kuuse K., Yakoreva M., Vaidla E., Mölter-Väär T., Reimand T., Kurg A., Õunap K. (2014). Chromosomal microarray analysis as a first-tier clinical diagnostic test: Estonian experience. Mol Genet Genomic Med. 2(2): 166–175.

II. **Zilina O.,** Reimand T., Zjablovskaja P., Männik K., Männamaa M., Traat A., Puusepp-Benazzouz H., Kurg A., Ounap K. (2012). Maternally and paternally inherited deletion of 7q31 involving the *FOXP2* gene in two families. Am J Med Genet A. 158A(1):254–256.

III. **Žilina O.,** Reimand T., Tammur P., Tillmann V., Kurg A., Õunap K. (2013). Patient with dup(5)(q35.2-q35.3) reciprocal to the common Sotos syndrome deletion and review of the literature. Eur J Med Genet. 56(4):202–206.

IV. **Žilina O.,** Kahre T., Talvik I., Õiglane-Šlik E., Tillmann V., Õunap K. (2014). Mosaicism for maternal uniparental disomy 15 in a boy with some clinical features of Prader-Willi syndrome. Eur J Med Genet. 57(6):279–283.

The articles listed above have been reprinted with the permission of the copyright owners.

My personal contribution to the articles referred to in the thesis is as follows:

Ref. I contributed to introduction of chromosomal microarray analysis into clinical practice in Estonia; participated in data analysis and interpretation; wrote the manuscript

Ref. II contributed to experimental design; conducted the experiments; performed data analysis; wrote the manuscript

Ref. III contributed to experimental design; conducted the experiments; performed data analysis; wrote the manuscript

Ref. IV performed data analysis; prepared the manuscript

ABBREVIATIONS

ACMG	American College of Medical Genetics
ADHD	attention deficit hyperactivity disorder
aCGH	array-based comparative genomic hybridization
AS	Angelman syndrome
ASD	autism spectrum disorders
BAC	bacterial artificial chromosome
BAF	B-allele frequency
CAS	childhood apraxia of speech
cDNA	complementary DNA
CGH	comparative genomic hybridization
CMA	chromosomal microarray analysis
CNV	copy-number variation
DD	developmental delay
DGV	Database of Genomic Variants
DVD	developmental verbal dyspraxia
FISH	fluorescence in situ hybridization
G-banding	Giemsa-banding
heteroUPD	uniparental heterodisomy
HL	hearing loss
IBD	identical by descent
ID	intellectual disability
IQ	intelligence quotient
ISCA	International Standards for Cytogenomic Arrays
ISCN	International System for Human Cytogenetic Nomenclature
isoUPD	uniparental isodisomy
LCRs	low-copy repeats
LCSH	long contiguous stretches of homozygosity
LRR	$\log_2$ R intensity ratio
MCA	multiple congenital anomalies
MLPA	multiplex ligation-dependent probe amplification
MS-MLPA	methylation-specific multiplex ligation-dependent probe amplification
MS-PCR	methylation-specific polymerase chain reaction
NAHR	non-allelic homologous recombination

NGS	next generation sequencing
nt	nucleotide
p	short arm of human chromosome
PAC	P1-derived artificial chromosome
PCR	polymerase chain reaction
PWS	Prader-Willi syndrome
q	long arm of human chromosome
qPCR	quantitative polymerase chain reaction
RT-qPCR	real-time reverse transcription quantitative polymerase chain reaction
SNP	single nucleotide polymorphism
UPD	uniparental disomy
VUCS	variant of uncertain clinical significance
YAC	yeast artificial chromosome

INTRODUCTION

The widespread utilization of array-based comparative genomic hybridization (aCGH) and single nucleotide polymorphism (SNP) genotyping arrays has revealed that DNA copy-number variations (CNVs) constitute a significant portion of the observed genetic variation [1, 2] and could predispose to common and complex disorders, including developmental delay/intellectual disability (DD/ID), multiple congenital anomalies (MCA), and autism spectrum disorders (ASD) [3, 4]. These classes of disorders have a high rate of prevalence in the general population (DD/ID 2–3%; ASD ~1:150 individuals) and account for the largest proportion of cytogenetic testing [5]. Chromosomal microarray analysis (CMA), which encompasses all types of array-based genomic copy-number analyses, offers the capacity to examine the entire human genome on a single chip with a resolution that is at least 10-fold greater than the best G-banded chromosome analysis, and is now established as the first-tier cytogenetic diagnostic test for fast and accurate detection of chromosomal abnormalities in patients with DD/ID, ASD and/or MCA [5].

The decision to replace the traditional G-banding method with the novel CMA method was made by comparing the diagnostic yields of the two techniques and calculating the total cost of the analyses per patient. G-banded karyotyping alone detects pathogenic genomic imbalances in ~3% of patients with developmental disorders (excluding clinically recognizable chromosomal syndromes, e.g., Down syndrome), whereas the diagnostic yield for CMA is 10–25% depending on the microarray platform and patient selection [4–6]. In general, the adoption of microarrays by research and clinical diagnostic laboratories to analyze DNA copy-number changes has had a great impact on the field of medical genetics by enabling to both clarify genotype-phenotype relationships in known disorders and identify novel syndromes [3, 7].

In Estonia, CMA was firstly applied in a high-resolution study of CNVs and uniparental disomy (UPD) in Estonian families with unexplained ID carried out in 2007–2009 [8]. The relevant structural aberrations were detected in 23% of the families analyzed, thus further demonstrating both the great value of CMA for the detection of genetic causes of ID and its role as a reliable and effective tool in research and diagnostics. In another study, CMA was applied in a small cohort of patients with hearing loss (HL) with a conclusion that this diagnostic method might be recommended for patients where HL occurs together with DD or congenital anomalies [9]. The successful application of CMA in research studies prompted the introduction of this state-of-the-art diagnostic possibility into clinical practice in 2009. At that time, the target population consisted of patients whose diagnosis remained unknown despite all routine genetic investigations. Since 2011, CMA is on the official service list of the Estonian Health Insurance Fund and is performed as the first-tier cytogenetic diagnostic test for patients with DD/ID, ASD, and/or MCA. This study summarizes the Estonian experience of using CMA for postnatal and prenatal diagnosis in routine clinical practice; both positive aspects and major drawbacks and challenges are discussed.

1. REVIEW OF THE LITERATURE

1.1. Chromosomal microarray analysis

1.1.1. Human cytogenetics: from microscopes to microarrays

The need to detect genomic copy-number changes originates from the knowledge that both chromosomal numerical and structural aberrations work to cause abnormal mental and physical development in humans [10–14]. In addition, recent advances in the field of DNA microarray technologies have demonstrated the importance of submicroscopic copy-number variations (CNVs) in both human evolution and genetic and hence phenotypic diversity between individuals [1, 2, 15–17]. Moreover, CNVs can also represent benign polymorphic changes that have no effect on human phenotype [15, 18]. The very first large-scale studies of CNVs in human population have proposed that up to 12% of the human genome could be involved in copy-number variation [15, 17]. However, in these initial studies, the CNV size was often overestimated because of the relative low resolution of array platforms used and the actual percentage of the genome affected by CNVs could be in the range of 0.78–1.28% [2, 21]. Deletions, amplifications, insertions, and translocations can all result in CNV that is defined as a segment of DNA that is 1 kb or larger in size and is present in a variable copy-number when compared with a reference human genome [19].

The ability to study chromosomal rearrangements, including CNVs, has changed dramatically over time. The field of cytogenetics – the study of the structure, function, and evolution of chromosomes – sprang up at the end of the 19th century, however, the period up to the early 1950s is known as "the dark ages". The first cytogenetic studies were difficult to conduct, and it was only in 1956 that the correct number of chromosomes in normal human cells was finally discovered and clinical cytogenetic practice was born [22, 23]. In the following years, several numerical and structural chromosomal abnormalities have been associated with specific disorders, for example, trisomy 21 in Down syndrome, 45,X in Turner syndrome, 47,XXY in Klinefelter syndrome, or "Philadelphia chromosome" in chronic myeloid leukemia [14, 24–26]. In 1963, the first inherited deletion syndrome, Cri du Chat, was recognized: it was noted that all the patients who displayed specific phenotypic features, including severe ID and a characteristic cat-like cry, were missing a portion of the short arm of chromosome 5 [27]. The rapid development of human cytogenetics in the ensuing years has become possible due to technological advances that have combined innovations in molecular biology, chemistry, and instrumentation (reviewed in [28]). In 1968, staining protocols were developed that generated reproducible chromosome-specific patterns of light and dark bands along the length of each chromosomes, thus enabling one to easily recognize chromosomes, detect deletions, inversions, insertions, translocations, fragile sites, and other more complex rearrangements, and refine breakpoints [29].

Although several different banding techniques have been developed, the Giemsa-banding method (G-banding), based on the application of trypsin followed by Giemsa staining, still remains the most widely used method in clinical cytogenetic laboratories [30, 31]. Typically, 400–800 bands per haploid genome, according to the International System for Human Cytogenetic Nomenclature (ISCN), are visible in prometaphase chromosomes. This means that, depending on the banding resolution and characteristics of a chromosomal region, CNVs smaller than 5–10 Mb cannot be detected. Nevertheless, the introduction of G-banding was a big step forward for clinical cytogenetics, and provided an opportunity to study not only patients but also healthy individuals as possible carriers of a balanced aberration, e.g., family members of a known carrier or couples suffering from recurrent spontaneous abortions [31].

Historically, visible structural aberrations could not be associated with the sequence. This challenge has been overcome with the development of a completely new approach: fluorescence in situ hybridization (FISH), which has given a birth to the field of molecular cytogenetics and has worked to narrow the gap between cytogenetics and molecular biology [32, 33]. FISH was developed as a refinement of radioactive labeling, where the isotopic labels were replaced with fluorescent ones, thus allowing FISH to be routinely applied under clinical laboratory settings. FISH allows one to reveal the presence and localization of specific labeled DNA probes that have bound to complementary sequences on their targets, traditionally the metaphase chromosome spreads. This technique was initially used to diagnose known deletion syndromes such as DiGeorge and Williams-Beuren syndromes. Currently, segments as small as 10 kb or even 1 kb may be localized, depending on the probes (e.g., cosmids, BACs, PACs, YACs, fosmids, oligonucleotide probes) and the substrates (e.g., metaphase and interphase chromosomes, extended chromatin fibers) used for hybridization. Further developments of FISH include multiplex-FISH (M-FISH) and spectral karyotyping (SKY) where all chromosomes are differentially colored in a single experiment. These techniques enable one to study very complex rearrangements that are typically associated with different types of solid tumors [34, 35]. Despite their usefulness, FISH studies are very time-consuming, difficult to automate, and can therefore only be applied for a limited number of probes [31].

In addition to FISH, alternative targeted approaches have been developed for CNV analysis, such as real-time quantitative polymerase chain reaction (qPCR), and multiplex ligation-dependent probe amplification (MLPA) [36, 37]. All targeted methods require prior knowledge of the region under study and are used in clinical genetics tests to detect recurrent events or to confirm the results of other studies or tests.

The next transformation of cytogenetics came with the introduction of a completely new approach – comparative genomic hybridization (CGH) and its further development – array-based CGH (aCGH) that both permit the detection of chromosomal copy-number changes without the need to culture cells [38–40].

The principle of CGH is to compare two differentially labeled (usually red and green) genomic DNA samples from two sources (e.g. tumor and normal tissue) hybridized to a normal human metaphase preparation. The ratio of red-to-green fluorescence is measured across the length of each chromosome and gains or losses can be therefore delineated. Compared to traditional molecular cytogenetic techniques, CGH does not need any prior knowledge of the region under investigation and can be used to quickly scan an entire genome for imbalances. However, the resolution of CGH is limited by the use of chromosomes as targets and achieves approximately 5–10 Mb, which is comparable with conventional G-banding. In aCGH, the principles of traditional CGH are combined with the use of microarrays and the metaphase chromosomes are replaced by an array of small segments of DNA (e.g., BAC or PAC clones, oligonucleotide probes). The resolution of this analysis is determined by the array probe size and the genomic distance between them.

Another microarray type that has been widely adopted in cytogenetic studies is SNP arrays. Developed initially for genotyping procedures, they have progressively started to be used to simultaneously genotype SNPs and detect CNVs [41, 42]. In contrast with aCGH that requires the simultaneous hybridization of reference DNA from a healthy person or a normal tissue and test DNA, in case of SNP arrays, only test DNA is needed. In addition to CNVs, SNP arrays can also detect copy-number neutral events, mosaicism, and determine the parental origin of *de novo* CNVs in trios [43].

The newest approach is the use of next generation sequencing (NGS), or massively paralleled sequencing, data for CNV analysis. Most NGS applications focus on the detection of single base variants and small insertions/deletions (indels), but indeed structural variations (>50 bp), including larger indels, CNVs and translocations, can be identified from the same data. NGS-based methods are developing rapidly and are becoming part of routine clinical genomic testing where they are often able to supplement array-based techniques and may potentially replace them in the future [44]. While CNV analysis could be relatively easily fulfilled using data obtained from whole-genome sequencing, none of the whole-exome sequencing-based CNV detection tools currently performs at a satisfactory level and all suffer from a limited power [45].

An overview of both conventional and molecular cytogenetic methods that are routinely used to detect chromosomal rearrangements is given in **Table 1** [46].

Table 1. Overview of conventional and molecular cytogenetic methods that are routinely used to detect chromosomal rearrangements (adapted from [46]).

Technique	Aberration detection			Max resolution	Sensitivity	Number of loci*
	Unbalanced	Balanced	LCSH and UPD			
Whole-genome wide chromosome-based						
G-banding	x	x	–	Low (>5Mb)	Low	High
CGH	x	–	–	Low (>3–5Mb)	High	High
SKY/M-FISH	x	x	–	Low (>several Mb)	High	High
Whole-genome wide array-based						
1-Mb BAC aCGH	x	–	–	Medium (>1Mb)	High	High
Tiling-path BAC aCGH	x	–	–	High (>50kb)	High	High
Oligonucleotide aCGH	x	–	–	High (up to kb)	High	Very high
SNP genotyping arrays	x	–	x	High (>5kb)	High	Very high
Targeted approaches						
FISH	x	x	–	High (<100kb)	High	Low
qPCR	x	–	–	High (~100bp)	Very high	Low
MLPA	x	–	–	High (~100bp)	High	Medium
Next-generation sequencing-based						
NGS	x	x	x	Very high (bp)	Very high	Very high

*number of genomic loci analyzed per experiment.
Abbreviations: LCSH – long contiguous stretches of homozygosity; UPD – uniparental disomy; CGH – comparative genomic hybridization; SKY – spectral karyotyping; M-FISH – multiplex FISH; FISH – fluorescence *in situ* hybridization; BAC – bacterial artificial chromosome; aCGH – array-based comparative genomic hybridization; SNP – single nucleotide polymorphism; qPCR – real-time quantitative polymerase chain reaction; MLPA – multiplex ligation-dependent probe amplification; NGS – next-generation sequencing.

1.1.2. Array-based techniques for analysis of DNA copy-number changes

Originally developed for gene expression profiling, DNA microarrays currently have a wide range of applications in molecular biology: CNV analysis, SNP genotyping, DNA methylation, alternative splicing, miRNA, and protein-DNA interactions studies [46]. Each array consists of thousands of small DNA sequences (e.g., cloned sequences or oligonucleotide probes) immobilized on a solid surface. The core principle behind microarrays is hybridization between two complementary nucleic acid strands: "probes" attached to the array surface and labeled DNA or RNA "targets" applied to the array. The main advantages of array-based methods are the ability to rapidly scan a huge amount of genetic targets in a single experiment, high sensitivity, and high specificity. In terms of clinical settings, the low amount of input sample (<1µg) material is also advantageous.

Chromosomal microarray analysis (CMA) is a general term used to refer to microarray-based genomic copy-number analysis, which can also be offered under other names, e.g., "molecular karyotyping". Array-based approaches enable one to analyze chromosomes for chromosomal rearrangements that result in changes in DNA copy-number, such as chromosomal loss or gain. Thus, CMA performs the function of conventional karyotyping but at a much higher resolution [5].

Two types of microarrays used for CNV analysis are available: CGH arrays and SNP arrays. Various array platforms for specific applications are produced by array manufactures, and usually some extent of flexibility in design is provided which allows one to adapt the content of the array to specific needs and purposes.

1.1.2.1. Comparative genomic hybridization arrays

The aCGH technique was introduced in 1997 as a further development of conventional CGH with the purpose of improving its resolution [39, 40]. The first array probes represented large genomic clones of typically 80–200 kb in length (BACs or PACs) robotically spotted onto a solid surface. The first diagnostic CGH arrays consisted of roughly 3000 BACs covering the entire genome with a mean resolution of approximately 1 Mb and were used to test patients with ID or MCA [47–49]. In 2004, the first whole-genome tiling path array was developed that contained 32,000 overlapping BAC clones and provided a mean resolution of roughly 100–150 kb [50]. The BAC array resolution limit corresponds to the average size of a BAC clone. BAC arrays are unable to accurately define a CNV that is smaller than the BAC clone which defines it. In addition, BAC arrays tend to overestimate the real size of CNVs detected due to the inaccurate determination of aberration breakpoints [17]. The use of shorter sequences and increasing the number of probes present on an array enables one to further improve the resolution. BAC clones were replaced

by complementary DNA (cDNA) probes, followed by PCR amplicons and finally, oligonucleotide probes that are currently widely used [51–54]. Although the short size of oligonucleotide probes (roughly 25–85 nt) assures a high resolution of CNV analysis, which now depends mainly on the number of probes and genomic distance between them not on their size, it can also be seen as a drawback because short sequences make the hybridization process variable and sequence-dependent [55]. To overcome this limitation and to reduce the number of false-positive results, the average signal intensity of several consecutive probes is taken into account during CNV calling procedures. This, however, reduces the effective resolution of an array platform. This effect is in-turn compensated by the great number of probes which can achieve up to 1 million probes per slide and allows the detection of chromosomal imbalances as small as just a few kilobases and more precise determination of the aberration breakpoints. Actually, the application of arrays with even higher resolution has been described [2].

The detection of CNVs using aCGH typically proceeds as follows: both test and reference DNA samples are labeled with different fluorophores (usually Cy3 and Cy5), and then co-hybridized onto an array in the presence of Cot-1 DNA to reduce the binding of repetitive sequences. The signal ratio between the test and reference sample is normalized and converted to a $\log_2$ ratio, which provides a means for copy-number estimation. An increased $\log_2$ ratio (theoretically, 0.58, as $\log_2 (3/2)=0.58$) indicates a one copy gain in test DNA compared with the reference, while a decreased $\log_2$ ratio (theoretically, –1, as $\log_2 (1/1)= -1$) indicates a one copy loss. A $\log_2$ ratio value of zero indicates a normal copy-number [40, 46]. Different algorithms can be used to normalize the array data to correct for artifacts such as GC-bias or dye-bias and to guarantee better CNV detection and reduce the number of false-positive results [46].

One advantage of oligonucleotide CGH arrays is the flexibility to create custom designs, even for a limited number of slides ordered. Specific regions and genes of interest can be covered with many probes in order to increase the local resolution and detect the tiniest pathological imbalances, down to one deleted exon [56, 57]. Usually, different options for custom design are provided: the user can choose oligonucleotide array probes from a database containing *in silico*-validated sequences or alternatively add any custom oligonucleotide sequences to an array. All procedures can be accomplished through online applications such as eArray (Agilent Technologies (Santa Clara, CA, USA)).

The main disadvantage of aCGH that does not depend on the choice of array platform is an inability to detect the presence of mosaicism, balanced aberrations (e.g., inversions, and balanced translocations), and copy-number neutral events such as long contiguous stretches of homozygosity (LCSH) and UPD. Also, these arrays are not helpful in determining the parental origin of *de novo* CNVs. To compensate these drawbacks, at least partially, some manufactures (e.g., Agilent Technologies) have attempted to develop CGH arrays containing additional SNP probes that allow one to screen for LCSH [58].

However, these do not provide robust SNP genotypes and thus cannot always delineate the parental origin of a *de novo* aberration.

1.1.2.2. Single nucleotide polymorphism arrays

SNP arrays were initially developed to conduct genotyping procedures. Single nucleotide polymorphisms – which represent a substantial part of human genetic variation – are genome positions at which two distinct nucleotide residues are possible and both appear in a significant portion of a human population. For simplicity, manufactures usually label the two alleles of a SNP as *A* and *B*, which means that an individual can have an *AA*, *AB,* or *BB* genotype at any particular SNP site [59]. According to the SNP database (http://www.ncbi.nlm.nih.gov/snp; dbSNP BUILD 140) there are more than 55 million SNPs in the human genome. The very first SNP arrays designed by Affymetrix, Inc. (Santa Clara, CA, USA) – the pioneer of commercial SNP arrays – genotyped 1494 SNPs [59]. Since then, SNP arrays have improved and currently several millions SNPs can be genotyped on a single chip. Large-scale studies such as the HapMap Project and the 1000 genomes project have provided an opportunity to improve the content of SNP arrays by incorporating more informative SNPs [60, 61]. Because SNP array probes represent short oligonucleotide sequences, the resolution of SNP arrays is similar to that of CGH arrays and depends both on the number of array probes and the genomic distance between them [57].

The two main manufactures of SNP arrays – Affymetrix, Inc. and Illumina, Inc. (San Diego, CA, USA) – use different chemistries in their genotyping protocols, however, they both rely on the same principle of complementary base pairing, which is also utilized by aCGH [59]. The copy-number analysis provided by SNP arrays relies on two parameters, the $\log_2$ R intensity ratio (LRR) and the "B-allele frequency" (BAF). The former is obtained by comparing the signal intensities from the test sample to a collection of reference hybridizations. The latter, in-turn, is calculated for each particular SNP from the observed allelic intensity ratio by interpolating the known B allele frequencies of the three canonical clusters (0, 0.5, and 1.0). The BAF has a significantly higher per-probe signal-to-noise ratio than LRR data and can be interpreted as follows: a BAF of 0 represents the genotype (*A/A* or *A*/–), whereas 0.5 represents (*A/B*) and 1 represents (*B/B* or *B*/–). Generally, the BAF may be used to accurately assign copy-numbers from 0 to 4 in diploid regions of the genome [62, 63].

Therefore, the main difference between CGH arrays and SNP arrays is that the latter provides genotype information in addition to copy numbers. While both CGH and SNP arrays can easily detect chromosomal aberrations associated with DNA copy-number changes, the additional genotype information provided by SNP arrays allows one to detect mosaicism, LCSH, UPD, and determine the parental origin of *de novo* aberrations [43]. Knowledge regarding copy-number

neutral events and the parental origin of *de novo* CNVs is often essential in terms of imprinting disorders. In addition, a substantial portion of a genome presented in a homozygous state may indicate a possible consanguinity between parents which may be useful in genetic counseling [64]. In contrast with aCGH, SNP arrays do not require simultaneous hybridization of test and reference samples because the fluorescent signal intensity of each probe is compared with a reference bioinformatic file [63, 65]. These reference files could either be provided by the manufactures for the general population, or custom population-specific files created by the users.

Despite the advantages provided by SNP arrays, some studies have found that dedicated CGH arrays tend to have better sensitivity when analyzing copy-numbers [46, 66]. However, no difference in CNV calls have been detected in other studies, so both array platforms can be considered equivalent and can be used in diagnostics for patients with the same clinical indications [67]. Nevertheless, to improve the efficiency of CNV discovery using SNP arrays and to increase marker density in CNV regions, manufactures now provide array designs that include non-polymorphic copy-number probes (which are used to examine log ratios but not BAF). For example, the Affymetrix Genome-Wide Human SNP Array 6.0 contains 1.8 million genetic markers, half of which are probes for the detection of CNVs. In addition, as with most CGH arrays, the user can order custom SNP arrays or add additional probes to standard chips to target specific regions of the genome. For instance, up to 50,000 markers can be added to an Illumina HumanOmniExpress array that initially contains ~700,000 fixed probes.

As with CGH arrays, one of the main disadvantages of SNP arrays is their inability to detect balanced rearrangements such as inversions and balanced translocations. Under clinical diagnostic settings, conventional cytogenetic methods are still commonly used to screen specific patient populations (e.g., couples suffering from recurrent spontaneous abortions) for this type of chromosomal alteration. In addition, both, CGH arrays and SNP arrays provide no information regarding the location of duplicated copies and are generally unable to resolve breakpoints at the single-base-pair level. Furthermore, microarrays are less sensitive when detecting single-copy gains compared with deletions, especially when analyzing shorter CNVs [62].

1.1.3. Application of chromosomal microarray analysis in research

Although structural variants in some genomic regions do not have obvious phenotypic consequences, others may cause genetic diseases, either alone or in combination with other genetic or environmental factors [15, 16, 68]. Disorders in which the clinical phenotype is a consequence of either abnormal dosage or dysregulation of one or more genes located within a rearranged segment of the genome are known as “genomic disorders” [20]. During the pre-array era, it was known that changes in both the number of chromosomes and their structure

cause clinical phenotypes, however, the introduction of array technologies, including whole-genome wide array-based methods, enabled high-resolution screening of entire genomes and led to the discovery of numerous submicroscopic changes in the chromosome structure that had not been previously observed. It has also become evident that despite the smaller size of submicroscopic aberrations, their overall potential contribution to human genetic variation and disease may be higher compared with microscopic changes because they occur at a higher frequency [19].

Because tumor samples have a high frequency of various chromosomal rearrangements, the primary application of both CGH and aCGH was mainly in cancer research. These technologies improved cancer diagnosis, prognosis, and classification of tumors, and also helped to identify candidate cancer genes [69, 70]. In addition, it has been found that submicroscopic subtelomeric chromosome alterations could be responsible for up to 5% of previously unexplained cases of ID, which affects 1–3% of the general population in developed countries and represents a leading socio-economic problem in health care [71–73]. The most frequent cause of ID is Down syndrome (or trisomy 21) with a frequency of ~12% in the ID population. The overall livebirth prevalence of Down Syndrome in Estonia – 1.17 per 1000 livebirths – is within the range reported in the literature and thus allows one to suppose that the frequency of this aneuploidy among Estonian ID patients may be approximately the same, although it has not been directly estimated [74]. Other numerical and chromosomal imbalances that are visible with light microscopy are less common and account for between 5–15% in the ID population [72, 75]. In addition, hundreds of monogenic forms of ID have been mapped (http://www.ncbi.nlm.nih.gov/omim). Despite these advances, the etiology in more than half of all ID cases remains unknown. Taking into account that many ID-associated aberrations found in subtelomeric regions were interstitial, it has been proposed that submicroscopic pathogenic alterations should exist elsewhere in the human genome. Vissers et al. demonstrated the utility of aCGH in detection submicroscopic aberrations with high sensitivity and specificity [48]. Since then, aCGH has become a method of choice for detecting deletions or duplications associated with neurodevelopmental disorders. Depending on the array platform used and the patient selection criteria, apparently pathogenic CNVs can be identified in up to 15% of patients with idiopathic ID and normal karyotype [76, 77]. A large comprehensive study of unexplained ID has also been carried out in Estonian patients using an SNP genotyping platform. This revealed the underlying reason of ID in 23% of the patients and allowed for the identification of 4 novel aberrations associated with ID. Genotype data from a cohort of 1000 individuals from the general population have been used to accurately interpret CNVs detected in ID patients [8]. Cooper et al. composed a copy-number variation morbidity map of developmental delay through the analysis of array data from more than 15,000 patients with ID and various congenital defects and comparing it with >8,000 healthy controls. This study

offers a great resource for researchers and clinical geneticists studying the etiology of ID/DD and ASD [78]. With this growing body of knowledge, ID could be considered as a model disease to study the clinical consequences of CNVs.

In addition, CNV analysis has been used to study multiple psychiatric disorders, such as autism, schizophrenia, bipolar disorder and others, because these disorders display a high degree of heritability which suggests a genetic component to susceptibility. Indeed, a strong association between autism and *de novo* CNVs has been demonstrated by different research groups who found that the frequency of *de novo* CNVs can reach up to 10% in patients with ASD, and simplex autism shows a stronger correlation than multiplex autism [79–81]. As in case of ID, the presence of additional clinical problems, such as dysmorphic features and/or ID, increases the chances of detecting an underlying genetic cause. Application of array-based technologies to study ASD has allowed for the detection of multiple autism susceptibility loci in the human genome, including 16p11.2, 15q13.3, and 7q11.23. However, the presence of numerous genes in these regions can hamper the identification of the genes that actually cause this disorder. Depending on both the function and the degree of dosage sensitivity of the affected genes, common autism-associated CNVs can demonstrate wide phenotypic expressivity. Those with higher penetrance for ASD (e.g., 15q11-q13 duplication) are usually of *de novo* origin, cause more severe symptoms, and are more prevalent among sporadic ASD. Some CNVs (e.g., 15q11.2 deletion) produce moderate or mild effects and are also found in unaffected family members and the general population, which suggests the involvement of modifying factors of genetic or non-genetic origin [82].

CNV studies with schizophrenic patients have confirmed that submicroscopic chromosomal rearrangements also play a role in this neuropsychiatric disorder [83, 84]. However, different studies display poor overlap and currently only a small number of specific CNV loci are enriched in schizophrenia cases, including 1q21.2, 15q13.3, 16p11.2, and 22q11.2. Interestingly, these genomic regions have been also associated with other neurodevelopmental and neuropsychiatric disorders. Recently, a meta-analysis of multiple CNV studies of schizophrenia was performed with the purpose of generating a map of genes affected by CNVs in schizophrenia, which will serve as a starting point for further functional studies leading to the discovery of targets for future therapeutics and diagnostics [85].

Moreover, rare CNVs with a strong phenotypic effect may contribute to common diseases and complex traits, and may help to explain the “missing heritability“ (the portion that cannot be explained by defects in individual genes). One such example is a 16p11.2 microdeletion associated with a highly-penetrant form of obesity. This was initially found in a relatively small cohort of patients that display this extreme phenotype, while further procedures included targeted follow-up studies in European population cohorts [86].

In general, the adoption of microarrays for the analysis of DNA copy-number changes by research and clinical diagnostic laboratories had a great impact on the field of medical genetics, by enabling both the clarification of genotype-phenotype relationships in known disorders and identification of novel syndromes [3, 7]. Genomic microarrays ease the process of obtaining genotype information and have opened up a new approach to identify novel syndromes, termed the “genotype-first” approach or “reverse phenotyping” [87]. The main idea is that patients with overlapping genotypes are collected together and then followed by accurate clinical examination and observation. This contrasts with the traditional approach of first starting with a collection of a phenotypically homogeneous group of patients. The genotype-first approach has proven to be very effective and has lead to the discovery of many microdeletion/ microduplication syndromes, some of which display a great degree of phenotypic variability and penetrance and could have remained undetected based purely on the phenotype assessment method [88–90]. In addition, array-based methods may also facilitate the identification of disease genes that cause known syndromes. This may be achieved by narrowing the critical region of a specific disease by detecting patients with a specific phenotype and overlapping CNVs. One example is the identification of the genetic cause of CHARGE syndrome, which appears to be caused by haploinsufficiency of the *CHD7* gene, either by microdeletions encompassing *CHD7*, or by mutations within the gene [91].

In addition, array-based methods have been used to study the global genetic variation in the general population. The first evidence that changes in DNA copy-number do not necessarily result in clinical phenotypes but are present in abundance in healthy individuals came in 2004 [15, 16]. In 2006, the first CNV map of the human genome was completed; however, due to imperfections in the array-platform used the size of detected CNVs was generally overestimated. Later, the CNV map was refined using specifically developed SNP-CNV genotyping arrays, and from this it became obvious that common copy-number polymorphisms tend to show patterns of allele frequency, linkage disequilibrium, and population differentiation just like SNPs [92]. A comprehensive summary of structural variation in the human genome can be found in the Database of Genomic Variants (DGV; http://dgv.tcag.ca/) which contains data obtained from the analysis of healthy control samples and is continuously updated with new data from peer reviewed research studies [93].

1.1.4. Application of chromosomal microarray analysis in clinical practice

Although, array-based techniques are used extensively in research to study CNVs associated with various clinical phenotypes, under clinical diagnostic settings conventional cytogenetic methods generally remained the methods of choice. Nevertheless, the both the usefulness of microarrays in detecting

pathogenic CNVs and technical advances in the field of microarrays and CNV calling algorithms have made genomic microarrays an attractive tool for clinical diagnostics. In 2010, a consensus statement was published that recommend using CMA as a first-tier diagnostic test for individuals with DD or MCA, whereas G-banding should be reserved for patients with obvious chromosomal syndromes or a family history of multiple miscarriages. The authors, which represent the International Standards for Cytogenomic Arrays (ISCA) Consortium, performed a meta-analysis of 33 original studies, including 21,698 patients tested by array-based methods, and concluded that CMA offers a much higher diagnostic yield (15–20%) for patients with unexplained ID/DD, ASD, or MCA than G-banding (~3%, excluding Down syndrome and other recognizable chromosomal syndromes). Although one can object that CMA is unable to detect balanced translocations, this type of chromosomal aberrations is not a frequent cause of ID and can be found in only about 0.3% of ID patients tested by karyotype analysis [5].

In contrast to its application in research, diagnostic microarrays ought not to contain the maximum number of probes nor should they be able to detect tiny CNVs because this would generate too much data that hampers the efficient assessment of the clinical relevance. According to the recommendations of the American College of Medical Genetics (ACMG), clinical copy-number microarrays should provide the minimum detection size of 400 kb throughout the genome with increased resolution in the regions of clinical relevance (e.g., dosage-sensitive genes and subtelomeric regions) [94]. A clinically effective CMA resolution should therefore represent a balance between sensitivity and specificity where the sensitivity is considered in relation to standard karyotyping and not in comparison to gene-specific molecular-based assays for Mendelian disorders [57]. Specific standardized microarrays for clinical applications are now offered by microarray manufactures with the content that meets diagnostic standards and needs, e.g. HumanCytoSNP-12 BeadChip by Illumina, Inc. and CytoSure ISCA arrays by Oxford Gene Technology (Oxfordshire, UK).

Recommendations for the clinical interpretation of CMA results are presented in **Table 2**. Even with these recommendations, assessment is not always straightforward and remains challenging in some cases because exceptions for some criteria have been observed. For example, many pathogenic CNVs are rare and nonrecurrent events which may mean that there is insufficient information in the literature and databases which complicate the assessment. Also, rare inherited CNVs pose a challenge because these may still be clinically relevant through, for example, incomplete penetrance, variable expressivity (e.g., 15q13.3, 16p11.2, and 22q11.21), or the presence of a second mutation not detectable by CMA. Thus, it is often imprudent to attribute clinical significance based on the inheritance pattern of a single CNV in a single family. In addition, CNVs in noncoding regions that have generally been considered

Table 2. Recommendations for assessing the clinical relevance of a CNV identified in patients with DD or congenital anomalies. Certain criteria define whether a CNV could be pathogenic, benign, or with uncertain clinical significance. Adapted from [5].

Primary criteria		**Pathogenic**	**Benign**
1. Inherited CNVs	Identical CNV inherited from a healthy parent		x[a]
	Expanded or altered CNV inherited from a parent	x	
	Identical CNV inherited from an affected parent	x	
2. Familial CNVs	Identical CNV inherited from a healthy relative		x[a]
	Identical CNV inherited from an affected relative	x	
3. CNVs present in databases or published	CNV has been reported in multiple publications or databases as a benign variant present in healthy individuals		x
	CNV overlaps with an imbalance defined with a high resolution technology in CNV database for patients with ID/DD, ASD, or MCA	x	
	CNV overlaps with well-recognized deletion or duplication syndrome region	x	
4. Gene content	CNV contains morbid OMIM gene(s)	x[b]	
	CNV is gene rich	x	
	CNV is gene poor		x
Minor criteria[c]		**Pathogenic**	**Benign**
1. Type of CNV	CNV is a heterozygous/homozygous deletion	x	
	CNV is a duplication containing no known dosage-sensitive genes		x
	CNV is an amplification	x	
2.	CNV is devoid of known regulatory elements		x

Abbreviations: UCS – uncertain clinical significance; DGV – database of genomic variants; ID – intellectual disability; DD – developmental delay; ASD – autism spectrum disorders; MCA – multiple congenital anomalies.
[a] with some exceptions; [b] CNV should produce the same type of mutation that is known to cause OMIM disease and the phenotype should be that expected for the OMIM disease; [c] exceptions to each case have been observed.

nonpathogenic because of the underlying disease mechanism appeared less evident, however, may also produce a clinical phenotype by affecting the regulation of one or more distant target genes [95, 96]. Generally, when assessing the clinical significance of a single CNV, one should keep in mind the mechanisms by which structural rearrangements convey clinical phenotypes (**Figure 1**). The ACMG recommends assigning any CNV reported in a patient to one of three main categories of significance: pathogenic, benign, or a variant of uncertain clinical significance (VUCS). The latter category should include all CNVs for which the clinical significance could not be properly determined due to insufficient knowledge at the time of reporting [97]. Later, these CNVs could be re-evaluated and reclassified as either pathogenic or benign. The application of this strategy was recently demonstrated when CMA results of the same cohort of patients were reanalyzed (first in 2010 and again in 2012) and found to display a statistically significant difference in the interpretation of CNVs (the percentage of CNVs assigned with a label "potentially pathogenic" increased from 19% in 2010 to 31% in 2012). The reasons for this difference include, for example, newly acquired knowledge regarding pathogenic and benign CNVs, both in the literature and databases, detection of novel candidate genes for various diseases, and remapping of gene locations in new human genome assemblies [98]. It should be emphasized that although CNVs with uncertain clinical significance could possibly aid in future diagnosis, at time of reporting the presence of VUCS, due to the uncertainty of their nature, often cause stress for both parents and patients.

Several countries, including Belgium and the Netherlands, have already adopted CMA as a first-tier cytogenetic test for patients with ID/DD, ASD, or MCA as recommended by the ISCA Consortium. The diagnostic yield varies between studies from 5% to 25% depending on the microarray platform used and applied patient selection criteria [5, 6, 100, 101]. Nevertheless, the main conclusion from these studies is the recognition that CMA has a high utility in clinical diagnostics. Recently, CMA was proposed as a first-line investigation in patients with neurological disorders where the initial clinical assessment did not indicate a likely etiology, especially in those with severe epilepsy and neurologically abnormal neonates. It was demonstrated that CMA could recognize an underlying genetic explanation in up to 10% of patients [102].

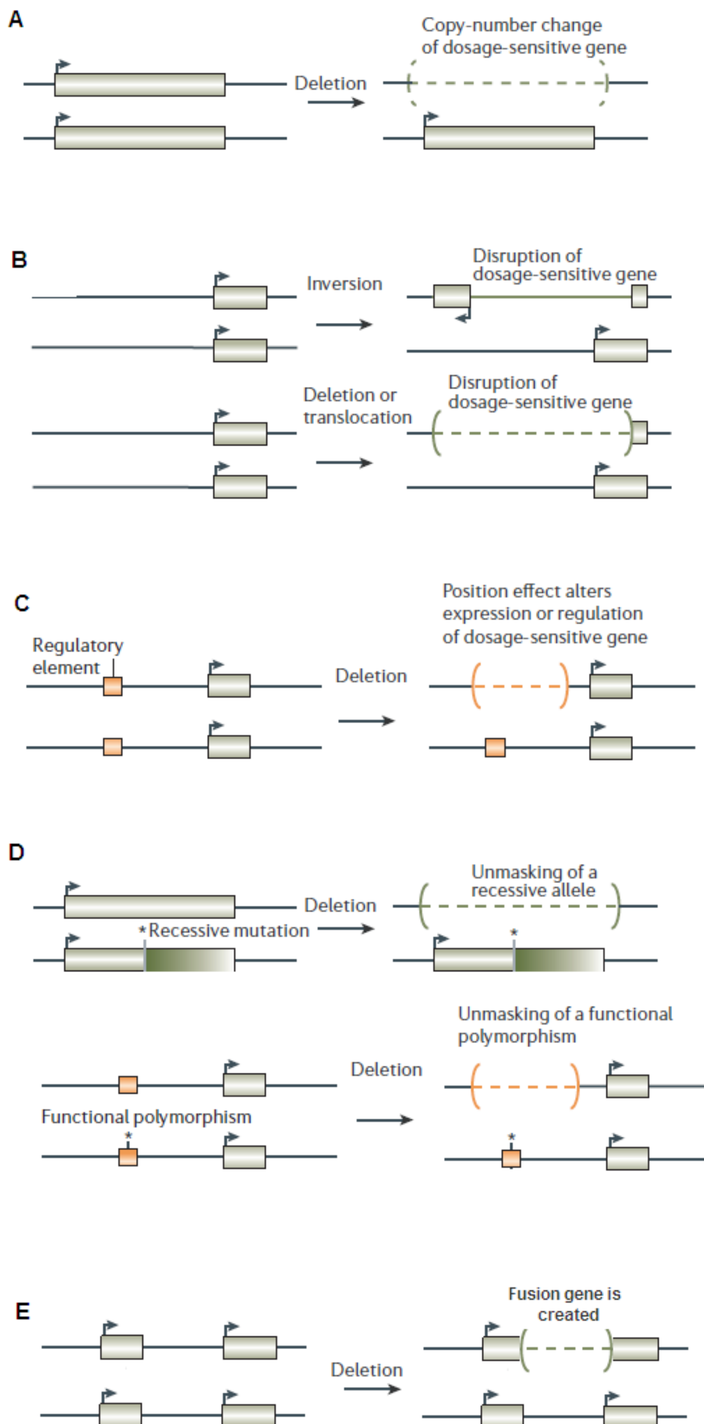


Figure 1. **Overview of the molecular mechanisms by which CNVs convey phenotypes**. Five models are depicted and include
A. gene dosage, where there is a dosage-sensitive gene within the rearrangement (a deletion is shown here);
B. gene interruption, wherein a dosage-sensitive gene is disrupted by inversion (upper panel), or translocation or deletion (lower panel);
C. position effect, in which the rearrangement has an effect on the expression/ regulation of a gene near the breakpoint, potentially by removing or altering a regulatory sequence;
D. unmasking recessive allele or functional polymorphism, where the deletion results in hemizygous expression of a recessive mutation (upper panel) or further uncovers effects of a functional polymorphism (lower panel);
E. a fusion event whereby a fusion gene is created at the breakpoint that either fuses coding sequences or a novel regulatory sequence to the gene. In each model, both chromosome homologs are depicted as horizontal lines. The rearranged genomic interval is enclosed by brackets. Each gene is depicted by a grey horizontal rectangle while regulatory elements are shown as orange rectangles. An asterisk denotes a point mutation. Adapted from [19, 99].

It is not surprising, however, that generating a large amount of CMA data could aid in the discovery of new knowledge that was not initially intended to be found and is unrelated to the patients' referring diagnoses. This additional knowledge includes CNVs that predispose one to adult-onset disorders or are associated with neoplasia. Some studies have assessed the potential of CMA to produce incidental findings that are primarily associated with cancer, and found that the frequency of CNVs that affect cancer genes was 0.18–0.6% in a general clinical cohort [103–105]. This raises an ethical question whether the incidental findings should be reported to the patients or not. It is generally proposed that pre-test genetic counseling should include the possibility of identifying presymptomatic conditions and discussion of disclosure policies. However, these recommendations are not always kept, especially in cases when a CMA test is ordered by a physician without written informed consent from the patient [3, 103]. In addition, SNP arrays could unravel cases with misattributed paternity or consanguinity. However, the specific familial relationship or degree of parental relatedness cannot always be extrapolated from the inbreeding coefficient; therefore, speculations regarding a specific relationship must be avoided in clinical laboratory reports [106].

1.1.5. Use of chromosomal microarray analysis in prenatal diagnostics

The utility of CMA in postnatal clinical diagnostics has stimulated interest in its potential prenatal applications; however, the use of CMA in prenatal diagnostics ought to be approached with caution due to possible ethical concerns and technical aspects of the analysis (e.g., choice of the type of array platform and resolution). Indeed, there are distinct differences in how genetic information from prenatal subjects is used. For example, genetic counseling is mostly based on incomplete knowledge of the fetal phenotype which contrasts postnatal referrals. In addition, most of the current literature describes the clinical effect of microdeletions/microduplications on the basis of postnatal recognition and only limited information is available regarding prenatal cases [107]. Several studies have attempted to estimate the potential benefits of applying CMA for prenatal testing in comparison with conventional fetal karyotyping, and have discussed the limitations and possible problems associated with this approach [108–111]. The main advantages of CMA over G-banding reported by the most authors include the higher resolution which yields more genetic information; the use of DNA samples extracted directly from the patient material which results in more rapid turn-over time; the possibility to analyze samples from stillbirth or fetal demise; avoiding artifacts associated with cell culturing; and the use of standardized protocols and automated analysis of the measured data which minimizes human error. Until 2013, karyotyping was recommended by the American College of Obstetrics and Gynecology and the Italian Society of Human Genetics as the principal cytogenetic test in prenatal diagnosis while

CMA was recommended to be used as an adjunct method when a structural anomaly is seen using ultrasound imaging [112, 113]. However, new recommendations based on the results of a comprehensive study on parallel application of karyotyping and CMA in a large cohort of pregnant women (n=4406) appeared near the end of 2013 [111, 114]. The main prenatal CMA target groups include pregnancies where ultrasound imaging reveals abnormalities, and in cases of fetal demise and stillbirth. For patients undergoing invasive prenatal diagnostic testing due to reasons other than fetal structural abnormality, either karyotyping or CMA can be performed [114].

The main positive aspect of applying CMA to prenatal diagnostics is its significantly higher diagnostic yield (two-fold to three-fold) compared with karyotyping. The presence of anomalies found using ultrasonography increases the possibility of detecting pathogenic CNVs: additional clinically relevant information is obtained by CMA in roughly 6–10% of cases with abnormal ultrasound findings and normal G-banding, and in ~1.7% of pregnancies with standard indications (e.g., advanced maternal age and positive aneuploid screening test) [111, 115]. Thus, prenatal CMA analysis is more beneficial when ultrasonographic examination identifies fetal structural abnormalities [114]. However, in Belgium, CMA is performed on all prenatal cases because clinically important findings are detected in normal fetuses in addition to those with ultrasound abnormalities [116]. This could be completely defensible because many microdeletion/microduplication syndromes associated with neurodevelopmental problems and ID manifest only subtle fetal abnormalities which cannot be detected using prenatal ultrasound imaging. Interestingly, no significant difference is observed in the diagnostic yield between fetal groups with a single malformation and multiple malformations [117].

The issue which causes the most ethical controversy is the potential of CMA to produce incidental findings and variants that have uncertain clinical significance. According to various studies, VUCS may occur in up to 6.7% of all cases [111, 115, 118]. This level of uncertainty could invoke considerable anxiety in prospective parents and creates the potential to terminate a pregnancy based on an uncertain finding. However, the latter seems to be unlikely scenario. Moreover, the issue of VUCS is not unique to CMA, because conventional karyotyping can also reveal findings of uncertain clinical impact, such as supernumerary marker chromosomes or apparently balanced *de novo* aberrations [119]. In Belgium, a national consensus was made to only report CMA findings to the patients when the variant(s) comes with a high suspicion of being pathogenic [116]. The choice of an appropriate microarray platform could be one possible way to regulate the amount of VUCS obtained by prenatal CMA. The use of targeted arrays designed specifically to identify known genetic conditions reduces the number of VUCS, yet at the same time means that rare or novel genomic imbalances may also be missed [119]. Therefore, the arrays should provide a uniform coverage of the entire human genome. The resolution, in turn, should be a balance between maximizing the detection rate of the test

and minimizing the number of VUCS, because a significant positive relationship between the increase in VUCS rate and the overall detection rate has been observed [115]. The Belgian national consensus proposes to use arrays with an average resolution of 400 kb, which is higher than that provided by karyotyping, but lower than the full potential of postnatal CMA [116]. Some studies propose to use the same platform for both prenatal and postnatal settings [120]. Also, sufficient communication between cytogeneticists and clinical geneticists, combined with parental analysis and in-house or international data, can significantly reduce the number of VUCS detected [117].

Another debatable issue is the use of incidental findings. These can be defined as any findings which are not directly related to the indications for which the test was performed (e.g., *BRCA1/2* deletions). Incidental findings can be divided into three main categories: late-onset disorders with clinical utility, late-onset disorders without therapeutic treatment possibilities, and evidence of consanguinity and nonpaternity. Generally, the testing of children for late-onset genetic diseases is strongly discouraged, however, when this sort of information is obtained incidentally through the use of CMA, it should not be withheld [97]. This approach could also be applied for prenatal CMA testing. In addition, one must consider whether this information could also be relevant to the health of the parents [116, 119]. Late-onset disorders for which therapeutic options are not available (e.g., Alzheimer disease) are an exception, however, these types of diseases are rarely found using CMA. Although SNP arrays have the capacity to detect consanguinity, it should be kept in mind that they are not designed to definitively assign a specific relationship between the parents or to be used as a paternity test. Laboratories are encouraged to return this type of information to the ordering clinician, however, speculation on a specific relationship should be avoided [106].

CNVs found in susceptibility loci are another group of finding that poses a challenge in genetic counseling. These represent genetic risk factors with reduced penetrance and/or variable expressivity and occur with a frequency of 1/250 [121]. In most cases, the degree of clinical manifestation of such finding is difficult to predict. In Belgium, seven CNVs for which the risk of a severe phenotype is sufficiently large and/or which are associated with structural malformations for which ultrasound follow-up is warranted (i.e., distal del1q21.1, distal dup1q21.1, proximal del1q21.1, distal del16p11.2, del16p11.2, del17q12 and dup22q11.2) are reported to the patients, while the remaining susceptibility CNVs are not reported [116].

The inability of CMA to detect balanced aberrations, such as balanced translocations and inversions, should not be a limiting factor that constrains the application of CMA in prenatal diagnostics. This type of chromosomal rearrangement occurs in roughly 0.08 to 0.09% of prenatal diagnosis samples [122]. However, familial events in all probability will have no consequences for the current pregnancy. As for *de novo* apparently balanced abnormalities, which are associated with a 6.7% risk of congenital abnormalities, the clinical

manifestation is often a consequence of a gain or loss at the breakpoint, which can be observed using CMA [111].

In general, what is required is a comprehensive patient pretest followed by posttest genetic counseling regarding the benefits, limitations, and results of CMA. This should also include a discussion of the potential to identify findings of uncertain clinical impact, nonpaternity, consanguinity, and adult-onset disease [114]. Prenatal CMA ought not to be ordered without informed consent, and various protocols have been proposed with an option of allowing the patients to determine the range of possible outcomes that will or will not be reported back to them. It is noteworthy that when pregnant women are given a choice between CMA and standard karyotyping, 70% choose CMA [119].

1.2. Speech and language disorders associated with 7q31 genomic region and *FOXP2* gene

Developmental disorders that impact speech, language, and communication represent one of the most common reasons for pediatric referrals and represent a wide range of conditions with overlapping but heterogeneous phenotypes and underlying etiologies [123]. They show a significant overlap with associated developmental conditions, such as attention deficit hyperactivity disorder (ADHD), dyslexia, and autism [124]. It is generally thought that the genetic reasons underlying susceptibility to speech and language disorders are multi-faceted in nature and could involve complex interactions between common genetic variants and environmental factors [125]. Despite this complexity, significant progress has been made in identifying and studying genes associated with risk and decoding the biological basis of human spoken language [126].

Forkhead-box P2 (*FOXP2*), located on 7q31, is the first gene implicated into the etiology of speech and language disorders through both the genetic analysis of a three-generation family (referred to as “KE”) and an independent individual (referred to as “CS”) presented with verbal dyspraxia [127, 128]. Several studies have identified additional individuals with *FOXP2* point mutations, deletions involving or translocations disrupting the gene, in which all patients exhibit verbal dyspraxia. This evidence provides further support to the involvement of *FOXP2* in speech and language impairment [129–134]. In total, there are at least 27 reported cases with *FOXP2* lesions associated with speech and language disorder. The core phenotype includes developmental verbal dyspraxia (DVD [MIM 602081]), also known as childhood apraxia of speech (CAS), which is a rare (<0.01%), severe, and persistent disorder [135, 136]. Individuals with CAS have difficulties controlling the movement and sequencing of orofacial muscles which cause deficits in the production of fluent speech. In addition, virtually every aspect of grammar and language is affected [137, 138]. The speech of those affected is often “unintelligible”, with omission, substitution, and distortion of consonants and vowels, inconsistent errors across multiple

repetitions and prosodic impairments [130, 132, 133]. A severe receptive and expressive language disorder usually occurs [131, 133]. In some patients, reading and spelling impairments are also reported. However, verbal skills are generally poorer than non-verbal skills [136, 139].

Currently, *FOXP2* remains the only known albeit rare cause of CAS and in total it contributes to a relatively small number of speech and language disorder cases. Therefore, it seems increasingly unlikely that this gene represents a general risk factor for genetically complex forms of language impairment. Nevertheless, *FOXP2* targets could be good candidates for involvement in more common forms of language impairment [124, 140].

The *FOXP2* gene encodes an evolutionally conserved transcription regulator (primary a repressor) containing a zinc-finger motif, a fork-head DNA-binding domain, and a polyglutamine tract, which is widely distributed in the fetal and adult brain (as well as other tissues) and regulates a variety of genes, most of which are involved in the cortical, basal ganglia and cerebellar circuits [141]. The effect of FOXP2 can vary greatly between tissues and developmental time points [127, 142–144]. However, the exact cellular mechanisms underlying the involvement of FOXP2 in speech and language development remain obscure. Data from expression studies, neuro-imaging, and animal studies suggest that FOXP2 may regulate neurite growth, dendritic morphology, synaptic physiology of basal ganglia neurons, and synaptogenesis [145–148].

Because of the known involvement in speech and language development, *FOXP2* has been intensively studied from an evolutionary perspective. Remarkable similarities of sequence and expression patterns have been noted across vertebrate species. Comparing mouse and human genomes, this is one of the most conserved genes which differs, besides the difference in polyglutamine tract length, by only three coding positions (reviewed in [124]). Interestingly, *FOXP2* is one of the few human genes that differ from the chimpanzee version. The two-substitution difference occurred after the divergence of humans from their common ancestor with chimpanzees roughly 4–7 million years ago and animal studies suggest that these amino acid changes could have contributed to the evolution of human speech and language by adapting cortico-basal ganglia circuits [147]. It was initially assumed that at least one of the two substitutions was the cause of the selective sweep and therefore this would not be present in Neanderthals, who split from humans at least 400,000 years ago. However, both nucleotide substitutions were found in Neanderthals [149]. Recently, a novel substitution shared by nearly all present-day humans but absent in Neanderthals was found and is likely to alter the regulation of *FOXP2* expression and, therefore, is a good candidate for having caused a recent selective sweep in the *FOXP2* gene [150].

Due to a high degree of conservation across species, the effects and functions of FOXP2 could be studied in animal models, which generally include mice and songbirds. For example, knocking down *FoxP2* in songbirds disrupts imitation of tutor song. In mice, this results in reduced growth, severe motor delays and

early death. No humans have been identified with homozygous *FOXP2* loss, indicating that this gene is crucial for normal development. Heterozygous mice carrying the KE family specific mutation show significant deficits in learning motor skills. At the same time, the relationship between *FoxP2* disruption and altered mouse vocalization is not straightforward and any conclusions should be made with caution in regard to the correspondence between mice vocalization and human speech (reviewed in [126, 141]). Obviously, no animal model could adequately mirror all of the multiple processes required for human language.

Intriguingly, a role for the differential parent-of-origin expression of *FOXP2* in human speech has been proposed by Feuk and colleagues. By assessing *FOXP2* expression in different groups of patients they concluded that the majority of *FOXP2* transcripts are of paternal origin, because patients with maternal UPD7 and patients with paternal deletions exhibit the lowest expression of *FOXP2*. Therefore, they hypothesize that *FOXP2* may represent a paternally expressed, maternally imprinted gene and that the absence of paternal *FOXP2* could cause more severe phenotypes with CAS, whereas the lack of a maternal allele should be relatively benign [132]. However, there is no direct evidence of imprinting of *FOXP2* itself, and, in addition, other studies have revealed that alterations of maternal origin are sufficient to cause the CAS phenotype [127, 129, 134].

1.3. Reciprocal deletions and duplications of 5q35.2-q35.3 and conveyed phenotypes

During the past 15–20 years, it has become evident that higher-order genomic architectural features can predispose one to DNA rearrangements [1]. Rearrangements can be classified as either recurrent or nonrecurrent depending on whether they can be identified in unrelated individuals. Recurrent rearrangements – those of common size and having clustered breakpoints – most frequently result from a mechanism of non-allelic homologous recombination (NAHR) between region-specific low-copy repeats (LCRs – also known as segmental duplications), which represent DNA fragments >1 kb in size that occur in two or more copies per haploid genome with the different copies sharing >90% DNA sequence identity. LCRs that are longer than 10 kb, share over ~97% of the sequence identity, and are located within <5–10 Mb of each other can lead to local genomic instability and stimulate CNV formation [68, 151]. The NAHR mechanism predicts that for every recurrent deletion with breakpoints mapping within directly oriented LCRs located *in trans*, a reciprocal duplication should exist (**Figure 2**) [20]. However, a relatively small number of reciprocal events have been described to date. One possible explanation for this phenomenon might be that there is a milder and less distinct phenotype associated with the duplications or even the absence of a specific clinical phenotype [152–154]. The number of reported patients with duplications

reciprocal to any specific deletion is also quite modest, however, this could be partially explained by the fact that NAHR mechanism theoretically produces more deletions than duplications which was confirmed experimentally [68, 155].

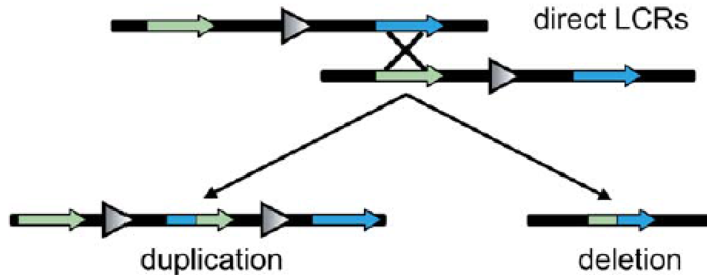


Figure 2. Schematic representation of non-allelic homologous recombination between directly oriented low-copy repeats (LCRs) located on homologues chromosomes. Predicted products of such rearrangements are reciprocal deletions and duplications. Colored arrows represent LCRs. Recombination is signified by "X". Shaded triangles with grey gradients represent unique directional segments within a chromosome (modified from [151]).

Implementation of array techniques has revealed the existence of recurrent reciprocal deletions and duplications in several genome loci [89, 156–161]. The phenotypic outcome of such reciprocal events can be variable [152–154]. The most intriguing feature is that in some cases a reciprocal or mirror phenotype between deletion and duplication cases has been observed [89, 157, 159, 162].

One such example occurs in the genomic region 5q35.2-q35.3. Deletions in this locus lead to Sotos syndrome (Sos) (OMIM# 117550) which is characterized by childhood overgrowth with advanced bone age, craniofacial dysmorphic features that include macrocephaly, and learning difficulties [163, 164]. Although, more than 30 genes are located in this region, the only known cause of Sos is haploinsufficiency of the *N*uclear receptor *S*et *D*omain containing protein 1 gene, *NSD1*, that occurs due to intragenic mutations, partial gene deletions, or 5q35 microdeletions [165]. Delineation of the precise architecture of the region has shown that the mechanism responsible for the generation of a common recurrent 1.9 Mb deletion in Sos patients is the NAHR mechanism which occurs between two directly oriented LCRs, which thus predicts the existence of a reciprocal duplication [166]. The condition was first clinically described in one family over three generations and is termed Hunter-McAlpine syndrome [167]. Later it was found to be caused by subtelomeric duplication of 5q35-qter [168]. Since then several reports describing patients with duplications in the Sos critical region have been published, and a novel syndrome has been proposed that is clinically opposite to Sos with regard to growth and head size [159, 169–173]. The phenotypes of all patients are very similar: short stature since the birth, microcephaly, brachydactyly, mild to

moderate ID, and mild facial dysmorphism. Delayed bone age is also evident in most cases where the X-ray investigation was done [159, 172, 173]. However, a patient with a larger duplication is affected more severely, and displays, in addition motor retardation, bilateral inguinal hernias, strabismus, and abnormal 5th fingers [170]. In general, phenotypic characteristics related to growth seem to be remarkably opposite comparing deletion and duplication of 5q35.2-q35.3.

The only gene known to be associated with Sos, and subsequently with tall stature and macrocephaly, is *NSD1*. To date, there is no evidence that disturbance of the dosage of genes neighboring *NSD1* have any specific effect on the phenotype [174]. Based on this evidence it was proposed that increased dosage of *NSD1* in patients with duplication is responsible for short stature and microcephaly, however, the molecular mechanism through which the deletion/duplication could influence the phenotype remains to be determined. Concurrently, it was shown that Sos patients carrying 5q35 microdeletions display more variable phenotypes compared with patients that have *NSD1* mutations and are often diagnosed with additional clinical problems [175]. Thus, it is possible that other genes within the Sos critical region could be dosage-sensitive and work to produce a greater variability of observed phenotypes. The smallest duplication identified to date is ~200 kb in size and encompasses, in addition to *NSD1*, nine additional genes, three of which are disease-associated and may therefore contribute to the duplication phenotype although the clinical effect of a duplication of these genes is not known [173]. Recently, a mouse model of Sos, *Df(13)Ms2Dja*$^{+/-}$*,* carrying a deletion syntenic to the human chromosome 5q35.2-q35.3 region was generated and phenotypically characterized, however, it was found that genetically deficient mice show significantly reduced postnatal growth and no changes in head and facial morphology. This study suggests a divergent role for *Nsd1/NSD1* in regulating growth in mouse and human. However, *Df(13)Ms2Dja*$^{+/-}$ mice may potentially recapitulate the learning disability characteristic for Sos and kidney abnormalities observed in a subset of Sos patients, which suggests a potential role in phenotype formation for other genes in the deleted region [176]. To date, no data exists relating mouse models of duplication reciprocal to the common Sos deletion. Obviously, gene expression studies could aid in our understanding of the exact mechanisms whereby structural changes in 5q35.2-q35.3 in general, and NSD1 in particular influence human growth.

1.4. Detection of mosaicism and uniparental disomy by single nucleotide polymorphism array analysis

CMA analysis using SNP microarrays does not only enable one to detect CNVs, but it also provides genotype information at multiple polymorphic loci throughout the genome. This genotype data further confirms CNV calls, and in addition, offers the capability to detect LCSH and provides the sensitivity to detect mosaicism. Although, the clinical utility of SNP arrays for CNV detection

is well accepted, the number of studies concerning their use in clinical settings for the detection of homozygosity and mosaicism is quite modest [43, 177, 178].

1.4.1. Genetic mosaicism

Genetic mosaicism is recognized as the presence of two or more populations of cells with different genotypes in a single individual who has developed from a single fertilized egg. This may be a consequence of meiotic or mitotic error, and generally the occurrence mechanism has a profound effect on the developing fetus as well as on the adult individual. Mosaicism has been reported for many types of chromosomal abnormalities with mosaic aneuploidy being the most common type of mosaicism [43]. Furthermore, application of array techniques has revealed that humans are commonly affected by somatic mosaicism for stochastic CNVs, which occur in a substantial fraction of cells [179]. It also appears that theoretically, genetically identical monozygotic twins can actually display different CNV profiles [180]. The phenotypic consequences of mosaicism depend on the type of chromosomal aberration, the mechanism and time of its occurrence, and the percentage of abnormal cells.

Although mosaicism has been historically detected using conventional cytogenetic methods, these methods have limitations in terms of resolution and an inability to identify copy-number neutral events, such as LCSH. Also, the identification of lower levels of mosaicism can be challenging because this requires one to analyze more cells than are typically analyzed in conventional laboratory practice and is not usually performed in the absence of specific indications. In addition, analysis of metaphases may provide a biased view of the true chromosomal constitution in a patient because in some types of mosaicism the abnormal cells, or in contrast normal cells, may not divide. These limitations are overcome by introduction of CMA because the number of cells investigated is much higher as compared with conventional chromosome analysis and the culture bias introduced by analyzing only metaphase cells is eliminated. Generally, it has been proposed that a low level of mosaicism may be more common than previously anticipated [43]. In the case of aCGH platforms, the presence of mosaicism is indicated by shifts in LRR, and the minimal detection of mosaicism is estimated to be 10–20%. Using SNP arrays, mosaic changes are detected by assessing if aberrations in LRR occur along with a shift in genotype frequencies of the SNP probes, and the minimal level of mosaicism that can be detected is less than 5% [43, 181]. This is explained by the fact that BAF is more sensitive to subtle loss or gain of haplotype than LRR is to the subtle shifts in intensity levels because the intensity data is normalized and logarithmically transformed. The percentage of mosaicism can be calculated from the relative shift in BAF. In addition, the information extracted from the genotypic data can be used to differentiate between a mitotic and meiotic origin of mosaicism which is essential to determine the risk of recurrence and thus proper counseling (**Figure 3**) [43].

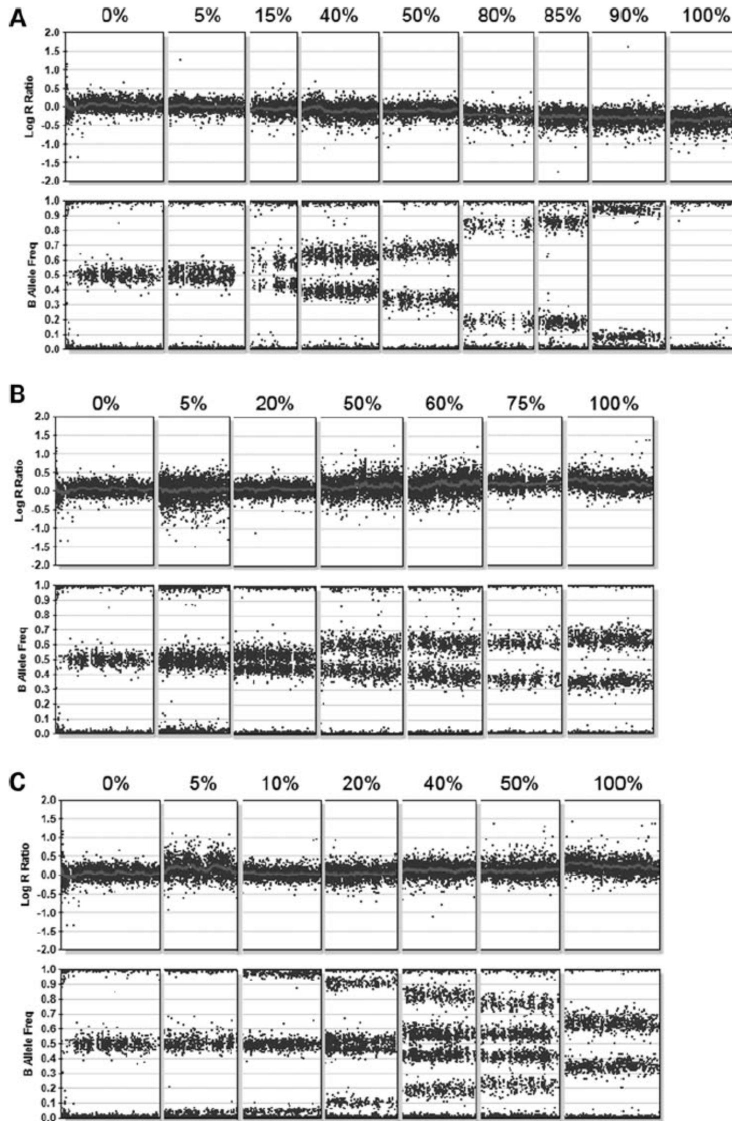


Figure 3. Examples of SNP array results for mosaic deletions and duplications. The figure shows segments from different chromosomes illustrating mosaicism from 0–100%. For all subfigures, the percentages above the data indicate the level of mosaicism, with 0% representing a patient with a normal copy-number, and 100% representing a non-mosaic patient. **A.** Nine patients with varying levels of mosaicism for deletions involving autosomes. **B.** Seven patients with varying levels of mosaicism for trisomies arisen from mitotic error. The pattern of BAF indicates that the same haplotypes present in the euploid cell line are also present in the triploid cell line at altered ratios. **C.** Seven patients with varying levels of mosaicism for trisomies arisen from meiotic error, which is indicated by the presence of additional genotypes in the trisomic cell line that are not present in the euploid cell line [43].

1.4.2. Excessive homozygosity and uniparental disomy

The phrase "long contiguous stretches of homozygosity" refers to uninterrupted regions of homozygous alleles with a genomic copy-number state of 2 [64]. In population genetic studies, the minimal threshold for LCSH calls is usually set at ~1 Mb, while in clinical settings this threshold is set to between 3 to 10 Mb [106, 177, 182]. Detection of excessive homozygosity, in and of itself, is not diagnostic of any underlying condition and may be clinically benign [64]. The clinical consequences of LCSH depend on the genomic context, such as the presence or absence of recessive pathogenic mutations or imprinted genes in the region. Potentially, LCSH regions of any size can be helpful during the diagnosis of autosomal recessive diseases through homozygosity mapping and the selection of candidate genes for sequence analysis. This, however, requires effective communication between the managing clinician and laboratory [178].

Generally, LCSH can indicate ancestral homozygosity, uniparental disomy, or parental consanguinity. Short homozygous regions (up to 5 Mb) are considered ancestral markers and are present in all outbreed populations. They tend to be co-localized in different populations and reflect the presence of long ancestral haplotypes that remain intact because of low rates of recombination locally [182, 183]. Multiple large blocks of homozygosity that are distributed across the genome can point towards a possible parental blood relationship, and the percentage of the genome that is in a homozygous state correlates with the degree of consanguinity [64, 106, 178]. Large LCSH (single or multiple) that are restricted to a single chromosome can be a hallmark of UPD, be it whole-chromosome UPD or segmental UPD [43, 177].

UPD is defined as the inheritance of two copies of a chromosome, or a part of chromosome, from one parent. The inheritance of two identical chromosomes from a single parent is termed uniparental isodisomy (isoUPD), whereas the inheritance of two homologous chromosomes is termed uniparental hetero-disomy (heteroUPD). Due to a meiotic crossover, partial isoUPD and partial heteroUPD (iso/heteroUPD) may co-exist within the same chromosome pair. Generally, the inheritance of both chromosomes from one parent does not cause any clinical consequences unless the chromosome either contains imprinted genes that have different expression patterns that depend upon the parent of origin, or if it unmasks mutations for autosomal recessive disease [177, 184]. UPD cannot be detected by conventional cytogenetic techniques and usually requires specific analysis, which may include microsatellite analysis and methylation analysis. Typically, this is performed only when UPD is suspected based on clinical or cytogenetic features (e.g., mosaic trisomy or marker chromosome) [185]. Therefore, the true incidence rate of UPD is currently unknown, as many asymptomatic cases remain undetected.

The use of SNP arrays enables one to identify the majority of UPD cases and gives insight into the mechanism underlying the formation of each specific case of UPD [43]. It has been found that both the size and location of LCSH correlate with UPD. The minimal threshold of ~20 Mb interstitially and 10 Mb

telomerically has been proposed to distinguish true UPD from other genomic events. This, however, requires further investigation and possible re-estimation [177]. Generally, most known UPD mechanisms can lead to mosaic formation. Therefore, knowing if additional cell lines are present and the type of UPD (hetero- or isoUPD) enables one to specify whether monosomy rescue, trisomy rescue, gamete complementation, or post-fertilization mitotic error took place in a particular case of UPD. However, the underlying mechanism often cannot be delineated if pure UPD without additional cell lines is observed. In turn, understanding UPD mechanisms can be useful to estimate the patient's prognosis and thus allows for effective counseling. In fact, the mechanism by which UPD occurred may have greater clinical implications than the presence of UPD itself [186]. For example, a UPD15 associated with mosaic trisomy 15 is linked with a more severe phenotype with a high incidence of congenital heart disease [187]. In those cases, it is difficult to separate the clinical contribution of the chromosomal abnormality and the UPD itself.

Because UPD detection using SNP arrays is generally based on the identification of long homozygous region(s) restricted to a single chromosome, this approach will fail to identify cases of complete heteroUPD if the parental DNA is not included into the analysis. For example, it has been estimated that as many as 8% of patients with Prader-Willi syndrome (PWS; OMIM #176270) will be missed with SNP-based CMA alone [184]. In addition, the detection of iso/heteroUPD cases depends on the minimal LCSH threshold set by the laboratory and thus it is clear that some cases will remain undetected. Therefore, a normal CMA result cannot exclude diagnosis in patients with a suspected imprinting disorder. Studying trios would identify all cases of UPD, including heteroUPD; however, it is not economically feasible to test the parents of each patient in a clinical laboratory setting [184]. Nowadays, there is no single approach that could provide the detection of all UPD cases in a single experiment. A step-wise or combined use of different methods (e.g., CMA with SNP arrays, MS-PCR (methylation-specific polymerase chain reaction), MS-MLPA (methylation-specific multiplex ligation-dependent probe amplification)) is recommended to differentiate between UPD, deletion/duplication, and/or methylation disturbances [185, 188]. The use of pure qualitative methylation analysis should be avoided because this method is uninformative in mosaic UPD cases.

2. AIMS OF THE STUDY

The aims of the current study are as follows:

1. To evaluate the application of CMA in routine clinical practice in Estonia as a first-tier diagnostic test for individuals with ID/DD, MCA, or ASD (Publication I).
2. To estimate the efficacy of CMA in prenatal diagnostics (Publication I).
3. To evaluate the clinical manifestation of 7q31 deletions that encompass the *FOXP2* gene and to determine whether there are any parent-of-origin-dependent differences in severity of speech and language problems (Publication II).
4. To investigate the clinical consequences of 5q35.2-q35.3 duplication reciprocal to common Sotos syndrome deletion (Publication III).
5. To use SNP-based CMA for diagnosing a patient with mosaic matUPD15 and to offer a workflow for evaluating imprinting disorders (Publication IV).

3. MATERIALS AND METHODS

3.1. Study subjects

3.1.1. Patients studied for chromosomal aberrations by chromosomal microarray analysis

All samples in this study were collected between January 2009 and December 2012, a total of 1191 patients (male/female ratio 58/42). The median age of the patients was 5 years (range: newborn – 83 years). Although probands constituted the vast majority of all analyses [1072 (95%)], in some cases additional family members were investigated [59 (5%)]. In addition to postnatal referrals [1131(95%)], CMA was performed for some prenatal cases [60 (5%) fetuses].

Since 2011, the cost of CMA is covered by the Estonian Health Insurance Fund, and is performed as the first-line cytogenetic diagnostic test in patients with DD/ID, ASD and/or MCA, as recommended by the ISCA Consortium [5]. The patient population could therefore be divided into two groups: those sent for CMA in 2009–2010 (a total of 188 individuals) and those analyzed in 2011–2012 (1003 patients). Informed consent for regular medical evaluation and investigation was taken from all patients by the referring doctor. The patient population sent for CMA before 2011 was very carefully selected and consists of patients with an unknown diagnosis despite all routine genetic investigations (e.g., standard G-banding karyotyping, routine metabolic analysis, test for fragile X syndrome, as well as tests for Prader-Willi/Angelman syndrome (PWS/AS) or other genetic disorders in the case of specific indications). In contrast, the spectrum of indications for patients tested by CMA between 2011 and 2012 was very wide, and ranged from severe disabilities to nonspecific milder problems such as learning difficulties at school. Generally, currently the patients could be referred to CMA by any physician, including general practitioners, however, most patients are sent by clinical geneticists, neurologists, and neonatologists.

Currently in Estonia, CMA is not applied as a first-line cytogenetic test in prenatal diagnosis. However, it was performed in a limited number of prenatal cases, mostly in parallel with conventional karyotyping. The main indications that led to prenatal CMA included an abnormal ultrasound finding, family history of chromosomal abnormalities, a positive aneuploidy screening result, and/or other exceptional conditions (e.g., complicated anamnesis and repeated miscarriages).

3.1.2. Detailed clinical description of patients with 7q31 deletion from Publication II

The patients and their families presented in Publication II were included in a study of ID in the Estonian population and were described for the first time as cases 7 and 8 in Mannik et al. [8]. This study was approved by the Ethics Committee on Human Research at the University of Tartu.

3.1.2.1. Family 1

The family 1 (**Figure 4A**) proband (F1-III-2) was born prematurely at the 37th week of pregnancy with a low birth weight 2131g (-2 SD), length 44 cm (–2 SD), head circumference (OFC) 32 cm (–1.5 SD) and an Apgar score of 8/8. At the age of four days, extrasystoles were observed; however, the echocardiogram revealed only persistent ductus arteriosus. At the age of six months she was hospitalized due to an acute urinary tract infection. A kidney ultrasound investigation revealed bilateral duplex kidneys and ureters, and megaureters that later required repeated surgical intervention. On that occasion her family received the first consultation by a clinical geneticist. She had DD, mild failure to thrive – length 62 cm (–2 SD), weight 5850g (–2 SD) and OFC 41.5 cm (–2 SD), a high and prominent forehead, hypertelorism, a small nose, a cupid-shaped upper lip, a high palate, a small mouth, a low-set left ear, overriding toes and partial syndactyly of toes II–III (refer to **Supplementary Figure 1A,B,C** in [189]). In addition, vision problems were observed: nystagmus (disappeared at 18 months), convergent strabismus, positive Graefe symptom, and myopia. A brain MRI at the age of eight months showed mild brain atrophy and mild bilateral hyperintensity in the white matter. She began to walk at the age of two years and eight months. At the age of three years her height was 90 cm (–1 SD), weight 11.82 kg (–2 SD) and OFC 49.5 cm (–1 SD). She was unable to sneeze, had problems chewing food and swallowing, and displayed pronounced drooling. Her cough reflex appeared at the age of 5–6 months. She had frequent upper-airway infections and bronchitis. A mild hand tremor was sometimes evident. Additionally, some autistic features were observed. Neuropsychological evaluation with Bayley-III (Bayley Scales of Infant and Toddler Development-Third Edition [190]) revealed moderate DD. The Composite scores were: Cognitive – 55, Language – 58 (Receptive Communication = Expressive Communication), Motor – 48 (Fine Motor < Gross Motor), Social-Emotional – 55, and corresponded to a developmental level under the age of 16 months. The child's vocalization activity was quite low. She used several words, but not always for an evident purpose. Once in a while she pointed to an object to express her needs. The child had difficulties repeating words, imitating actions or initiating play. Her exploration and manipulative activity with objects was quite limited, and she had attention difficulties. Deficits in Adaptive Behavior were particularly considerable in the areas of leisure, self-care, self-

direction and social skills. The substantiation of a specified level of ID, however, required follow-up assessments.

Chromosomal analysis from peripheral blood lymphocytes, tests for 7q11.23 and a 22q11.2 deletion, urine organic acids, as well as serum and urine amino acids were performed on the patient and reported as normal.

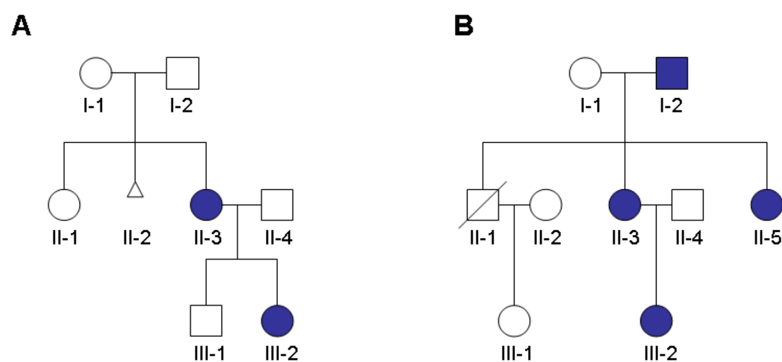


Figure 4. A. Pedigree of Family 1 (F1); **B**. pedigree of family 2 (F2) [189]

The mother of the proband (F1-II-3) had a similar clinical phenotype. She was born prematurely at the 37th week of pregnancy with an Apgar score of 7/8. Shortly after birth she had breathing problems and required oxygen therapy. Nystagmus was observed during the first three months and she began to walk at 14 months. During the first two years of her life she had difficulties swallowing and had pronounced drooling. Chewing food was also problematic. Her cough reflex developed at the age of 5–6 months, and she was unable to sneeze. She had moderate speech delay (her first words came at the age of 3.5 years) and verbal dyspraxia. During the first years of life, a mild hand tremor and some autistic features were also observed. Evaluation at the age of 28 years showed a tall height of 178 cm (+2.5 SD), mild obesity (BMI=26.8), a high and wide forehead, hypertelorism, mild divergent strabismus, a high palate, and many lentigines on her whole body (refer to **Supplementary Figure 1D** in [189]). She currently has a problem with speech expression, especially when nervous. Some lack of social skills and emotional lability were also present. She has a primary school education (9 grades). The neuropsychological evaluation using the KAIT test at the age of 28 (The Kaufman Adolescent and Adult Intelligence Test [191]) resulted in a below average general intelligence (intelligence quotient (IQ) 88). Her language abilities (based on verbal expression, verbal comprehension or word storage and retrieval) were more impaired than abilities based on visual processing modality.

3.1.2.2. Family 2

The family 2 (**Figure 4B**) proband (F2-III-2) was born normally by cesarean section with a normal birth length of 48 cm (–1 SD), weight 3220 g (–0.5 SD), OFC 34 cm (–0.5 SD) and an Apgar score of 7/8. Evaluation at the age of six years revealed a normal height of 121.5 cm (+0.5 SD), weight 20 kg (–0.5 SD), OFC 53 cm (+1 SD), long face, high forehead, broad eyebrows, long eye-lashes, short philtrum, thin upper lip, high palate and dental anomalies (refer to **Supplementary Figure 1E,F** in [189]). In addition she had a narrow chest, mild chest deformity, broad toes I–III, and mild syndactyly of toes II–III. Swallowing, chewing, coughing or sneezing problems were not observed by caregivers in the child care institution where she lives. She had mild ataxia and sometimes aggressive behavior. Brain MRI and EEG were found to be normal. Her developmental delay excluded administration of age-appropriate tests in the assessment of abilities. Neuropsychological evaluation with Bayley-III at the age of 6 revealed DD in all areas. Her developmental level was under the age of 24 months on the Cognitive and Language Scale (Receptive Communication > Expressive Communication), and under 28 months on the Motor Scale (Fine Motor < Gross Motor). She had significant pronunciation difficulties. Her speech was inarticulate and hardly understandable to others. Her vocabulary was plain and poor. She used some words and phrases, accompanied by restricted nonverbal communicational cues. She was not likely to repeat words, but was likely to repeat a few simple activities. Her cognitive ability, exploring behavior and manipulation of objects were all limited. Her attention was easily distracted and she showed considerable impairment in fine motor skills. The expression of emotions was limited. Compared to other skills, her self-care skills were less impaired (as appropriate for the age of 29 months). The evaluation of socio-emotional development and adaptive behaviors were not possible due to the absence of appropriate information (e.g., caregiver's reports).

Tests to identify the etiology of the patient's problems included chromosomal analysis from peripheral blood lymphocytes, fragile X testing, DNA analysis for spinal muscular atrophy and acylcarnitine tandem MS analysis. No abnormalities were detected.

The mother of the proband (F2-II-3) had moderate MR, aggressive behavior and verbal dyspraxia. Due to a mood disorder, she refused clinical and molecular investigation. The maternal aunt of the proband (F2-II-5) had mild MR and aggressive violent behavior. Her blood sample was available for molecular studies, but detailed clinical investigation was impossible due to her imprisonment. The grandmother of the proband (F2-I-1) had a primary school education (7 grades) and has been working as a farm worker her entire life. The grandfather of the proband (F2-I-2) completed only 4 grades at school. According to the grandmother, he had a severe speech defect, aggressive behavior and balance problems (e.g. inability to bicycle). It was impossible to perform a clinical and molecular investigation of him.

3.1.3. Clinical description of a patient with 5q35.2-q35.3 duplication from Publication III

The patient was born prematurely after 35 weeks of gestation with a low birth weight 2110g (–1.5 SD), length 45cm (–1.5 SD), and microcephalic skull circumference 29 cm (–2 SD). The Apgar score was 7/7. At the age of 9 months her height was 62 cm (–4 SD), weight 5600g (–4 SD) and occipitofrontal circumference (OFC) 41 cm (–3 SD). At the age of 6.5 years she received her first consultation by a clinical geneticist due to her ID and short stature (104 cm, –4 SD). Clinical evaluation revealed microcephaly (OFC 48 cm, –2 SD), almond shape eyes, epicantic folds, wide nasal bridge, broad and prominent nasal tip, thin lips, small mouth, dysplastic ears, short neck, and brachydactyly.

The second time she was sent for genetic counseling was at the age of 13.5 years by a child endocrinologist who had followed her growth since age six. All endocrine investigations were found to be normal excepting diagnosis of a delayed bone age by two years (at age six) and idiopathic short stature. Clinical evaluation of the patient at 13.5 years revealed short stature (137.5 cm, –4 SD), normal growth velocity (5.1 cm/year) microcephaly (50 cm, –3.5 SD), brachydactyly, delayed bone age (12 years), moderate ID, similar facial dysmorphic features as described earlier (refer to **Figure 1a, b, c** in [192]). In addition she had clinodactyly, mild syndactyly of IV–V toes, hirsutism in the upper lip and back. X-ray investigation showed brachydactyly and cone shape epiphyses (refer to **Figure 1d** in [192]). Brain MRT was normal. Her hormonal status was within a normal range and she experienced a normal and relatively early development of puberty.

Because regular chromosomal analysis, test for 1p36.3 and 17p11.2 microdeletions, as well as standard metabolic tests showed normal results, the patient was further analyzed by CMA (in 2010).

Family history: the proband has a three-year younger sister, who has mild DD, epilepsy, and normal growth. Her father has a normal height (174 cm, –1.0 SD) and studied in the school for mentally retarded. Unfortunately her father was not available for a genetic evaluation.

3.1.4. Clinical description of a patient with mosaic maternal UPD15 from Publication IV

The patient (male, 18 years old) was born normally after 37 weeks of gestation with a birth weight of 3000g. His growth and development were normal up to school age when his growth rate slowly started to decrease. At the age of 10, he was tested for possible growth hormone deficiency, however, his anterior pituitary function tests were normal, including a normal pituitary on MRI and a normal male karyotype (46,XY). He was overweight (BMI > 85th centile), had small genitalia, and delayed bone age by three years. Over the next three years he was followed by a pediatric endocrinologist. During these years he grew normally (height along 25th centile, weight along 75th centile), his puberty

started normally and at the age of 13, he was discharged from constant observation.

At the age of 16, he was referred to a child neurologist due to multiple complaints, which were observed during the previous 3–4 months. He had recurrent balance problems, headaches, muscle weakness, uncoordinated movements, tremors and diplopia. The severity of symptoms was variable: on some days, he was not even able to go to the toilet, however, on other days he was completely symptom-free.

Six months later (the patient was then 16.5 years old) he was hospitalized for further tests and investigations. His height was 172.8 cm (25th centile; mid-parental height 50th centile), weight 75.0 kg (75th centile), BMI 25 kg/m^2 (> 85th centile i.e. he was overweight), and head circumference 58 cm (+0.5 SD). His pubertal development was normal (Pu5, G5, Ax2), with a testicular volume of 20 ml (dex=sin) by Tanner stages. However, mild thoracic kyphosis, hyperlordosis in the lumbar region, female body shape, gynecomastia, hyperpigmented skin, and highly pitched voice were observed. He had some microanomalies – broad eyebrows, synophrysis, epicanthal folds, small up-turned nose, crowded teeth, small chin, anteverted dysplastic ears, clinodactyly of T2-5, hypermobile joints and flat feet (refer to **Figure 1** in [193]). Neurological examination revealed muscular hypotonia, brisk deep tendon reflexes, and slight balance abnormalities. A brain MRI showed mild dilatation of lateral ventricles, and an ENMG did not reveal pathological changes. All hormonal investigations were in a normal range, except for mildly increase of 17-OH progesterone (7.3 nmol/L; normal 0.2–3.0) that was interpreted as a clinically insignificant increase. The patient was also seen by a child psychiatrist, who noted moderately increased anxiety, preoccupation with health problems, and fatigue, and diagnosed him with undifferentiated somatoform disorder. His IQ was normal, however, he had some learning difficulties and occasionally experienced bullying in school.

CMA analysis was performed in 2011 with a DNA sample from the patient.

Family history: the paternal grandfather is of Mari origin with similarly hyperpigmented skin as in the proband. He has one healthy older brother. Otherwise, the family history was unremarkable.

3.2. Methods

3.2.1. Analysis of genomic aberrations by SNP-based chromosomal microarray analysis

Screening for chromosomal rearrangements was performed using HumanCNV370 (January 2009–August 2009) or HumanCytoSNP-12 BeadChips (September 2009–up to date) (Illumina, Inc.). The HumanCNV370 BeadChip contains ~370,000 markers and covers the entire human genome with an average spacing of 5 kb, allowing an average effective resolution of 50 kb

(i.e. 10 consecutive markers). HumanCytoSNP-12 BeadChips were developed for cytogenetic diagnosis purposes; they contain ~300,000 probes and allow for an average effective resolution of ~60 kb (i.e. 10 consecutive markers). Genotyping procedures were performed according to the manufacturer's protocol. Genotypes were called by BeadStudio v.3.1 or GenomeStudio v2009.1 software GT module (Illumina, Inc.); LRR and BAF values were extracted from BeadStudio/GenomeStudio software and used in further CNV analysis and breakpoint mapping with Hidden Markov Model-based QuantiSNP software (ver. 1.1 or 2.1) [194]. The parameters suggested by the software developers were used. Only samples with a call rate >98% that passed the QuantiSNP quality control parameters were analyzed. In mosaic cases, the level of mosaicism was determined based upon visual estimation of the BAF distribution pattern [43].

Based on the recommendations provided by the ISCA Consortium and ACMG, all chromosomal aberrations detected were assigned the labels "pathogenic/likely pathogenic", "benign/likely benign" or "VUCS" [5, 97]. For this, the aberrations were compared with known CNVs listed in the Database of Genomic Variants (DGV) [15] and studied for genomic content using either the UCSC genome browser or ENSEMBL. The potential clinical significance of CNVs not present in normal individuals was estimated using both the DECIPHER and OMIM databases and peer-reviewed literature searches were performed using the PubMed database [195]. A chromosomal aberration was defined as pathogenic or likely pathogenic if it 1) overlapped with a genomic region associated with a well-established syndrome, 2) was large in size (>5 Mb) and contained a gene-rich content, or 3) contained a gene or a part of a gene implicated in a known disorder. CMA findings were considered benign or likely benign if they 1) were present in healthy individuals [e.g. healthy family members (with some exceptions) or DGV], 2) were gene-poor and did not encompass any known disease-causing genes, or 3) had not been previously reported in association with any disorders (**Table 2**). All remaining findings were categorized as VUCS.

Either FISH, qPCR, G-banding, or MLPA was used in confirmation studies. The inheritance pattern was examined either by CMA or other methods.

3.2.2. CNV confirmation by qPCR

qPCR was applied to confirm the detected CNVs and to investigate the inheritance pattern. This analysis was performed on a 7900HT Real-Time PCR system (Applied Biosystems, Carlsbad, CA, USA) using either ready-to-use Maxima™ SYBR-Green qPCR Master Mix (Thermo Fisher Scientific, Vilnius, Lithuania) or HOT FIREPol® EvaGreen® qPCR Mix Plus (Solis BioDyne, Tartu, Estonia). Primers were designed using the web-based service qRTDesigner 1.2 (http://bioinfo.ut.ee/gwRTqPCR/). To confirm each aberration, a total of eight specific primer pairs were designed: four outside the

aberration and four within the aberrant area. To eliminate non-specific variations, such as differences in the amount of DNA input or presence of PCR inhibitors, C_t values were normalized using the C_t values of two reference regions with theoretical copy-number of 2 (**Table 3**). Amplification mixtures (15 µl) contained SYBR-Green Master Mix/EvaGreen Mix, 250 nM of each forward and reverse primer, 4 ng template DNA. Each assay included test DNA, a mixture of ten healthy female DNA samples, a mixture of ten healthy male DNA samples (all in triplicate), and a no-template control (in duplex). The cycling conditions were as follows: 15 min at 95°C, 40 cycles at 95°C for 15 s and 60°C for 60 s. After PCR amplification, a melting curve was generated to check the specificity of PCR reaction (absence of primer dimers and other non-specific amplification products). The data was acquired using SDS 2.2.2 software (Applied Biosystems) and was further processed using either a spreadsheet program or qBase+ (Biogazelle, Ghent, Belgium). Analysis was performed as relative quantification using the Pfaffl method of calculation while taking into account the amplification efficiencies of each primer pair [196].

Table 3. qPCR reference primer sequences and positions.

Primer ID [reference]	Position (hg18)/ Amplicon length	Forward primer (5'→ 3')	Reverse primer (5'→ 3')
Ref1	1:28533747 / 70bp	TGCGAAACTGCGTGGACATT	ATGCGGAAGCCCATTTCCAT
Ref2 [197]	3:115437566 / 120bp	CTGTGACCTGCAGCTCATCCT	TAAGTTCTCTGACGTTGACTGATGTG
Ref3 [197]	3:99734152 / 101bp	GGTCCCTGGTGGCCTTAATT	TTGCTGGTAATGGGCACACA

3.2.3. CNV breakpoints analysis by sequencing in Family 1 and 2 with 7q31 deletion

For the exact mapping of deletion breakpoints determined *in silico* in F1 and F2 (Publication II), we used qPCR followed by Sanger sequencing. In the case of F1, four consecutive qPCR assays were required to narrow the junction fragments to between 2–3 kb, while only two qPCR assays were required when analyzing F2. The predicted junction fragments spanning the proximal and the distal breakpoints were amplified by long-range PCR using a Herculase®II Fusion DNA Polymerase (Agilent Stratagene, Santa Clara, CA, USA) according to the manufacturer's suggested protocol. Sequencing was performed using a 3730xl DNA Analyzer (Applied Biosystems). The sequences obtained were aligned against the human genome (NCBI Build 36.1) in the UCSC Genome Browser (BLAT on DNA). The breakpoints were analyzed with the Lalign web-

service which enables one to compare and determine the homology between two sequences. Palindromic sequences were analyzed using the software Dotlet [198].

3.2.4. Gene expression analysis by RT-qPCR for 5q35.2-q35.3 region

Real-time reverse transcription-qPCR (RT-qPCR) was applied to study the expression levels of genes encompassed by 5q35.2-q35.3 duplication reciprocal to a common Sos deletion. Three patients were enrolled in the study, including the Estonian patient described in Publication III, as well as two additional patients from Belgium and Germany (**Table 4**). Sex- and age-matched controls were used for all patients.

Table 4. Patients with 5q35.2-q35.3 duplication enrolled in gene expression study.

	Dup size (Mb)	Sex	Age	Reference
Patient 1	~2	F	13	[192]
Patient 2	>1.1	M	33	[159] (Case 1)
Patient 3	~1.6	M	15	[173] (Patient 1)

Total RNA was extracted from whole-blood using a Tempus™ Spin RNA Isolation Kit (Applied Biosystems) (Patient 1 and 3). In case of Patient 2, total RNA was extracted from a lymphoblastoid cell line using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). RNA samples were treated with a TURBO DNA-free™ Kit (Ambion, Carlsbad, CA, USA) and used as a template for the synthesis of complementary DNA (cDNA) with oligo(dT) primers and a First Strand cDNA synthesis Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA).

The expression of 12 genes from the Sos critical region (*HIGD2A*; *FAF2; RNF44; UIMC1* OMIM 609433; *NSD1* OMIM 606681; *PRELID1* OMIM 605733; *LMAN2* OMIM 609551; *GRK6* OMIM 600869; *PDLIM7* OMIM 605903; *DDX41* OMIM 608170; *TMED9*) and flanking regions (*HNRNPAB* OMIM 602688; *CLK4* OMIM 607969; *SFXN1* OMIM 615569), as well as two known *NSD1* target genes (*HSD17B10* OMIM 300256 and *ZMYM3* OMIM 300061) were investigated. Primers were designed using the online tool Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and are provided in **Table 5**.

qPCR was performed on a 7900HT Real-Time PCR system (Applied Biosystems) using EvaGreen® qPCR mix (Solis BioDyne). The amplification conditions were as follows: 15 min at 95°C, 40 cycles at 95°C for 15 s and 60°C for 60 s. All reactions were repeated six times. Threshold cycle values were obtained and processed using SDS 2.2.2 software (Applied Biosystems) and further analysis was performed using qBase+ software (Biogazelle) followed by meta-analysis with METAL software (http://www.sph.umich.edu/csg/abecasis/metal/).

Table 5. Genes and primer sequences used for gene expression study by RT-qPCR.

Primer ID	RefSeq ID	Gene	Forward primer (5'→ 3')	Reverse primer (5'→ 3')	Amplicon length (bp)
NSD1 target genes					
ZMYM3	NM_201599.2	*ZMYM3*	GGTCCTGGAAAACGGAAGAG	CGGAGGCTTTCAGGACATTT	128
HSD17B10	NM_004493.2	*HSD17B10*	GGTTGGACAAGCTGCATACT	AAACAGACCTGGGGCAATG	118
Genes outside the Sos critical region					
SFXN1_1	NM_022754.5	*SFXN1*	AAGAGGTTCCCATGGATGAGT	GTGTAGCAAACACCAAACAGAAG	70
SFXN1_2	NM_022754.5	*SFXN1*	CAAGCCATCACGCAAGTTG	TGAATGGAGGGATGGCCAT	70
HNRNPAB	NM_031266.2	*HNRNPAB*	TATGGCGGCTACGACTACTC	GTAGTTTGTACTACCCTGACTGTAG	81
CLK4_1	NM_020666.2	*CLK4*	CTGCAAACCGTTGAAGGAATT	GCTGCAATGCTTCATCCAAG	125
CLK4_2	NM_020666.2	*CLK4*	CACAGTCTTTCAGACTCATGATAGT	TCATGTGTTGTGGTATGGGTC	80
Genes inside the Sos critical region					
HIGD2A	NM_138820.2	*HIGD2A*	GTGGTACCCATAGGTTGCC	GCATCATGAGCTGAGAGCG	97
FAF2	ENST00000261942	*FAF2*	TGATGACCCTGAAAGTGTC	AGTCGTGGATTACTGTTAGAG	100
RNF44	NM_014901.4	*RNF44*	GAACTATGAGGCCCTCCTGAA	GAGCTGCTCTATGTCTGCTTT	85
UIMC1	NM_001199297.1	*UIMC1*	AACCCAAAGGAAAAAGGCCA	CCTGAAGGCTCCTGTTTCTG	111
NSD1	NM_022455.4	*NSD1*	GATGCTGGTCCCAAAGGAAA	GGTAAGTTCAGTGCCTGCTT	144
PRELID1	NM_013237.2	*PRELID1*	ACAGTGGCTGGACTGAAATC	GCAAGACCAAATTCCTGGACA	88

Primer ID	RefSeq ID	Gene	Forward primer (5’→ 3’)	Reverse primer (5’→ 3’)	Amplicon length (bp)
LMAN2	NM_006816.2	*LMAN2*	AACTGCATTGACATCACGGG	GTCATGATTGTCAGACAGGTCG	90
GRK6_1	NM_002082.3	*GRK6*	CAGAACGAGATGGTGGAGAC	TGCAGCAGTCCCTTTTTAGG	125
GRK6_2	NM_002082.3	*GRK6*	AGTAGGTTTGTAGTGAGCTTGG	ATGTGGTAGATGTGGAACTTGAG	107
PDLIM7	NM_005451.3	*PDLIM7*	CTGCTTCGTCTGTGCGATAT	GGCTCACACATGAGAGAAGG	106
DDX41_1	NM_016222.2	*DDX41*	TCCAGCACGTCATCAATTATGA	CACTGACTCATCACACGCTTT	131
DDX41_2	NM_016222.2	*DDX41*	GTCATCCAGGAGGTAGAATATGTG	TTCTCTGCAAAGATGAGTACAGG	101
TMED9	NM_017510.4	*TMED9*	GAGGCATGCTGAGAGTTCAC	CGTAGCTGCAACTCACTCAA	100
Reference genes					
GAPDH	NM_002046.3	*GAPDH*	GTCTCCTCTGACTTCAAC	TGTCATACCAGGAAATGAG	103
YWHAZ*	NM_145690.2	*YWHAZ*	ACTTTTGGTACATTGTGGCTTCAA	CCGCCAGGACAAACCAGTAT	94
ACTB	NM_000194.2	*ACTB*	CTGGAACGGTGAAGGTGACA	CGGCCACATTGTGAACTTTG	65

*[199]

3.2.5. Methylation-specific MLPA for 15q11 chromosomal region

Methylation-specific MLPA (MS-MLPA) was applied to study the methylation status and potential DNA copy-number changes in chromosome region 15q11 in the patient with matUPD15 from publication IV. MS-MLPA was performed using SALSA® MS-MLPA® probemix P028-B2 PWS/AS (MRC-Holland, Amsterdam, The Netherlands) according to the manufacturer's instructions. All reactions were performed in quadruplet. Fragments were separated by capillary electrophoresis using an ABI Prism® 3100 and analyzed using GeneScan® Analysis Software (Applied Biosystems).

SALSA® MS-MLPA® probemix P028-B2 contains 32 probes specific for sequences in or near the PWS/AS critical region of chromosome 15q11, which can be used to detect copy-number changes in this region. Five of these probes are specific for an imprinted sequence and contain a recognition site for the methylation sensitive HhaI enzyme. Four of the methylation-sensitive probes target the *SNRPN* gene and one probe targets the *NDN* locus. These five probes enable one to detect aberrant methylation patterns in the 15q11 locus, caused either by UPD or imprinting defects. To aid in the analysis, 14 reference probes targeting genes located outside the PWS/AS region are also included. In addition, three digestion control probes are included to indicate if complete digestion was accomplished by the HhaI enzyme during the methylation quantification reaction (MRC-Holland; product description version 44; 24-02-2012).

MLPA data analysis was performed using Coffalyser software (MRC-Holland): a spreadsheet-based program which is able to perform all data normalization steps, correct for signal sloping, and calculate the standard deviation (SD) for each sample. For copy-number analysis, the expected normalized values are 0.85–1.15 in the absence of any change, 0.35–0.65 in case of heterozygous deletion, and 1.35–1.55 in case of heterozygous duplication. For methylation analysis, an average expected methylation index for normal *SNRPN* methylation status is 0.5 (MS-MLPA® protocol version MSP-004).

4. RESULTS AND DISCUSSION

4.1. Application of chromosomal microarray analysis for the diagnosis of Estonian patients with developmental delay, multiple congenital anomalies and autism spectrum disorders (Publication I)

During four years – from January 2009 until December 2012 – a total of 1191 CMA tests were ordered at Tartu University Hospital, and in 1003 cases CMA was performed as the first-line cytogenetic test. Postnatal analyses comprised 95% of all referrals (90% patients and 5% family members), while the remaining 5% were prenatal analyses. The overall success rate was 99.5%. A repeat analysis was required in six cases: five did not pass the quality control, and in one mosaic uniparental disomy (UPD) case adjustment analysis was required.

On average 5–10 chromosomal aberrations were detected per investigated genome, thus making the interpretation of CMA findings the limiting factor that hampered the selection of truly causative variants. Generally, the chromosomal imbalances associated with well-established microdeletion/microduplication syndromes are not a matter of concern while abnormalities identified in genomic regions that are not yet associated with human diseases might present some difficulties. Based on the criteria described in section 3.2.1 and summarized in **Table 2**, all chromosomal aberrations found were assigned one of three labels: "pathogenic/likely pathogenic", "benign/likely benign" and "VUCS". Excluding benign findings, chromosomal aberrations were reported back to 298 (25%) patients with a total of 351 findings (1–3 per individual, with a size range from tens of kbs to entire chromosomes): 147 (42%) deletions, 106 (30%) duplications, 89 (25%) regions of LCSH (>5 Mb), and 9 (3%) aneuploidies. Six (2%) aberrations were in a mosaic state. Over 80% of the reported CNVs (not including regions of LCSH) were <5 Mb and would likely have been missed by traditional karyotyping and 39% were <1 Mb. If the two time periods of the study are examined separately – 2009–2010, when CMA was applied only for patients with otherwise normal routine genetic/metabolic test results; and 2011–2012, when CMA was performed as the first-tier cytogenetic test for patients with DD/ID, ASD or MCA – a difference in the number of reported results can be observed: 32% and 24%, respectively, which is due to patient selection criteria differences.

Out of 351 reported findings, 143 (41%) can be defined as pathogenic or likely pathogenic while the clinical significance of 208 (59%), most of which were LCSH, remained unknown at the time of reporting. However, 61 (30%) of the reported findings defined initially as VUCS can now be reclassified as benign/likely benign due to advances in the field of molecular clinical genetics and the addition of new entries to the publicly available databases. Similarly, Palmer et al. demonstrated a statistically significant difference in the

interpretation of CMA findings over the time, with an increase in putatively pathogenic CNVs [98]. In our study, we did not estimate the difference with regards to pathogenic/likely pathogenic findings. Nevertheless, it seems that the systematic re-evaluation of CMA results would be useful and could resolve the cases that remained without a specific diagnosis at the time of reporting or, on the contrary, exclude the possibility that the clinical phenotype of a patient is a consequence of one or more CNVs. Therefore, it should be emphasized in all CMA laboratory reports that the interpretation of CMA finding is based on current knowledge and may evolve over time [98].

In general, clinically relevant findings were detected in 11% of all patients analyzed, which is in concordance with previous reports and demonstrates once again the clinical utility of CMA in the diagnosis of patients with DD/ID, ASD or MCA [5, 200–202]. The diagnostic yields for the periods 2009–2010 and 2011–2012 were 15% and 10%, respectively, which is explained by a much higher phenotypic heterogeneity of the patient population analyzed in 2011–2012. This indicates that a more prominent phenotype of unknown etiology increases the chance of detecting clinically relevant aberrations using CMA. Despite this, the use of CMA as a first-line diagnostic test in a more heterogeneous population still improves the diagnostic yield at least two-fold compared to conventional karyotyping.

Verification studies using independent methods, such as qPCR, FISH, karyotyping or MLPA, were performed for more than half of the CNVs found (148 of 262) and confirmed the majority of these; yet four CNVs detected (10–970 kb; one reported initially as pathogenic and three VUCS) were found to be false-positives. It cannot be excluded that there may be more false-positives in our patient group. In two cases, the size of CNV was below the effective resolution level of the array platform used. Generally such small findings were not reported to the patients, however, in some cases an exception was made if a CNV spanned one or more known disease-causing genes. In all of these cases, confirmation studies were recommended. In case of a 970 kb duplication which has not been confirmed, there may be a problem with the qPCR method that provided the result due to either improper primers or difficulties with amplification of the particular genomic region. One could argue that there is no need to confirm CMA results because this technology has already been proven to be quite accurate and robust. However, our experience demonstrates that tiny CNVs with size ranges close to the effective resolution of the array platform still require confirmation by an independent method. Also, both karyotyping and FISH could provide additional insight if there is a need for precise characterization of a chromosomal aberration.

4.1.1. Clinically relevant CMA findings (Publication I)

Altogether, 143 CMA findings (42% of all reported findings) were assigned with the label "pathogenic or likely pathogenic" (for criteria used refer to

Table 2). Out of these, 106 aberrations were associated with known microdeletion or microduplication syndromes, or represented deletions that encompass a gene or a part of a gene implicated in human disease (most of those were <1 Mb) (**Table 6**). The most frequent genomic disorders found in our dataset were 15q13.3 microdeletion/microduplication syndrome (9 cases), 16p11.2 microdeletion/microduplication syndrome (5 cases), 1p36 microdeletion syndrome (4 cases), Silver-Russell/Beckwith-Wiedemann syndrome (4 cases, including one case of 11p15.5–15.4 UPD), PWS/AS (4 cases, including one case of maternal UPD 15). In general, the chromosomal imbalances associated with well-established microdeletion/microduplication syndromes are not difficult to interpret, and are usually unequivocal and straightforward, and thus ease the process of genetic counseling. We also discovered a relatively large number of aberrations in the recurrent microdeletion/microduplication loci that have well-established associations with abnormal phenotypes but also have incomplete penetrance and variable expressivity, e.g. the deletions and duplications in genomic loci 1q21.1 (4 cases), 16p13.1 (6 cases), and 15q11.2 (7 cases) responsible for increased susceptibility to neuropsychiatric and neurodevelopmental disorders and associated with variable dysmorphism [88, 203, 204]. Generally, CNVs with low effect size represent a challenge in genetic counseling because in addition to being enriched in individuals with various developmental disorders, they are also found in the normal population. Moreover, these CNVs are often inherited from a healthy or mildly affected parent. Remarkably, all 1q21.1 aberrations detected in our patient cohort were inherited, as well as 15q11.2 deletions for which inheritance analyses was carried out; the inheritance studies for 16p13.1 imbalances have not been performed. One can speculate that these variants are critical to, but not the sole determinants of, a given phenotype. Indeed, it has been demonstrated that the phenotypic variation of some genomic disorders (e.g., 15q11.2 deletion, 16p12.1 deletion, and 17p13.3 duplication) may be partially explained by the presence of additional large CNVs [205]. Still, genetic counseling and assessing the risk of recurrence is complicated due to an incomplete understanding of the mechanisms responsible for phenotypic variability and additionally hampered by the absence of specific guidelines. Recently, recommendations were published to aid in the genetic counseling regarding the deletion of susceptibility loci 15q11.2. These suggest that family-based studies with detailed genotyping and phenotyping of all family members, both carrier and non-carrier, could contribute to proper counseling [206].

Aneuploidies were discovered in eight (2%) of the patients (one trisomy 13, two monosomies X, two triple X syndromes, one Klinefelter syndrome, two XYY syndromes), which shows that aneuploidies, especially sex chromosome aneuploidies, are sometimes not easily recognizable from a clinical perspective.

Multiple LCSH distributed across the entire genome that obviously influence the phenotype by unmasking recessive mutations in disease-causing genes were observed in four cases, including two fetuses. The percentage of the genome

that is identical by descent (IBD) varied from 4% to 22%. However, this is clearly an underestimate because only those segments of homozygosity meeting a threshold of 5 Mb set by our laboratory were included in the calculation. The presence of multiple LCSH distributed across different chromosomes can indicate a familial relationship between the proband's parents and usually represents an unexpected finding [64, 106, 178]. Generally, a high percentage of the genome that is IBD (>10%) indicates a close parental relationship, and in this case the laboratory report should indicate that the results could be associated with possible consanguinity. However, a specific familial relationship or degree of parental relatedness cannot always be extrapolated from the inbreeding coefficient; therefore, speculations of a specific relationship must be avoided in laboratory reports [106].

Also, four cases of UPD associated with the patients' clinical phenotypes were found, including three mosaic cases: 4q31.3-q35.2 (50%), 11p15.5–p15.4 (50%) – Beckwith-Wiedemann syndrome, UPD 14, and maternal UPD 15 (50%) – PWS. All of these were detected through the presence of large LCSH restricted to a single chromosome, and remarkably, none of these cases was suspected to be caused by UPD based purely on clinical evaluation. This demonstrates that SNP-based CMA could be recommended to be included into a workflow of diagnosis of imprinting disorders.

One approximately 45 kb size deletion in 2q33.1 that was initially reported as likely pathogenic because it spans the disease-causing gene *NDUFB3* (OMIM #252010) and, in addition, overlaps with 2q33.1 deletion syndrome (OMIM #612313), was later found to be a false-positive result.

The remaining 20 aberrations classified as pathogenic or likely pathogenic did not overlap with any known syndrome but were large in size (at least several Mb) and located in gene-rich areas, which alone provides a reason to assume that they could be responsible for abnormal phenotypes. A summary of our clinically relevant findings, based on CMA analysis, with their corresponding reasons for referral are provided in **Table 7**.

Table 6. Aberrations that overlap with critical genomic regions for microdeletion and microduplication syndromes, or encompass genes implicated in human diseases [207].

Cytoband	Syndrome/Disease	OMIM No.	Gene(s)	No. of deletion cases	No. of duplication cases
1p36	1p36 microdeletion	607872		4	–
1q21.1	1q21.1 deletion/duplication[1]	612474/612475	Contiguous gene deletion syndrome, incl. *GJA5*	2	2
1q43-q44	megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome	603387	*AKT3*	1	–
2p16.3	2p16.3 deletion	614332	*NRXN1*	1	–
2q11.2	2q11.2 microdeletion	–	*LMAN2L, ARID5A*	1	–
2q31.2	2q31.2 deletion	612345	Contiguous gene deletion syndrome	1	–
2q37	2q37 microdeletion	600430	Contiguous gene deletion syndrome	1	–
3p25-pter	Distal 3p deletion	613792	Contiguous gene deletion syndrome	1	–
3p25.3	Von Hippel-Lindau syndrome	193300	*VHL*	1	–
3p13-p14	Waardenburg syndrome	193510	*MITF*	1	–
3q22.3	Blepharophimosis-ptosis-epicanthus inversus syndrome	110100	*FOXL2*	1	1
4p16.3	Wolf-Hirschhorn Syndrome	194190	Contiguous gene deletion syndrome	1	2
4q22.1	Parkinson disease	168601	*SNCA*	–	2
5p15.2	Cri-du-Chat syndrome	123450	Contiguous gene deletion syndrome, incl. *TERT*	2	–
5p15.2	Mental retardation in Cri-du-Chat syndrome	123450	*CTNND2*	1	–

Cytoband	Syndrome/Disease	OMIM No.	Gene(s)	No. of deletion cases	No. of duplication cases
5q35.2-q35.3	Sotos syndrome/5q35 microduplication	117550/-	*NSD1*	2[2]	1
6q25.1-q25.2	Emery-Dreifuss muscular dystrophy 4, autosomal dominant	612998	SYNE1	1	–
7p21.1	Saethre-Chotzen syndrome	101400	*TWIST1*	1	–
7p14.1	Greig cephalopolysyndactyly syndrome/Pallister-Hall syndrome	175700/146510	GLI3	1	–
7q11.23	Williams-Beuren syndrome	609757	Contiguous gene deletion syndrome, incl. *ELN*	1	–
7q21.2-q21.3	Split-hand/foot malformation 1 with sensorineural hearing loss	220600	*DLX5*	1	–
7q36.3	Polydactyly, preaxial II	174500	*LMBR1*	1	–
8q24.13	Spastic paraplegia 8, autosomal dominant	603563	*KIAA0196*	1	–
10q23	Juvenile polyposis syndrome + 10q23 deletion	174900/612242	*NRG3, GRID1, PTEN, BMPR1A*	1	–
10q26	10q26 deletion	609625	Contigous gene deletion syndrome	1	1
11p15.5	Beckwith-Wiedemann/Silver-Russell syndrome	130650/180860	Contigous gene deletion syndrome, incl. *CDKN1C, H19, LIT1*	3	1
11q23	Jacobsen syndrome/Thrombocytopenia, Paris-Trousseau type	147791/188025	Contigous gene deletion syndrome	1	–
12p12.1	DD, language delay, behavioral problems	–	*SOX5*	1	–
15q11.2[4]	Prader-Willi/Angelman syndrome (Type 1)	176270/105830	*NDN, SNRPN, UBE3A*	1	1

Cytoband	Syndrome/Disease	OMIM No.	Gene(s)	No. of deletion cases	No. of duplication cases
15q11.2	Prader-Willi syndrome/Angelman syndrome (Type 2)	176270/105830	*NDN, SNRPN, UBE3A*	1	–
15q11.2	15q11.2 microdeletion/microduplication[1]			6	1
15q13.3	15q13.3 microdeletion/microduplication[3]	612001	Contigous gene deletion syndrome, incl. *CHRNA7*	7	2
16p11.2	16p11.2 microdeletion/microduplication	611913/614671	Contigous gene deletion syndrome	4	1
16p12.1	16p12.1 microdeletion[1]	136570	Contigous gene deletion syndrome	2	–
16p13.11	16p13.11 microdeletion/microduplication[1]	–	Contiguous gene deletion syndrome, incl. *MYH11*	1	5
16p13.2	Epilepsy with neurodevelopmental defects	613971	*GRIN2A*	1	–
17p13.3	17p13.3 distal deletion	–	*YWHAE*	2	–
17p12	Hereditary neuropathy with liability to pressure palsies	162500	*PMP22*	2	–
17p11.2	Smith-Magenis syndrome	182290	*RAI1*	1	–
17q11.2	Neurofibromatosis I	162200	*NF1*	2	–
17q21.31	Koolen-De Vries syndrome	610443	Contiguous gene deletion/duplication syndrome, incl. *MAPT*	1	–
18p	Chromosome 18p deletion syndrome	146390	Contigous gene deletion syndrome	1	–
18p11.31	Holoprosencephaly 4	142946	*TGIF*	2	–

Cytoband	Syndrome/Disease	OMIM No.	Gene(s)	No. of deletion cases	No. of duplication cases
18q22.3-q23	Congenital aural atresia	607842	*TSHZ1*	1	–
22q11.2	DiGeorge/Velocardiofacial/Chromosome 22q11.2 duplication syndrome	188400/192430/ 608363	Contiguous gene deletion syndrome, incl. *TBX1* and *COMT*	3	1
22q13	Phelan-McDermid syndrome	606232	Contiguous gene deletion/duplication syndrome, incl. *SHANK3*	3	–
Xp22.31	Ichthyosis	308100	*STS*	1	–
Xp21.3-p21.2	X-linked mental retardation	300143	*IL1RAPL1*	3	–
Xp21.1	Duchenne muscular dystrophy	310200	*DMD*	3	–
Xq28	Rett syndrome	312750	*MECP2*	–	1
Yq11.21-q11.23	Spermatogenic failure	415000	*USP9Y, DB9*	1	–

[1] Susceptibility locus; [2]One of the patients with a deletion of exons 3–8 of *NSD1* did not display a clinical phenotype of Sotos syndrome, but rather a phenotype of 5q35 microduplication; [3]Duplication represents a susceptibility locus; [4]In one case maternal UPD was diagnosed.

Table 7. Summary of clinical and molecular data of patients with potentially pathogenic findings not overlapping with known genetic syndromes.

Aberration [hg19]	Length (Mb)	Confirmation/inheritance studies	RefSeq genes	OMIM genes	Gender, age	Referral reason
1p36.33(781,258-1,965,445)x1	1.18	Confirmed/*de novo*	70	31	M, 6y	ID, speech delay, febrile seizures
1p13p11.2(112,773,948-121,184,898)x1	8.41	Confirmed/*de novo*	85	54	M, 10y	Severe ID, dysmorphism
1q22q23.1(155,424,065-157,907,476)x3	2.48	Not performed	76	44	M, 5y	DD, ID
4q11q13.1(52,423,081-60,940,717)x3	8.52	Karyotype has been done previously	47	32	M, 1y	Karyotype 47,XY+mar. Refinement study
7q11.22(67,037,628-71,954,937)x3	4.92	Confirmed/*de novo*	5	3	F, 1y	DD, dysmorphism
8p23.3p23.1(1-6,961,460)x1	6.96	Confirmed	44	15	M, 0y	DD, hypotonia, dysmorphism
8p23.3p23.1(1-8,615,874)x1	8.62	Confirmed	87	21	F, 12y	DD, epilepsy
8p23.3p23.2(1-6,802,066)x3	6.8	Confirmed	31	12	F, 0y	DD, heart anomaly, seizures
8p23.1p11.1(12,595,570-43,765,570)x3	31.17	Confirmed	226	150	M, 0y	DD, hypotonia, dysmorphism
8p11.21q11.21(40,540,918-49,805,796)x3	9.26	Confirmed (marker chromosome)	45	28	F, 4y	DD, speech delay
12p11.21p13.31(9,758,340-31,359,210)x3	21.6	Confirmed	181	113	M, 6y	DD, microanomalies
12q11q12(38,094,338-42,411,344)x3	4.32	Karyotype has been done previously	10	7	M, 4y	Karyotype 47,XYY,+mar. Refinement study

Aberration [hg19]	Length (Mb)	Confirmation/inheritance studies	RefSeq genes	OMIM genes	Gender, age	Referral reason
13q33.1q34(101,733,731-115,169,878)x1	13.44	Confirmed	85	38	F, 4y	DD, speech delay
13q33.3q34(108,360,217-115,169,878)x1	6.81	Not performed	63	26	M, 9y	ID, seizures, diabetes
14q21.1q21.2(38,419,315-46,613,386)x1	8.19	Not performed	25	13	M, 0y	DD, microanomalies, hypospadias
17q25.3(78,175,854-81,047,565)x3	2.87	Confirmed (marker chromosome)	95	52	M, 4y	DD, microanomalies
18p11.31p11.31(212,756-5,529,007)x1	5.32	Not performed	36	17	F, 6y	DD
18q21.32q22.1(58,105,704-66,013,818)x3	7.91	Karyotype has been done previously	33	21	M, 21y	Karyotype 46,XY,–18,+r(18). Refinement study
20q11.21q11.22(29,843,779-32,202,779)x3	2.36	Confirmed	52	27	M, 4y	Speech delay, microanomalies, behavioural problems
20q13.12q13.33(46,262,908-62,909,908)x3	16.65	Confirmed	211	111	M, 0y	DD, epicanthus, telecanthus

4.1.1.1. Deletions in 7q31 encompassing *FOXP2* cause speech and language disorder (Publication II)

The patients described in Publication II were enrolled into a study of ID in Estonian patients using SNP genotyping microarrays [8]. This study is the first to use CMA for the purpose of studying CNVs associated with ID in Estonian patients and the general population.

Family 1

DNA samples of six individuals from F1 (I-1, I-2, II-1, II-3, III-1 and III-2) (**Figure 4A**) were studied using a whole-genome genotyping array. In patient F1-III-2, a ~8.3 Mb deletion on 7q31.1-q31.31 (arr[hg18] 7q31.1q31.31 (111,784,188–120,139,346)x1) was detected, and this was found to have been inherited from her affected mother (F1-II-3). The genotyping data showed that the mother's (F1-II-3) rearrangement was located on the paternally derived chromosome. A further breakpoint refinement mapped the proximal breakpoint at genomic position 7:111,781,517 within a short interspersed nuclear element (SINE) repeat (AluSx), and the distal breakpoint at position 7:120,142,536 within a short palindromic sequence, while any repetitive sequences were missing in those regions.

Because the affected daughter had previously been suspected of having cystic fibrosis caused by mutations in the *CFTR* gene located within the deleted region (7q31.2), an analysis for the detection of *CFTR* mutations was performed but did not reveal any mutations.

Family 2

Three individuals from F2 (I-1, II-5 and III-2) (**Figure 4B**) were studied using SNP-based CMA. A DNA copy-number analysis revealed a ~6.5 Mb deletion in 7q31.1-q31.2 (arr[hg18] 7q31.1q31.2(108,290,244-114,759,023)x1) in proband (F2-III-2) and her maternal aunt (F2-II-5). DNA samples of the patient's parents were not available for testing, but the family anamnesis and the analysis of the genotyping data of three individuals (I-1, II-5 and III-2) from this family made it possible to establish the pattern of inheritance. Presumably, the proband inherited the 7q31 deletion from her mother, and her aunt carried the deletion on a paternally inherited chromosome. The fine-mapping of the deletion boundaries revealed that the proximal breakpoint occurred at genomic position 7:108,278,292 and the distal at position 7:114,819,017. As in the case of family 1, the proximal breakpoint was located within an AluSX element, and the distal breakpoint within a short palindromic sequence.

Both deletions included the *FOXP2* gene implicated previously in autosomal dominant CAS [137]. There are at least 27 previously published cases of interstitial deletions encompassing 7q31, including one prenatally diagnosed case [133–136, 208–221]. It has been suggested that patients with chromosomal deletions involving 7q31 may define a new contiguous gene deletion syndrome

characterized by aspects of verbal dyspraxia [133]. However, excepting impairments in speech and language, no common features are consistently observed in patients carrying a deletion of the *FOXP2* locus [132]. Moreover, 7q31 deletions seem to be nonrecurrent with scattered breakpoints. Based on the sequencing data we suggest that nonhomologous end-joining (NHEJ) may be responsible for the formation of the 7q31 deletions in our patients because no substantial homologies were detected between the DNA sequences located across the breakpoints. This hypothesis is supported by the fact that in both cases the distal breakpoints were located within short palindromic sequences, which are known to be able to induce a curvature in the DNA molecule, predisposing it to recombination [222, 223]. The patients with intragenic point mutations as well as a recently reported patient with a small intragenic deletion confirm that the haploinsufficiency of *FOXP2* is responsible for formation of the core phenotype observed in patients with a 7q31 deletion [135]. The additional problems or CA could therefore be attributed to the other genes located within the deleted region or elsewhere in genome. For example, the leading cause for consultation by a clinical geneticist in patient F1-III-2 was congenital anomalies of the urinary system and facial dysmorphism. None of the reported patients with a 7q31 deletion have been reported to have any renal problems. No renal problems were observed in the proband's mother who carries the same deletion. To the best of our knowledge, none of the genes encompassed by the deletion could contribute to the development of nephrogenesis or other parts of the urinary system, and it appears that some other reason not detected in this study is the underlying cause.

In addition, the families presented in this work provided an opportunity to estimate the phenotypic effect of the same deletion of opposite parental source. The analysis of genotyping data in our study revealed that in both families described, the probands' deletions were of maternal origin, while the mothers themselves carried the rearrangements on paternally-inherited chromosomes. It has previously been proposed that the *FOXP2* locus may be differentially regulated in a parent-of-origin manner and that the absence of paternal *FOXP2* could cause a more severe phenotype with CAS, whereas the absence of the maternal allele may result in a milder phenotype [132]. However, the detailed phenotype descriptions and speech and language assessments of our patients demonstrated that there is no significant difference in the severity of each observed phenotype. Moreover, this conclusion is further supported by the work of Thomas et al. who demonstrated biallelic expression of *FOXP2* which disproves the imprinting hypothesis. They tested the imprinting status of *FOXP2* by estimating its allelic expression using transcribed SNPs and found that *FOXP2* was biallelically expressed in fetal brain and fetal liver tissues, thus verifying that it is not imprinted in these tissues at least [224]. The parent-of-origin specific phenotypic effect reported by Feuk et al. could therefore be coincidental.

4.1.1.2. Reciprocal deletions and duplications in 5q35.2-q35.3 cause mirror phenotypes related to growth (Publication III and unpublished data)

The patient presented in Publication III was tested by CMA in 2010 due to a proportionally short stature and facial dysmorphism of unknown etiology. CMA revealed a duplication on the long arm of chromosome 5 (arr[hg18] 5q35.2q35.3(174,950,741–176,979,615)x3), which was confirmed by FISH analysis (46,XX.ish dup(5)(q35q35)(NSD1++)). The duplication was not detected in either the patient's mother or sister. The father was not available for testing.

The detected 5q35.2-q35.3 duplication is reciprocal to the common Sos deletion and has been proposed to cause a specific syndrome with a recognizable phenotype that is clinically opposite to Sos with regard to growth and head circumference [159, 169–173]. Still, although the 5q35.2-q35.3 duplication syndrome is often referred to in the literature as "reversed Sos", this term does not seem to be suitable because of the absence of reversed facial phenotype and the presence of unspecific findings such as short stature and microcephaly [173]. The characteristic features of the condition include short or low normal stature, microcephaly or low normal OFC, learning disability or mild to moderate ID which may be associated with behavioral problems, and distinctive facial features comprising periorbital fullness, short palpebral fissures, long nose, long or bulbous nasal tip, flat philtrum and thin upper lip. Delayed bone age was also evident in most cases where X-ray investigation was performed. Minor digital anomalies, such as clinodactyly, brachydactyly, syndactyly, and polydactyly, were described in some cases. No major malformations are associated with 5q35.2-q35.3 duplications. [159, 172, 173]. In general, phenotypic characteristics related to growth seem to be remarkably opposite in cases of 5q35.2-q35.3 deletion and duplication. However, intra- and interfamilial variation is observed in terms of body measurements, intelligence and dysmorphic features, which point to the possibility that the duplication may be underdiagnosed [173].

Generally, the *NSD1* gene located in this region is the only gene known to be associated with Sos. Subsequently, it has been proposed that increased *NSD1*-dosage should cause the reduced growth parameters observed in patients with duplication. Recent publications have described additional patients with a 5q35.2-q35.3 duplication which allow for a more precise determination of the corresponding phenotype and narrows down the region of overlap [172, 173]. Combining the phenotypic and molecular data from these two studies it could be hypothesized that the duplication of *NSD1* alone could be sufficient to cause growth retardation. At the same time, the phenotype of the Sos patients carrying 5q35 microdeletions is more variable compared with patients that carry *NSD1* mutations, which suggests a possible dosage-sensitivity of genes neighboring *NSD1* and their possible effect on phenotype. For example, the digital

anomalies described in a subset of patients may be attributable to the duplication of *PDLIM7* (PDZ and LIM domain 7) [172]. It was demonstrated that when present in excess, Pdlim7 lowers nuclear levels of Tbx5 thus affecting Tbx5 target gene expression [225]. In turn, haploinsufficiency of *TBX5* causes Holt-Oram syndrome (OMIM #142900) which is characterized by thumb anomalies and heart defects. However, digital anomalies are not observed in all patients with a 5q35.2-q35.3 duplication encompassing *PDLIM7* [173]. Heart defects are also an infrequent finding in those patients [172].

So far, there have been no functional studies on the influence of 5q35.3 structural changes on human growth. Obviously, gene expression studies could reveal information regarding the mechanisms that underlie the phenotypic effect of reciprocal deletions and duplication of this particular locus. Using RT-qPCR, we analyzed the expression profiles of 11 genes from the Sos critical region (*HIGD2A, FAF2, RNF44, UIMC1, NSD1, PRELID1, LMAN2, GRK6, PDLIM7, DDX41, TMED9*) and flanking regions (*HNRNPAB, SFXN1,* and *CLK4*), and two known NSD1 target genes (*HSD17B10* and *ZMYM3*) in three patients with a 5q35.2-q35.3 duplication (**Figure 5**). All genes excepting one (*UIMC1)* within the duplicated region showed significant overexpression. The most significant changes in expression were observed in *HIGD2A, PRELID1, GRK6, DDX41,* and *TMED9*. Gene expression in the flanking region was also perturbed. Both NSD1 target genes studied also show overexpression (Unpublished data). However, because of the small number of patients analyzed as well as the use of RNA samples from various sources, such as whole-blood and lymphoblastoid cell line, we cannot make any statistically significant conclusions. These results only hint at a tendency that requires further confirmation. Additionally, it is not clear whether the observed overexpression manifests itself also at the protein level and subsequently on the phenotype.

Interestingly, locus 5q35.2-q35.3 is not the only known genomic region where reciprocal deletions and duplications are associated with opposite phenotypes involving body parameters. For instance, proximal 16p11.2 microdeletion and microduplication display mirror phenotypes with regards to both body mass index and head circumference, while a reciprocal genomic events in the distal 1q21.1 locus have opposite effects on head size [89, 162]. While major drivers of mirror neuroanatomical phenotypes associated with the 16p11.2 CNVs have been identified, the genes implicated in 1q21.1 micro-deletion/microduplication syndrome remain to be elucidated. Manipulation of zebrafish embryos has been proposed as an attractive approach to discover dosage-sensitive genes responsible for mirror anatomical phenotypes that are detectable during early development [226]. In addition, there are several examples of single-gene disorders in which activating mutations and haploinsufficiency or dominant negative mutations have opposite phenotypic effects related to growth, e.g. the *AKT* genes and *FGFR3* (reviewed in [172]). This supports the idea that disturbance of *NSD1* alone can manifest itself in

opposite growth parameters, while other genes in the region are responsible for other features.

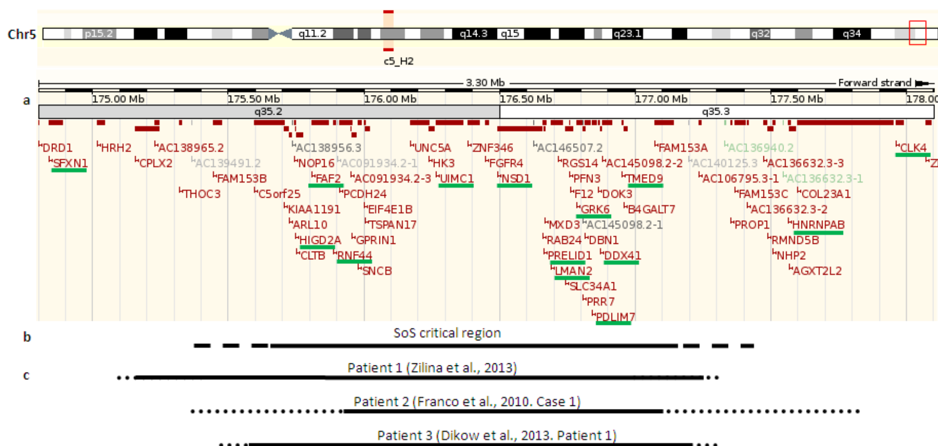


Figure 5. Scale representation of the 5q35.2-q35.3 region (http://ensembl.org; NCBI36): (a) location of Ensembl genes in the region. Genes selected for gene expression analysis are underlined; (b) location of Sotos critical region and the flanking low-copy repeats; (c) duplicated fragments in Patients 1–3. Solid lines indicate the minimum boundaries and dashed lines indicate the maximum boundaries of duplication.

4.1.1.3. Mosaic matUPD15 in a patient with partial Prader-Willi syndrome phenotype as an example of the utility of SNP-based arrays in detecting uniparental disomy and mosaicism (Publication IV)

The patient was referred for CMA analysis in 2011 by a child neurologist due to being overweight, having muscular hypotonia, and psychiatric problems. There was no reason to suspect PWS in the patient based solely on the phenotype.

Chromosomal microarray analysis and regular karyotyping

CMA using an SNP array revealed mosaic loss of heterozygosity of the entire chromosome 15. The percentage of abnormal cells was estimated to be 55–60% in peripheral lymphocytes [43]. Karyotyping from peripheral blood lymphocytes (100 mitoses) and from skin fibroblasts (30 mitoses) showed normal male karyotype 46,XY. No monosomy or trisomy was detected by CMA, which confirmed the G-banding results thus it was assumed that the patient was mosaic for uniparental isodisomy 15, and had a normal biparental cell line.

Methylation analysis of 15q11-q13 region

Non-quantitative MS-PCR within the PWS/AS region showed a normal result; *SNRPN* alleles from both parents were present. Quantitative MS-MLPA

analysis showed hypermethylation of *SNRPN* and *NDN* loci in the PWS/AS critical region. The *SNRPN* locus was analyzed using four MS-MLPA probes and provided average methylation quotients of 0.8 (in normal controls 0.45–0.6), and the *NDN* locus was analyzed using one probe and provided methylation quotients of 0.69 (mean of normal controls 0.38). In a patient previously diagnosed with PWS and a typical 15q11-q13 deletion, the mean methylation quotient for the *SNRPN* locus was 1.0 and for the *NDN* locus was 0.68. There were no copy number variations within the 15q11-q13 region in our patient. These results confirm that the patient has a mosaic UPD within chromosome 15 (maternal) in an approximate proportion of 50–55%.

Taking into consideration the MS-MLPA results, and the presence of some PWS features in the patient, we conclude that he had mild PWS caused by mosaic matUPD15, although the patient's parents were not enrolled in the study. Using the results of both CMA and MS-MLPA, we estimate that the proportion of cells with the abnormality is within the range of 50–60% in whole blood. However, we cannot predict the proportion of mosaicism in other tissues.

According to CMA and karyotyping, no trisomic or monosomic cells were present. Both peripheral blood lymphocytes and skin fibroblasts were studied by karyotyping and revealed no signs of mosaic aneuploidy. To the best of our knowledge, only two PWS cases with mosaic maternal isodisomy 15 and without trisomic/monosomic cell lines have been reported to date [227–229]. In addition, one patient was recently described who is mosaic for isochromosome 15q associated with matUPD15 and a normal cell line, [46,XX/45,i(15)(q10)]. This patient displays a minimal PWS-like phenotype [230]. Usually, the presence of additional cell lines as well as determining the type of UPD (hetero- or isodisomy) enables one to delineate the mechanism responsible for the formation of UPD, however, it should be noted that this is impossible in the case of pure UPD without additional cell lines. In general, most known UPD mechanisms can lead to mosaic formation (**Figure 6**). In turn, understanding the mechanism underlying the formation of UPD can be useful for prognosis estimation and counseling.

Our patient is similar to patient No 1 described by Izumi et al. [228], who was found to be mosaic for maternal isodisomy 15 and had heterodisomic cell lines. In that case, the percentage of abnormal cells was up to 85–90% which apparently leads to a more prominent PWS phenotype compared with our patient who had 50–60% abnormal cells. Post-fertilization mitotic error was proposed to be responsible for mosaic UPD in that patient and this is very likely the cause of UPD in our patient as well. The signs that point to possible post-fertilization mitotic error are the type of UPD (isodisomy) and the absence of any degree of aneuploidy (**Figure 6**). In the two other reported cases [227, 230], the mechanism of UPD formation was different. Although, in the case reported by Horsthemke et al. [227] no numerical or structural chromosome aberrations were detected by karyotyping and FISH, the microsatellite analysis revealed three alleles at five of the nine loci tested, which indicates that trisomy rescue

should be responsible for the UPD (**Figure 6**). Moreover, regarding the severe ID of the patient, the authors speculate that the presence of trisomic cell lines might simply have remained undetected because of their possible low level, or perhaps trisomy may also be present in unstudied body tissues [227]. In the case described by Wang et al. [230], the level of the cells with i(15)(q10) associated with maternal UPD15 was 30%, while the remaining 70% were cells with a normal female karyotype. The patient herself had an abbreviated PWS phenotype, including morbid obesity, small hands and feet, short stature, and behavioral problems, however, lacked significant facial dysmorphisms, hypotonia, and cognitive delays. Normal cognitive functions are obviously a consequence of the comparatively low-level mosaicism for the abnormality. Two possible UPD mechanisms were proposed in this case: 1) a mitotic error in a normal diploid cell followed by trisomy rescue; or 2) a trisomy rescue followed by monosomy rescue [230].

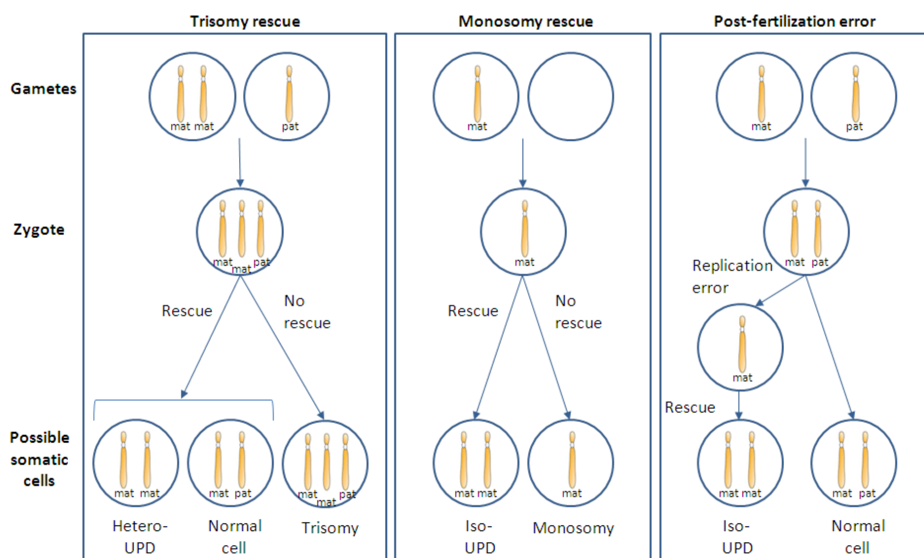


Figure 6. A schematic representation of mechanisms that lead to mosaic uniparental disomy. The three panels are: trisomy rescue, monosomy rescue, and post-fertilization mitotic error. The type of UPD (iso- or heterodisomy) and the presence of additional cell lines provide clues when determining the mechanism responsible for UPD formation in a particular patient [193].

It is remarkable that all cases described to date have been diagnosed using a panel of different methods, including conventional cytogenetic analysis (karyotyping, FISH), a molecular cytogenetic tool (CMA) and molecular genetic methods (microsatellite analysis, methylation analysis) in various combinations. Obviously, the complex nature of those aberrations and the underlying mechanisms would not have been elucidated if only one method had been applied. Generally, a broad panel of molecular tests for imprinting

disorders is available, however, none of these detects all known mutations and epimutations [185]. A combination of different methods (e.g. CMA with SNP arrays, MS-PCR, MS-MLPA) is recommended to differentiate between UPD, deletion/duplication, and/or methylation disturbances [185, 188]. The use of pure qualitative methylation analyses should be avoided because they are not informative in mosaic UPD cases. We propose that a combination of SNP-based CMA and quantitative methylation analysis is appropriate for diagnosing patients with putative imprinting disorders.

Currently, the majority of cases with mosaic matUPD15 may remain undetected because of an abbreviated PWS phenotype observed in the patients. Arriving to the correct diagnosis can also be complicated by a lack of diagnostic tools in some diagnostic laboratories, technical difficulties associated with their application and difficulties associated with interpretation of results. Apparently, the wider application of SNP arrays and quantitative methylation analysis methods in molecular diagnostics would clarify whether mosaic UPD is truly rare or is simply underdiagnosed in PWS patients [228].

4.1.2. Findings with uncertain clinical significance (Publication I)

The clinical relevance of 144 (41%) reported findings remains unclear, including 64 deletions/duplications and 80 regions of LCSH. Most CNVs were <1 Mb and have not been previously implicated in human diseases. One of the primary tests recommended for the estimation of the pathogenicity of VUCS's (primarily deletions and duplications) are inheritance studies, although it is often imprudent to attribute clinical significance based on the inheritance pattern of a CNV within a single family [97]. In this study, inheritance analyses were completed for about half of the deletions/duplications cases with uncertain clinical relevance, and only three imbalances out of 28 appeared to be *de novo*. Still, the pathogenicity of inherited CNVs cannot be excluded before more information on those genomic regions is available, because a growing number of recurrent CNVs display variable penetrance or expressivity and may confer susceptibility or risk, rather than being directly causative [78, 231]. In addition, it should be kept in mind that parentally segregated CNVs could contribute to a proband's phenotype through epigenetic effects, or by unmasking a recessive mutations on a non-deleted allele [97, 202]. The situation with *de novo* mutations is also not straightforward. Although the "*de novo*" status is usually taken as evidence supporting pathogenicity, it has been demonstrated that many regions of the genome have significantly elevated mutation rates, and some CNVs may indeed be *de novo* mutations yet do not infer clinical significance [232].

According to laboratory policy, stretches of homozygosity larger than 5 Mb were reported. However, in most cases this turned out to be diagnostically unhelpful because the vast majority of reported LCSH were classified as VUCS. Single LCSH events, especially smaller ones, are generally more difficult to

interpret. Most detected cases of LCSH likely represent regions of suppressed recombination or linkage disequilibrium, although they may potentially be associated with recessive diseases. The genomic content of the region should be carefully evaluated with regards to the patient's clinical problems, which assumes a close collaboration between clinical and laboratory staff [231]. Subsequently, the confirmation of the pathogenicity of such chromosomal abnormalities requires sequencing of the candidate gene or genes of interest. Nevertheless, most LCSH detected in our patients were classified as VUCS because it was impossible to establish a link between a phenotype and the CMA findings. The most promising finding was a 12 Mb homozygosity stretch in 3q13.13-q21.1 that encompassed the *CASR* gene implicated in epilepsy, and this was found to correlate well with the patient's phenotype [233]. However, Sanger sequencing of *CASR* was not performed in this case.

Even when LCSH results are excluded, the proportion of VUCS among all CMA findings remains quite high (24%). It seems that our ability to detect CNVs has far outpaced our ability to understand their role in disease [3]. Proper treatment of most VUCS cases typically requires additional tests that are time-consuming and expensive, and furthermore require recurrent counseling. All of these increase the cost of providing health care. Due to cost restrictions, the number of confirmation studies performed on Estonian patients has been limited over time. Currently, only family studies are performed and small CNVs with a size close to the array platform resolution are confirmed to exclude false-positives. It should also be emphasized that, although the interpretation of CMA results tends to change over time, CMA findings of unknown clinical relevance for which proper counseling, recurrence risk assessment and management cannot be offered, often cause stress and uncertainty in patients and their families, and sometimes even mutual intrafamilial accusations. Therefore, the referring pediatricians or neurologists should be aware of the possibility that CMA provides results which are often random or difficult to interpret. Open-access databases of clinically relevant (e.g., DECIPHER) as well as nonpathogenic CNVs (e.g., DGV) are extremely helpful for interpreting CMA results, therefore it is very important that as many centers as possible contribute to the development and completion of these resources. It should be mentioned that due to our expanding knowledge, including the addition of new entries to the publicly available databases during 2009–2012, a significantly large portion (61 out of 351) of chromosomal aberrations reported to our patients that were initially labeled "VUCS" were re-categorized as benign or likely benign. Generally, the primary duty of monitoring the medical literature and relevant databases that is required to re-categorize VUCS lies with the physician with an ongoing patient relationship [97].

In turn, the amount of VUCS can be minimized by choosing an array platform that strikes a balance between sensitivity and specificity. Obviously, there is no need for maximum resolution in a genome-wide clinical test, because this is accompanied by an increase in the number of findings of uncertain

clinical relevance. It is recommended to use a resolution of ~400 kb throughout the genome with probe enrichment in regions of known clinical relevance. This level enables one to reliably identify all known recurrent microdeletion and microduplication syndromes and most nonrecurrent imbalances that are unequivocally pathogenic [5, 94]. In addition, one can choose between two possible options: SNP arrays and aCGH, which are both highly efficient tools to conduct both research and clinical diagnostics. However, and in contrast with aCGH, the genotype information provided by SNP-based arrays allows for the recognition of copy-number-neutral events such as LCSH. It should be discussed whether a particular diagnostic centre is interested in detection of this kind of aberration, because they usually represent an issue of concern with regards to interpretation and counseling. Also, the genotype data obtained by SNP-arrays is useful when knowledge of the parental origin of an aberration is critical and must be determined, although in this case the patient and both parents (a trio) should be analyzed. In addition, SNP arrays are able to distinguish between a mitotic and meiotic origin of mosaicism which also may be required for appropriate counseling.

4.1.3. Estonian experience of chromosomal microarray analysis application in prenatal diagnosis (Publication I)

During 2009–2012, CMA with fetal DNA was performed in 60 cases, eight of which were ordered after the termination of the pregnancy. Indications that lead to prenatal CMA testing are presented in **Table 8**. In most cases, array analysis was performed simultaneously with karyotyping to enable better characterization of potential CMA findings and to detect aberrations that could have been missed using CMA. In eight cases an abnormal result was reported (**Table 9**).

Table 8. Prenatal CMA testing in Estonia during 2009–2012 (including fetuses analyzed after the termination of pregnancy) [207].

Indication for prenatal diagnosis	**Number of cases (%)**
Familial balanced rearrangement	18 (30%)
Anomaly on ultrasonography	13 (22%)
Termination of pregnancy due to abnormal fetus	8 (13%)
Positive triple test	5 (8%)
Isolated abnormal nuchal translucency	5 (8%)
Other child(ren) with chromosomal disease	4 (7%)
Other child or parent with unspecified genetic pathology	3 (5%)
Unspecified	3 (5%)
Recurrent spontaneous abortions	1 (2%)
Total	**60 (100%)**

In case 1, a duplication encompassing exons 45–51 of the *DMD* gene was detected in a male fetus (46,XY) and was confirmed by MLPA analysis using the SALSA MLPA P034-A2 and P035-A2 probe mix (MRC-Holland). The mother did not carry the duplication and the pregnancy was terminated after counseling. However, later it was found that the father was a carrier of Xp21.1 duplication. Because chromosome X generally cannot be transferred to male offspring through the paternal line, the duplicated segment is likely to be inserted into some other chromosome, however, this theory has not been tested. Potentially, there is still a possibility that the male fetus obtained an X chromosome from his father resulting in an XXY karyotype at an early stage of development followed by trisomy rescue and elimination of the maternal X chromosome [234]. We attempted to verify this case, but found no supporting evidence.

CMA was carried out in case 2 due to recurrent spontaneous abortions of unknown etiology in the family. The analysis was performed after the termination of the pregnancy and revealed a 5.6 Mb LCSH on chromosome 8; however, its association with clinical problems remains uncertain.

In two cases (3 and 8) multiple regions of LCSH distributed across the entire fetal genome were discovered and point to possible parental relationship (the percentage of genome that is IBD was 6% and 20%, respectively).

In case 4, a low-level mosaic trisomy 7 (~13% and ~10%, respectively) was detected by G-banding and CMA using amniotic fluid cell culture. Although most cases with this chromosomal abnormality have no symptoms or only subtle clinical symptoms, a maternal UPD7 that is strongly associated with severe growth restriction cannot be excluded. Because some symptoms were observed using ultrasonography, additional amniocentesis was performed. FISH analysis revealed the presence of an additional chromosome 7 in 5% of the cells, while G-banding revealed a normal karyotype. However, a normal female was born at term with normal birth weight and length.

In case 5, an approximately 3 Mb deletion in 7p14.1-p13 was found, thus disrupting the *GLI3* gene associated with Greig cephalopolysyndactyly syndrome (OMIM #175700), and was concordant with the fetal dysmorphic phenotype.

Cases 6 and 7 were referred due to familial balanced rearrangements. In case 6, a terminal duplication of 4p (14 Mb) and a terminal deletion of 4q (2 Mb) were detected and treated as pathogenic due to their size and pregnancy was terminated. In case 7, the fetus was found to inherit an inv(10)(p11.2q21.2) from his mother, and no CNVs in inversion adjacent regions or elsewhere in the genome were detected by CMA. However, a 5.5 Mb LCSH with unclear clinical relevance was identified. The outcome of this pregnancy is not known.

Interpretation of CMA results is particularly challenging in prenatal testing where limited information of the fetal phenotype is compounded with time pressure. CMA ability to produce results of uncertain clinical relevance and incidental findings present grounds for many ethical debates. However, the

issue of VUCS is not unique to CMA, because conventional karyotyping can also reveal findings of uncertain clinical impact, such as supernumerary marker chromosomes or apparently balanced *de novo* aberrations [119]. As with postnatal diagnostics, the choice of an appropriate array platform is critical in terms of decreasing the number of prenatal VUCS. Several possibilities exist and include the use of the same platforms for both prenatal and postnatal settings, the use of microarrays with reduced average resolution for prenatal diagnosis compared with postnatal diagnosis, and the use of targeted microarrays [116, 119, 120]. Some authors propose that targeted platforms developed specifically for prenatal settings should be used to avoid VUCS and thereby facilitate autonomous reproductive choice for pregnant women (and their partners). In turn, undirected microarrays should be reserved for research purposes only due to their lack of clinical validity, because generation of VUCS serves a role in research and does not contribute to the aims of autonomous reproductive decision-making [235]. However, there is a lack of consent forms or specific guidelines regarding the choice of array platform. Overall, targeted design seems to be a convenient option, however, the design would need to be constantly updated to include new relevant genomic regions and this could make this approach less cost-effective and possibly impractical for clinical applications. A close collaboration between cytogeneticists and clinical geneticists, combined with parental analysis, is also very important and can reduce the amount of VUCS. In similarity with postnatal testing, CNVs found in susceptibility loci with reduced penetrance and/or variable expressivity represent a true challenge in counseling because the degree of clinical manifestation is difficult to predict. The Belgian experience shows that it may be reasonable to report only those CNVs for which the risk of a severe phenotype is sufficiently large and/or which are associated with structural malformations for which an ultrasound follow-up is warranted [116]. Incidental CMA findings should generally be treated in the same way as in case of postnatal CMA testing. The information of late-onset genetic diseases with clinical utility should not be withheld, and, in addition, it should be estimated whether this information may be relevant for the health of adults related to the fetus [97, 116, 119]. Knowledge regarding possible consanguinity should be returned to the ordering clinician and clinical reports should avoid speculating on specific parental relationships [106].

The inability of CMA to detect balanced aberrations could be considered as one of the major drawbacks that confine the application of CMA in prenatal diagnostics. However, balanced rearrangements occur in only between 0.08 to 0.09% of prenatal diagnosis samples [122]. Moreover, while familial events mostly lack any negative consequences for the current pregnancy, the increased risk of congenital anomalies associated with *de novo* and apparently balanced aberrations is often caused by the presence of tiny CNVs at the breakpoints not detectable by karyotyping but discoverable by CMA [111].

Table 9. CMA findings in prenatal tests (including cases tested after the termination of pregnancy) [207].

Case	Indication	Karyotype	CMA	Clinical significance	Outcome
1	Isolated increased nuchal translucency	–	arr[hg19] Xp21.1(31,665,779-32,096,779)x3	UCS	Termination of pregnancy
2	Recurrent spontaneous abortions	–	arr[hg19] 8q11.1q11.23 (47,060,977-52,693,165)x2 hmz	UCS	Tested after the termination of pregnancy
3	Positive triple test	46,XX	multiple LCSH across whole genome	Pathogenic	Termination of pregnancy
4	Positive triple test	46,XX[64]/ 47,XX,+7[9]	arr(7)x2~3 (10–20%)	Likely benign	Normal female at term
5	Abnormal ultrasound	–	arr[hg19] 7p14.1p13 (42,179,377-44,932,538)x1	Pathogenic (Greig syndrome, OMIM #175700)	Tested after the termination of pregnancy
6	Familial balanced rearrangement	46,XX,rec(4)dup (4p)del(4q)inv(4) (p15.3q35)pat	arr[hg19] 4p13.33p16.3(1–13,912,694)x3, 4q35.2(188,730,709–190,880,409)x1	Pathogenic	Termination of pregnancy
7	Familial balanced rearrangement	46,XY,inv(10) (p11.2;q21.2)mat	arr[hg19] 12q14.2q15 (63,291,364–68,794,078)x2 hmz	UCS	Not known
8	Dysmorphic foetus	–	multiple LCSH across whole genome	Pathogenic	Tested after the termination of pregnancy

UCS – uncertain clinical significance; LCSH – long contiguous stretches of homozygosity.

In our prenatal cohort of 52 high-risk pregnancies and 8 fetuses tested after the termination of pregnancy, CMA was used in conjunction with conventional karyotyping in most cases. As expected, the unbalanced changes observed on G-banding were also seen by CMA, while balanced rearrangements remained undetected. Low-level mosaic trisomy 7 (~10%) was also identified by both CMA and karyotyping. In addition, CMA was able to identify multiple LCSH in two cases and a small pathogenic deletion that would be missed by traditional methods. Because of the relatively small prenatal cohort we avoided making any conclusions about applying CMA as a first-line test in prenatal diagnosis. Currently, CMA is mainly applied in parallel with traditional cytogenetic analyses, and a number of reports that compare the diagnostic efficacy of these approaches have been published [111, 115, 236, 237]. However, the application of CMA in prenatal diagnosis remains controversial. Until recently, CMA has not been recommended as a principal cytogenetic tool in prenatal diagnosis, but instead it should be used as an additional test in light of abnormal ultrasound findings [112, 113]. However, this opinion has been re-evaluated based on the results of a comprehensive study on parallel application of karyotyping and CMA in a large cohort of pregnant women [111]. According to The American College of Obstetrics and Gynecology, the main prenatal CMA target groups should include pregnancies with abnormal ultrasound findings, and cases of fetal demise and stillbirth. In patients undergoing invasive prenatal diagnostic testing due to reasons other than fetal structural abnormalities, either karyotyping or CMA can be performed [114]. It has been demonstrated that the presence of an anomaly using ultrasonography increases the chance of detecting pathogenic CNVs [111, 115]. Interestingly, it has been reported recently that the degree of additional information provided by CMA in those cases depends on the organ system affected, with anomalies of the renal and cardiac systems being significantly associated with the presence of (potentially) pathogenic CNVs [238]. Still, clinically relevant findings could be detected in fetuses without any ultrasonographically visible structural anomalies which can be explained by the complete absence or presence of only subtle fetal abnormalities detectable by ultrasound examination in many microdeletion/ microduplication syndromes associated with neurodevelopmental problems and ID. Due to this fact, the application of CMA for all pregnancies without additional indications should be considered [236]. In all cases testing should include a comprehensive patient pre-test followed by post-test genetic counseling regarding the benefits, limitations, and results of CMA analysis. The counseling should include a discussion of the potential to identify findings of uncertain clinical impact, nonpaternity, consanguinity, and adult-onset diseases [114].

5. SUMMARY AND CONCLUSIONS

Chromosomal aberrations are a well-known cause of various pathological conditions, such as ID/DD, diverse malformations, congenital anomalies, and neurological disorders. Due to the capacity of CMA to analyze the entire genome with regards to DNA copy number changes as well as its relatively high diagnostic yield, this approach is currently recommended as the first-tier cytogenetic diagnostic test for the rapid and accurate detection of chromosomal abnormalities in patients with ID/DD, MCA, or ASD.

The first aim of this dissertation is to evaluate the application CMA in routine clinical practice in Estonia for the detection of chromosomal aberrations associated with developmental disorders. In summary, clinically relevant findings were detected in 11% of the patients analyzed during 2009–2012 which is a rate comparable with previous reports. We found that applying CMA to analyze patients with more prominent phenotypes of unknown etiology resulted in a higher diagnostic yield (15%), while the use of CMA as a first-line diagnostic test in a more heterogeneous population allowed for the discovery of the underlying reason in 10% of the patients. In all cases, the diagnostic yield of CMA is at least two-fold higher compared to conventional karyotyping. At the same time, the proportion of CMA findings of unknown clinical relevance remained quite high (41% of all reported findings, including LCSH; 24% when LCSH are excluded). It has been reported elsewhere and noted in the current study that interpretation of CMA results tends to change over time due to refinements in our knowledge regarding pathogenic and benign CNVs both in the literature and databases, detection of novel candidate genes for various diseases, and remapping of gene locations in new human genome assemblies. Still, it should be kept in mind that VUCS reported to patients and their families often cause stress and serious doubts. Close collaboration between cyto-geneticists and clinical geneticists is often required to decipher the maximum number of CMA findings and thus aids in providing proper counseling for families.

Secondly, the Estonian experience in the application of CMA for prenatal diagnosis is presented. CMA was mostly used in conjunction with conventional karyotyping and was able to detect all unbalanced rearrangements observed by G-banding. In addition, CMA enabled the identification of pathogenic changes which are not detectable by karyotyping, and thus provides novel opportunities in the field of prenatal diagnostics. However, because of the small prenatal cohort studied we have avoided making any conclusions with regards to the application of CMA as a first-line test in prenatal diagnosis. Although CMA is now recommended as a principal cytogenetic tool for pregnancies with abnormal ultrasound findings, controversial issues remain with regard to its application in prenatal diagnostics, including ethical considerations, target populations, challenges associated with VUCS, and the choice of array platform.

Thirdly, two families were described with a 7q31 deletion encompassing the *FOXP2* gene associated with speech and language disorder. It has previously been proposed that the *FOXP2* locus may be differentially regulated in a parent-of-origin manner, and lesions of paternal and maternal alleles could have differential impact on the phenotype. Therefore, these particular families allow, for the first time, the opportunity to estimate potential intrafamilial parent-of-origin-dependent differences in the severity of phenotype observed. Following detailed clinical evaluation and speech and language assessment, it was concluded that there are no differences in the severity of phenotype and that alterations of both maternal and paternal origins cause severe speech and language disorder.

Fourthly, a patient with 5q35.2-q35.3 duplication reciprocal to the common Sos deletion added to a growing body of knowledge regarding phenotypes associated with reciprocal deletions and duplications. Although 5q35.2-q35.3 duplication syndrome cannot be referred as a "reversed Sos", the observed phenotypic characteristics related to body parameters, including growth and head circumference, seem to be remarkably opposite comparing deletion and duplication of this region. Although it seems that the *NSD1* gene located within this chromosomal region is responsible for overgrowth/short stature associated with 5q35.2-q35.3 deletion/duplication, the molecular mechanism underlying these phenotypic changes remains unknown. It is also unclear whether other genes within the 5q35.2-q35.3 region play a role in the formation of observed clinical phenotypes. We analyzed the expression of genes from the Sos critical region and flanking areas in three patients with 5q35.2-q35.3 duplication and found that nearly all the genes analyzed were overexpressed. However, both the design of the study and the use of a small number of patients do not allow us to draw any statistically significant conclusions. It is obvious that when studying cases of genetic variation with low frequency, multi-center collaboration is required to both describe their phenotypic effects and to establish the precise mechanisms responsible for their clinical manifestation.

Lastly, a patient with mosaic matUPD15 was included in this work as an example of the utility of SNP-based CMA to detect uniparental disomy and mosaicism. The patient was found to have mosaic isodisomy 15 (55–60% in peripheral blood) and a normal biparental cell line. The molecular mechanism underlying this particular mosaic UPD formation was proposed based on the results of CMA, karyotyping, and methylation analysis. So far, only three PWS cases with mosaic maternal isodisomy 15 and without trisomic/monosomic cell lines have been reported. Still, it is unclear whether mosaic UPD is truly rare or simply underdiagnosed in PWS patients due to either the partial PWS phenotype observed in those patients, or a lack of diagnostic tools in various laboratories, or difficulties associated with the combined interpretation of multiple results. Although, a wide range of different methods for imprinting disorders is available, none of these are able to detect all possible (epi)mutations. Using a combination of CMA and quantitive methylation

analysis method, we propose that the combination of these two techniques could be useful when evaluating putative imprinting disorders.

In summary, our experience has confirmed that CMA is a reliable and effective tool for the detection of chromosomal aberrations responsible for developmental disorders and its application as a primary cytogenetic test is completely justified. However, it should be kept in mind that the interpretation of CMA findings is based on current knowledge and may evolve over time. In this respect it is important for the referring physician to re-evaluate results for each patient with CMA findings of unknown clinical relevance at a regular interval and consult with the patient if the diagnosis changes.

REFERENCES

[1] Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med*. 2010; 61:437–455.

[2] Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, et al. Origins and functional impact of copy number variation in the human genome. *Nature*. 2010; 464:704–712.

[3] Coughlin CR, 2nd, Scharer GH, Shaikh TH. Clinical impact of copy number variation analysis using high-resolution microarray technologies: advantages, limitations and concerns. *Genome Med*. 2012; 4:80.

[4] Vissers LE, de Vries BB, Veltman JA. Genomic microarrays in mental retardation: from copy number variation to gene, from research to diagnosis. *J Med Genet*. 2010; 47:289–297.

[5] Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*. 2010; 86:749–764.

[6] Ahn JW, Bint S, Bergbaum A, Mann K, Hall RP, Ogilvie CM. Array CGH as a first line diagnostic test in place of karyotyping for postnatal referrals – results from four years' clinical application for over 8,700 patients. *Mol Cytogenet*. 2013; 6:16.

[7] Bejjani BA, Shaffer LG. Clinical utility of contemporary molecular cytogenetics. *Annu Rev Genomics Hum Genet*. 2008; 9:71–86.

[8] Mannik K, Parkel S, Palta P, Zilina O, Puusepp H, Esko T, Magi R, Noukas M, Veidenberg A, Nelis M, et al. A parallel SNP array study of genomic aberrations associated with mental retardation in patients and general population in Estonia. *Eur J Med Genet*. 2011; 54:136–143.

[9] Teek R, Kruustuk K, Zordania R, Joost K, Kahre T, Tonisson N, Nelis M, Zilina O, Tranebjaerg L, Reimand T, et al. Hearing impairment in Estonia: an algorithm to investigate genetic causes in pediatric patients. *Adv Med Sci*. 2013; 58:419–428.

[10] McFadden DE, Friedman JM. Chromosome abnormalities in human beings. *Mutat Res*. 1997; 396:129–140.

[11] Chance PF, Alderson MK, Leppig KA, Lensch MW, Matsunami N, Smith B, Swanson PD, Odelberg SJ, Disteche CM, Bird TD. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. *Cell*. 1993; 72:143–151.

[12] Lupski JR, de Oca-Luna RM, Slaugenhaupt S, Pentao L, Guzzetta V, Trask BJ, Saucedo-Cardenas O, Barker DF, Killian JM, Garcia CA, et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell*. 1991; 66:219–232.

[13] Lupski JR, Wise CA, Kuwano A, Pentao L, Parke JT, Glaze DG, Ledbetter DH, Greenberg F, Patel PI. Gene dosage is a mechanism for Charcot-Marie-Tooth disease type 1A. *Nat Genet*. 1992; 1:29–33.

[14] Lejeune J, Gautier M, Turpin R. [Study of somatic chromosomes from 9 mongoloid children]. *C R Hebd Seances Acad Sci*. 1959; 248:1721–1722.

[15] Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. Detection of large-scale variation in the human genome. *Nat Genet*. 2004; 36:949–951.

[16] Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004; 305:525–528.

[17] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al. Global variation in copy number in the human genome. *Nature*. 2006; 444:444–454.

[18] Conrad DF, Andrews TD, Carter NP, Hurles ME, Pritchard JK. A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet*. 2006; 38:75–81.

[19] Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet*. 2006; 7:85–97.

[20] Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*. 1998; 14:417–422.

[21] Pang AW, MacDonald JR, Pinto D, Wei J, Rafiq MA, Conrad DF, Park H, Hurles ME, Lee C, Venter JC, et al. Towards a comprehensive structural variation map of an individual human genome. *Genome Biol*. 2010; 11:R52.

[22] Ford CE, Hamerton JL. The chromosomes of man. *Nature*. 1956; 178:1020–1023.

[23] Tjio HJ, Levan A. The chromosome numbers of man. *Hereditas*. 1956; 42:1–6.

[24] Ford CE, Jones KW, Polani PE, De Almeida JC, Briggs JH. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet*. 1959; 1:711–713.

[25] Jacobs PA, Strong JA. A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature*. 1959; 183:302–303.

[26] Nowell PC, Hungerford DA. A minute chromosome in human chronic granulocytic leukemia. *Science*. 1960; 132:1497–1501.

[27] Lejeune J, Lafourcade J, Berger R, Vialatte J, Boeswillwald M, Seringe P, Turpin R. [3 Cases of Partial Deletion of the Short Arm of a 5 Chromosome]. *C R Hebd Seances Acad Sci*. 1963; 257:3098–3102.

[28] Trask BJ. Human cytogenetics: 46 chromosomes, 46 years and counting. *Nat Rev Genet*. 2002; 3:769–778.

[29] Caspersson T, Farber S, Foley GE, Kudynowski J, Modest EJ, Simonsson E, Wagh U, Zech L. Chemical differentiation along metaphase chromosomes. *Exp Cell Res*. 1968; 49:219–222.

[30] Seabright M. A rapid banding technique for human chromosomes. *Lancet*. 1971; 2:971–972.

[31] Smeets DF. Historical prospective of human cytogenetics: from microscope to microarray. *Clin Biochem*. 2004; 37:439–446.

[32] Langer-Safer PR, Levine M, Ward DC. Immunological method for mapping genes on Drosophila polytene chromosomes. *Proc Natl Acad Sci U S A*. 1982; 79:4381–4385.

[33] Van Prooijen-Knegt AC, Van Hoek JF, Bauman JG, Van Duijn P, Wool IG, Van der Ploeg M. In situ hybridization of DNA sequences in human metaphase chromosomes visualized by an indirect fluorescent immunocytochemical procedure. *Exp Cell Res*. 1982; 141:397–407.

[34] Speicher MR, Gwyn Ballard S, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet*. 1996; 12:368–375.
[35] Schrock E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, et al. Multicolor spectral karyotyping of human chromosomes. *Science*. 1996; 273:494–497.
[36] Bieche I, Olivi M, Champeme MH, Vidaud D, Lidereau R, Vidaud M. Novel approach to quantitative polymerase chain reaction using real-time detection: application to the detection of gene amplification in breast cancer. *Int J Cancer*. 1998; 78:661–666.
[37] Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*. 2002; 30:e57.
[38] Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*. 1992; 258:818–821.
[39] Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer*. 1997; 20:399–407.
[40] Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet*. 1998; 20:207–211.
[41] Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigorova M, Jones KW, Wei W, Stratton MR, et al. High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res*. 2004; 14:287–295.
[42] Huang J, Wei W, Zhang J, Liu G, Bignell GR, Stratton MR, Futreal PA, Wooster R, Jones KW, Shapero MH. Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics*. 2004; 1:287–299.
[43] Conlin LK, Thiel BD, Bonnemann CG, Medne L, Ernst LM, Zackai EH, Deardorff MA, Krantz ID, Hakonarson H, Spinner NB. Mechanisms of mosaicism, chimerism and uniparental disomy identified by single nucleotide polymorphism array analysis. *Hum Mol Genet*. 2010; 19:1263–1275.
[44] Abel HJ, Duncavage EJ. Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches. *Cancer Genet*. 2013; 206:432–440.
[45] Tan R, Wang Y, Kleinstein SE, Liu Y, Zhu X, Guo H, Jiang Q, Allen AS, Zhu M. An evaluation of copy number variation detection tools from whole-exome sequencing data. *Hum Mutat*. 2014; 35:899–907.
[46] Le Scouarnec S, Gribble SM. Characterising chromosome rearrangements: recent technical advances in molecular cytogenetics. *Heredity (Edinb)*. 2012; 108:75–85.
[47] Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet*. 2001; 29:263–264.
[48] Vissers LE, de Vries BB, Osoegawa K, Janssen IM, Feuth T, Choy CO, Straatman H, van der Vliet W, Huys EH, van Rijk A, et al. Array-based comparative

genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet*. 2003; 73:1261–1270.
[49] Menten B, Maas N, Thienpont B, Buysse K, Vandesompele J, Melotte C, de Ravel T, Van Vooren S, Balikova I, Backx L, et al. Emerging patterns of cryptic chromosomal imbalance in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of published reports. *J Med Genet*. 2006; 43:625–633.
[50] Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, et al. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet*. 2004; 36:299–303.
[51] Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet*. 1999; 23:41–46.
[52] Mantripragada KK, Tapia-Paez I, Blennow E, Nilsson P, Wedell A, Dumanski JP. DNA copy-number analysis of the 22q11 deletion-syndrome region using array-CGH with genomic and PCR-based targets. *Int J Mol Med*. 2004; 13:273–279.
[53] Brennan C, Zhang Y, Leo C, Feng B, Cauwels C, Aguirre AJ, Kim M, Protopopov A, Chin L. High-resolution global profiling of genomic alterations with long oligonucleotide microarray. *Cancer Res*. 2004; 64:4744–4748.
[54] Carvalho B, Ouwerkerk E, Meijer GA, Ylstra B. High resolution microarray comparative genomic hybridisation analysis using spotted oligonucleotides. *J Clin Pathol*. 2004; 57:644–646.
[55] Ylstra B, van den Ijssel P, Carvalho B, Brakenhoff RH, Meijer GA. BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res*. 2006; 34:445–450.
[56] Boone PM, Bacino CA, Shaw CA, Eng PA, Hixson PM, Pursley AN, Kang SH, Yang Y, Wiszniewska J, Nowakowska BA, et al. Detection of clinically relevant exonic copy-number changes by array CGH. *Hum Mutat*. 2010; 31:1326–1342.
[57] Keren B, Le Caignec C. Oligonucleotide microarrays in constitutional genetic diagnosis. *Expert Rev Mol Diagn*. 2011; 11:521–532.
[58] Schaaf CP, Wiszniewska J, Beaudet AL. Copy number and SNP arrays in clinical diagnostics. *Annu Rev Genomics Hum Genet*. 2011; 12:25–51.
[59] LaFramboise T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res*. 2009; 37:4181–4193.
[60] IHMC TIHC. The International HapMap Project. *Nature*. 2003; 426:789–796.
[61] Durbin RM, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. *Nature*. 2010; 467:1061–1073.
[62] Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nat Rev Genet*. 2011; 12:363–376.
[63] Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res*. 2006; 16:1136–1148.
[64] Kearney HM, Kearney JB, Conlin LK. Diagnostic implications of excessive homozygosity detected by SNP-based microarrays: consanguinity, uniparental

disomy, and recessive single-gene mutations. *Clin Lab Med*. 2011; 31:595–613, ix.

[65] Cooper GM, Zerr T, Kidd JM, Eichler EE, Nickerson DA. Systematic assessment of copy number variant detection via genome-wide SNP genotyping. *Nat Genet*. 2008; 40:1199–1203.

[66] Pinto D, Darvishi K, Shi X, Rajan D, Rigler D, Fitzgerald T, Lionel AC, Thiruvahindrapuram B, Macdonald JR, Mills R, et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat Biotechnol*. 2011; 29:512–520.

[67] Tucker T, Montpetit A, Chai D, Chan S, Chenier S, Coe BP, Delaney A, Eydoux P, Lam WL, Langlois S, et al. Comparison of genome-wide array genomic hybridization platforms for the detection of copy number variants in idiopathic mental retardation. *BMC Med Genomics*. 2011; 4:25.

[68] Stankiewicz P, Lupski JR. Genome architecture, rearrangements and genomic disorders. *Trends Genet*. 2002; 18:74–82.

[69] Weiss MM, Kuipers EJ, Postma C, Snijders AM, Siccama I, Pinkel D, Westerga J, Meuwissen SG, Albertson DG, Meijer GA. Genomic profiling of gastric cancer predicts lymph node status and survival. *Oncogene*. 2003; 22:1872–1879.

[70] Oostlander AE, Meijer GA, Ylstra B. Microarray-based comparative genomic hybridization and its applications in human genetics. *Clin Genet*. 2004; 66:488–495.

[71] De Vries BB, Winter R, Schinzel A, van Ravenswaaij-Arts C. Telomeres: a diagnosis at the end of the chromosomes. *J Med Genet*. 2003; 40:385–398.

[72] Ropers HH. Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet*. 2010; 11:161–187.

[73] Flint J, Wilkie AO, Buckle VJ, Winter RM, Holland AJ, McDermid HE. The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nat Genet*. 1995; 9:132–140.

[74] Reimand T, Ounap K, Zordania R, Ilus T, Uibo O, Sitska M, Talvik T. Descriptive epidemiology of Down's syndrome in Estonia. *Paediatr Perinat Epidemiol*. 2006; 20:512–519.

[75] Ropers HH. Genetics of intellectual disability. *Curr Opin Genet Dev*. 2008; 18:241–250.

[76] Koolen DA, Pfundt R, de Leeuw N, Hehir-Kwa JY, Nillesen WM, Neefs I, Scheltinga I, Sistermans E, Smeets D, Brunner HG, et al. Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. *Hum Mutat*. 2009; 30:283–292.

[77] de Vries BB, Pfundt R, Leisink M, Koolen DA, Vissers LE, Janssen IM, Reijmersdal S, Nillesen WM, Huys EH, Leeuw N, et al. Diagnostic genome profiling in mental retardation. *Am J Hum Genet*. 2005; 77:606–616.

[78] Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, et al. A copy number variation morbidity map of developmental delay. *Nat Genet*. 2011; 43:838–846.

[79] Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007; 316:445–449.

[80] Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, et al. Functional impact of global rare

copy number variation in autism spectrum disorders. *Nature*. 2010; 466:368–372.

[81] Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, Marks S, Lakshmi B, Pai D, Ye K, et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*. 2011; 70:886–897.

[82] Devlin B, Scherer SW. Genetic architecture in autism spectrum disorder. *Curr Opin Genet Dev*. 2012; 22:229–237.

[83] Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet*. 2008; 40:880–885.

[84] Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, Fossdal R, Sigurdsson E, Sigmundsson T, Buizer-Voskamp JE, et al. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008; 455:232–236.

[85] Luo X, Huang L, Han L, Luo Z, Hu F, Tieu R, Gan L. Systematic Prioritization and Integrative Analysis of Copy Number Variations in Schizophrenia Reveal Key Schizophrenia Susceptibility Genes. *Schizophr Bull*. 2014; Mar 24. [Epub ahead of print].

[86] Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, Falchi M, Chen F, Andrieux J, Lobbens S, et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature*. 2010; 463:671–675.

[87] Shaffer LG, Theisen A, Bejjani BA, Ballif BC, Aylsworth AS, Lim C, McDonald M, Ellison JW, Kostiner D, Saitta S, et al. The discovery of microdeletion syndromes in the post-genomic era: review of the methodology and characterization of a new 1q41q42 microdeletion syndrome. *Genet Med*. 2007; 9:607–616.

[88] Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K, Huang S, Maloney VK, Crolla JA, Baralle D, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med*. 2008; 359:1685–1699.

[89] Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, Lalani SR, Graham B, Lee B, Shinawi M, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet*. 2008; 40:1466–1471.

[90] Sharp AJ, Mefford HC, Li K, Baker C, Skinner C, Stevenson RE, Schroer RJ, Novara F, De Gregori M, Ciccone R, et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet*. 2008; 40:322–328.

[91] Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC, et al. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet*. 2004; 36:955–957.

[92] McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker PI, Maller JB, Kirby A, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet*. 2008; 40:1166–1174.

[93] MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014; 42:D986–992.

[94] Kearney HM, South ST, Wolff DJ, Lamb A, Hamosh A, Rao KW. American College of Medical Genetics recommendations for the design and performance expectations for clinical genomic copy number microarrays intended for use in the postnatal setting for detection of constitutional abnormalities. *Genet Med*. 2011; 13:676–679.

[95] Spielmann M, Klopocki E. CNVs of noncoding cis-regulatory elements in human disease. *Curr Opin Genet Dev*. 2013; 23:249–256.

[96] Klopocki E, Mundlos S. Copy-number variations, noncoding sequences, and human phenotypes. *Annu Rev Genomics Hum Genet*. 2011; 12:53–72.

[97] Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med*. 2011; 13:680–685.

[98] Palmer E, Speirs H, Taylor PJ, Mullan G, Turner G, Einfeld S, Tonge B, Mowat D. Changing interpretation of chromosomal microarray over time in a community cohort with intellectual disability. *Am J Med Genet A*. 2014; 164A:377–385.

[99] Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet*. 2005; 1:e49.

[100] Battaglia A, Doccini V, Bernardini L, Novelli A, Loddo S, Capalbo A, Filippi T, Carey JC. Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *Eur J Paediatr Neurol*. 2013; 17:589–599.

[101] Ellison JW, Ravnan JB, Rosenfeld JA, Morton SA, Neill NJ, Williams MS, Lewis J, Torchia BS, Walker C, Traylor RN, et al. Clinical utility of chromosomal microarray analysis. *Pediatrics*. 2012; 130:e1085–1095.

[102] Howell KB, Kornberg AJ, Harvey AS, Ryan MM, Mackay MT, Freeman JL, Rodriguez Casero MV, Collins KJ, Hayman M, Mohamed A, et al. High resolution chromosomal microarray in undiagnosed neurological disorders. *J Paediatr Child Health*. 2013; 49:716–724.

[103] Boone PM, Soens ZT, Campbell IM, Stankiewicz P, Cheung SW, Patel A, Beaudet AL, Plon SE, Shaw CA, McGuire AL, et al. Incidental copy-number variants identified by routine genome testing in a clinical population. *Genet Med*. 2013; 15:45–54.

[104] Adams SA, Coppinger J, Saitta SC, Stroud T, Kandamurugu M, Fan Z, Ballif BC, Shaffer LG, Bejjani BA. Impact of genotype-first diagnosis: the detection of microdeletion and microduplication syndromes with cancer predisposition by aCGH. *Genet Med*. 2009; 11:314–322.

[105] Pichert G, Mohammed SN, Ahn JW, Ogilvie CM, Izatt L. Unexpected findings in cancer predisposition genes detected by array comparative genomic hybridisation: what are the issues? *J Med Genet*. 2011; 48:535–539.

[106] Rehder CW, David KL, Hirsch B, Toriello HV, Wilson CM, Kearney HM. American College of Medical Genetics and Genomics: standards and guidelines for documenting suspected consanguinity as an incidental finding of genomic testing. *Genet Med*. 2013; 15:150–152.

[107] Wapner RJ, Driscoll DA, Simpson JL. Integration of microarray technology into prenatal diagnosis: counselling issues generated during the NICHD clinical trial. *Prenat Diagn*. 2012; 32:396–400.

[108] Evangelidou P, Sismani C, Ioannides M, Christodoulou C, Koumbaris G, Kallikas I, Georgiou I, Velissariou V, Patsalis PC. Clinical application of whole-genome array CGH during prenatal diagnosis: Study of 25 selected pregnancies with abnormal ultrasound findings or apparently balanced structural aberrations. *Mol Cytogenet*. 2010; 3:24.

[109] Schmid M, Stary S, Springer S, Bettelheim D, Husslein P, Streubel B. Prenatal microarray analysis as second-tier diagnostic test: single-center prospective study. *Ultrasound Obstet Gynecol*. 2013; 41:267–273.

[110] Van den Veyver IB, Patel A, Shaw CA, Pursley AN, Kang SH, Simovich MJ, Ward PA, Darilek S, Johnson A, Neill SE, et al. Clinical use of array comparative genomic hybridization (aCGH) for prenatal diagnosis in 300 cases. *Prenat Diagn*. 2009; 29:29–39.

[111] Wapner RJ, Martin CL, Levy B, Ballif BC, Eng CM, Zachary JM, Savage M, Platt LD, Saltzman D, Grobman WA, et al. Chromosomal microarray versus karyotyping for prenatal diagnosis. *N Engl J Med*. 2012; 367:2175–2184.

[112] ACOG. ACOG Committee Opinion No. 446: array comparative genomic hybridization in prenatal diagnosis. *Obstet Gynecol*. 2009; 114:1161–1163.

[113] Novelli A, Grati FR, Ballarati L, Bernardini L, Bizzoco D, Camurri L, Casalone R, Cardarelli L, Cavalli P, Ciccone R, et al. Microarray application in prenatal diagnosis: a position statement from the cytogenetics working group of the Italian Society of Human Genetics (SIGU), November 2011. *Ultrasound Obstet Gynecol*. 2012; 39:384–388.

[114] ACOG. Committee Opinion No. 581: the use of chromosomal microarray analysis in prenatal diagnosis. *Obstet Gynecol*. 2013; 122:1374–1377.

[115] Hillman SC, McMullan DJ, Hall G, Togneri FS, James N, Maher EJ, Meller CH, Williams D, Wapner RJ, Maher ER, et al. Use of prenatal chromosomal microarray: prospective cohort study and systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2013; 41:610–620.

[116] Vanakker O, Vilain C, Janssens K, Van der Aa N, Smits G, Bandelier C, Blaumeiser B, Bulk S, Caberg JH, De Leener A, et al. Implementation of genomic arrays in prenatal diagnosis: The Belgian approach to meet the challenges. *Eur J Med Genet*. 2014; 57:151–156.

[117] Liao C, Fu F, Li R, Xie GE, Zhang YL, Li J, Li DZ. Implementation of high-resolution SNP arrays in the investigation of fetuses with ultrasound mal-formations: 5 years of clinical experience. *Clin Genet*. 2013; Sep 3. doi: 10.1111/cge.12271. [Epub ahead of print].

[118] Ganesamoorthy D, Bruno DL, McGillivray G, Norris F, White SM, Adroub S, Amor DJ, Yeung A, Oertel R, Pertile MD, et al. Meeting the challenge of interpreting high-resolution single nucleotide polymorphism array data in prenatal diagnosis: does increased diagnostic power outweigh the dilemma of rare variants? *BJOG*. 2013; 120:594–606.

[119] Stark Z, Gillam L, Walker SP, McGillivray G. Ethical controversies in prenatal microarray. *Curr Opin Obstet Gynecol*. 2013; 25:133–137.

[120] Vetro A, Bouman K, Hastings R, McMullan DJ, Vermeesch JR, Miller K, Sikkema-Raddatz B, Ledbetter DH, Zuffardi O, van Ravenswaaij-Arts CM. The introduction of arrays in prenatal diagnosis: a special challenge. *Hum Mutat*. 2012; 33:923–929.

[121] Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. *Genet Med*. 2013; 15:478–481.

[122] Giardino D, Corti C, Ballarati L, Colombo D, Sala E, Villa N, Piombo G, Pierluigi M, Faravelli F, Guerneri S, et al. De novo balanced chromosome rearrangements in prenatal diagnosis. *Prenat Diagn*. 2009; 29:257–265.

[123] Harel S, Greenstein Y, Kramer U, Yifat R, Samuel E, Nevo Y, Leitner Y, Kutai M, Fattal A, Shinnar S. Clinical characteristics of children referred to a child development center for evaluation of speech, language, and communication disorders. *Pediatr Neurol*. 1996; 15:305–311.

[124] Newbury DF, Fisher SE, Monaco AP. Recent advances in the genetics of language impairment. *Genome Med*. 2010; 2:6.

[125] Newbury DF, Monaco AP. Genetic advances in the study of speech and language disorders. *Neuron*. 2010; 68:309–320.

[126] Graham SA, Fisher SE. Decoding the genetics of speech and language. *Curr Opin Neurobiol*. 2013; 23:43–51.

[127] Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001; 413:519–523.

[128] Lai CS, Fisher SE, Hurst JA, Levy ER, Hodgson S, Fox M, Jeremiah S, Povey S, Jamison DC, Green ED, et al. The SPCH1 region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *Am J Hum Genet*. 2000; 67:357–368.

[129] MacDermot KD, Bonora E, Sykes N, Coupe AM, Lai CS, Vernes SC, Vargha-Khadem F, McKenzie F, Smith RL, Monaco AP, et al. Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet*. 2005; 76:1074–1080.

[130] Shriberg LD, Ballard KJ, Tomblin JB, Duffy JR, Odell KH, Williams CA. Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. *J Speech Lang Hear Res*. 2006; 49:500–525.

[131] Tomblin JB, O'Brien M, Shriberg LD, Williams C, Murray J, Patil S, Bjork J, Anderson S, Ballard K. Language features in a mother and daughter of a chromosome 7;13 translocation involving FOXP2. *J Speech Lang Hear Res*. 2009; 52:1157–1174.

[132] Feuk L, Kalervo A, Lipsanen-Nyman M, Skaug J, Nakabayashi K, Finucane B, Hartung D, Innes M, Kerem B, Nowaczyk MJ, et al. Absence of a paternally inherited FOXP2 gene in developmental verbal dyspraxia. *Am J Hum Genet*. 2006; 79:965–972.

[133] Zeesman S, Nowaczyk MJ, Teshima I, Roberts W, Cardy JO, Brian J, Senman L, Feuk L, Osborne LR, Scherer SW. Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves FOXP2. *Am J Med Genet A*. 2006; 140:509–514.

[134] Lennon PA, Cooper ML, Peiffer DA, Gunderson KL, Patel A, Peters S, Cheung SW, Bacino CA. Deletion of 7q31.1 supports involvement of FOXP2 in language impairment: clinical report and review. *Am J Med Genet A*. 2007; 143A:791–798.

[135] Turner SJ, Hildebrand MS, Block S, Damiano J, Fahey M, Reilly S, Bahlo M, Scheffer IE, Morgan AT. Small intragenic deletion in FOXP2 associated with

childhood apraxia of speech and dysarthria. *Am J Med Genet A*. 2013; 161:2321–2326.

[136] Rice GM, Raca G, Jakielski KJ, Laffin JJ, Iyama-Kurtycz CM, Hartley SL, Sprague RE, Heintzelman AT, Shriberg LD. Phenotype of FOXP2 haplo-insufficiency in a mother and son. *Am J Med Genet A*. 2012; 158A:174–181.

[137] Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME. Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet*. 1998; 18:168–170.

[138] Vargha-Khadem F, Watkins KE, Price CJ, Ashburner J, Alcock KJ, Connelly A, Frackowiak RS, Friston KJ, Pembrey ME, Mishkin M, et al. Neural basis of an inherited speech and language disorder. *Proc Natl Acad Sci U S A*. 1998; 95:12695–12700.

[139] Vargha-Khadem F, Gadian DG, Copp A, Mishkin M. FOXP2 and the neuro-anatomy of speech and language. *Nat Rev Neurosci*. 2005; 6:131–138.

[140] Newbury DF, Bonora E, Lamb JA, Fisher SE, Lai CS, Baird G, Jannoun L, Slonims V, Stott CM, Merricks MJ, et al. FOXP2 is not a major susceptibility gene for autism or specific language impairment. *Am J Hum Genet*. 2002; 70:1318–1327.

[141] Fisher SE, Scharff C. FOXP2 as a molecular window into speech and language. *Trends Genet*. 2009; 25:166–177.

[142] Vernes SC, Spiteri E, Nicod J, Groszer M, Taylor JM, Davies KE, Geschwind DH, Fisher SE. High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *Am J Hum Genet*. 2007; 81:1232–1250.

[143] Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J, Vernes SC, Fisher SE, Ren B, Geschwind DH. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet*. 2007; 81:1144–1157.

[144] Shu W, Yang H, Zhang L, Lu MM, Morrisey EE. Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. *J Biol Chem*. 2001; 276:27488–27497.

[145] Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM, Ho J, Mombereau C, Brewer A, Lowy E, et al. Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genet*. 2011; 7:e1002145.

[146] Enard W, Gehre S, Hammerschmidt K, Holter SM, Blass T, Somel M, Bruckner MK, Schreiweis C, Winter C, Sohr R, et al. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*. 2009; 137:961–971.

[147] Reimers-Kipping S, Hevers W, Paabo S, Enard W. Humanized Foxp2 specifi-cally affects cortico-basal ganglia circuits. *Neuroscience*. 2011; 175:75–84.

[148] Sia GM, Clem RL, Huganir RL. The human language-associated gene SRPX2 regulates synapse formation and vocalization in mice. *Science*. 2013; 342:987–991.

[149] Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, Burbano HA, Hublin JJ, Hanni C, Fortea J, de la Rasilla M, et al. The derived FOXP2 variant of modern humans was shared with Neandertals. *Curr Biol*. 2007; 17:1908–1912.

[150] Maricic T, Gunther V, Georgiev O, Gehre S, Curlin M, Schreiweis C, Naumann R, Burbano HA, Meyer M, Lalueza-Fox C, et al. A recent evolutionary change

affects a regulatory element in the human FOXP2 gene. *Mol Biol Evol*. 2013; 30:844–852.

[151] Lee JA, Lupski JR. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. *Neuron*. 2006; 52:103–121.

[152] Hannes FD, Sharp AJ, Mefford HC, de Ravel T, Ruivenkamp CA, Breuning MH, Fryns JP, Devriendt K, Van Buggenhout G, Vogels A, 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*. 2009; 46:223–232.

[153] Schubert C. The genomic basis of the Williams-Beuren syndrome. *Cell Mol Life Sci*. 2009; 66:1178–1197.

[154] Van der Aa N, Rooms L, Vandeweyer G, van den Ende J, Reyniers E, Fichera M, Romano C, Delle Chiaie B, Mortier G, Menten B, et al. Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome. *Eur J Med Genet*. 2009; 52:94–100.

[155] Turner DJ, Miretti M, Rajan D, Fiegler H, Carter NP, Blayney ML, Beck S, Hurles ME. Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet*. 2008; 40:90–95.

[156] Ensenauer RE, Adeyinka A, Flynn HC, Michels VV, Lindor NM, Dawson DB, Thorland EC, Lorentz CP, Goldstein JL, McDonald MT, et al. Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet*. 2003; 73:1027–1040.

[157] Somerville MJ, Mervis CB, Young EJ, Seo EJ, del Campo M, Bamforth S, Peregrine E, Loo W, Lilley M, Perez-Jurado LA, et al. Severe expressive-language delay related to duplication of the Williams-Beuren locus. *N Engl J Med*. 2005; 353:1694–1701.

[158] Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboa K, Kashork CD, et al. Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet*. 2000; 24:84–87.

[159] Franco LM, de Ravel T, Graham BH, Frenkel SM, Van Driessche J, Stankiewicz P, Lupski JR, Vermeesch JR, Cheung SW. A syndrome of short stature, microcephaly and speech delay is associated with duplications reciprocal to the common Sotos syndrome deletion. *Eur J Hum Genet*. 2010; 18:258–261.

[160] Ballif BC, Theisen A, Coppinger J, Gowans GC, Hersh JH, Madan-Khetarpal S, Schmidt KR, Tervo R, Escobar LF, Friedrich CA, et al. Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Mol Cytogenet*. 2008; 1:8.

[161] Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med*. 2008; 358:667–675.

[162] Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z, Martinet D, Shen Y, Valsesia A, Beckmann ND, et al. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature*. 2011; 478:97–102.

[163] Tatton-Brown K, Rahman N. Clinical features of NSD1-positive Sotos syndrome. *Clin Dysmorphol*. 2004; 13:199–204.

[164] Sotos JF, Dodge PR, Muirhead D, Crawford JD, Talbot NB. Cerebral Gigantism in Childhood. A Syndrome of Excessively Rapid Growth and Acromegalic

Features and a Nonprogressive Neurologic Disorder. *N Engl J Med*. 1964; 271:109–116.
[165] Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y, et al. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat Genet*. 2002; 30:365–366.
[166] Kurotaki N, Stankiewicz P, Wakui K, Niikawa N, Lupski JR. Sotos syndrome common deletion is mediated by directly oriented subunits within inverted Sos-REP low-copy repeats. *Hum Mol Genet*. 2005; 14:535–542.
[167] Hunter AG, McAlpine PJ, Rudd NL, Fraser FC. A 'new' syndrome of mental retardation with characteristic facies and brachyphalangy. *J Med Genet*. 1977; 14:430–437.
[168] Hunter AG, Dupont B, McLaughlin M, Hinton L, Baker E, Ades L, Haan E, Schwartz CE. The Hunter-McAlpine syndrome results from duplication 5q35-qter. *Clin Genet*. 2005; 67:53–60.
[169] Kirchhoff M, Bisgaard AM, Bryndorf T, Gerdes T. MLPA analysis for a panel of syndromes with mental retardation reveals imbalances in 5.8% of patients with mental retardation and dysmorphic features, including duplications of the Sotos syndrome and Williams-Beuren syndrome regions. *Eur J Med Genet*. 2007; 50:33–42.
[170] Chen CP, Lin SP, Lin CC, Chen YJ, Chern SR, Li YC, Hsieh LJ, Lee CC, Pan CW, Wang W. Molecular cytogenetic analysis of de novo dup(5)(q35.2q35.3) and review of the literature of pure partial trisomy 5q. *Am J Med Genet A*. 2006; 140:1594–1600.
[171] Zhang H, Lu X, Beasley J, Mulvihill JJ, Liu R, Li S, Lee JY. Reversed clinical phenotype due to a microduplication of Sotos syndrome region detected by array CGH: microcephaly, developmental delay and delayed bone age. *Am J Med Genet A*. 2011; 155A:1374–1378.
[172] Rosenfeld JA, Kim KH, Angle B, Troxell R, Gorski JL, Westemeyer M, Frydman M, Senturias Y, Earl D, Torchia B, et al. Further Evidence of Contrasting Phenotypes Caused by Reciprocal Deletions and Duplications: Duplication of NSD1 Causes Growth Retardation and Microcephaly. *Mol Syndromol*. 2013; 3:247–254.
[173] Dikow N, Maas B, Gaspar H, Kreiss-Nachtsheim M, Engels H, Kuechler A, Garbes L, Netzer C, Neuhann TM, Koehler U, et al. The phenotypic spectrum of duplication 5q35.2-q35.3 encompassing NSD1: is it really a reversed Sotos syndrome? *Am J Med Genet A*. 2013; 161:2158–2166.
[174] Tatton-Brown K, Douglas J, Coleman K, Baujat G, Cole TR, Das S, Horn D, Hughes HE, Temple IK, Faravelli F, et al. Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. *Am J Hum Genet*. 2005; 77:193–204.
[175] Kurotaki N, Harada N, Shimokawa O, Miyake N, Kawame H, Uetake K, Makita Y, Kondoh T, Ogata T, Hasegawa T, et al. Fifty microdeletions among 112 cases of Sotos syndrome: low copy repeats possibly mediate the common deletion. *Hum Mutat*. 2003; 22:378–387.
[176] Migdalska AM, van der Weyden L, Ismail O, Rust AG, Rashid M, White JK, Sanchez-Andrade G, Lupski JR, Logan DW, Arends MJ, et al. Generation of the Sotos syndrome deletion in mice. *Mamm Genome*. 2012; 23:749–757.
[177] Papenhausen P, Schwartz S, Risheg H, Keitges E, Gadi I, Burnside RD, Jaswaney V, Pappas J, Pasion R, Friedman K, et al. UPD detection using

homozygosity profiling with a SNP genotyping microarray. *Am J Med Genet A*. 2011; 155A:757–768.

[178] Sund KL, Zimmerman SL, Thomas C, Mitchell AL, Prada CE, Grote L, Bao L, Martin LJ, Smolarek TA. Regions of homozygosity identified by SNP microarray analysis aid in the diagnosis of autosomal recessive disease and incidentally detect parental blood relationships. *Genet Med*. 2013; 15:70–78.

[179] Piotrowski A, Bruder CE, Andersson R, Diaz de Stahl T, Menzel U, Sandgren J, Poplawski A, von Tell D, Crasto C, Bogdan A, et al. Somatic mosaicism for copy number variation in differentiated human tissues. *Hum Mutat*. 2008; 29:1118–1124.

[180] Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, Diaz de Stahl T, Menzel U, Sandgren J, von Tell D, Poplawski A, et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Hum Genet*. 2008; 82:763–771.

[181] Ballif BC, Rorem EA, Sundin K, Lincicum M, Gaskin S, Coppinger J, Kashork CD, Shaffer LG, Bejjani BA. Detection of low-level mosaicism by array CGH in routine diagnostic specimens. *Am J Med Genet A*. 2006; 140:2757–2767.

[182] Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. *Hum Mol Genet*. 2006; 15:789–795.

[183] Li LH, Ho SF, Chen CH, Wei CY, Wong WC, Li LY, Hung SI, Chung WH, Pan WH, Lee MT, et al. Long contiguous stretches of homozygosity in the human genome. *Hum Mutat*. 2006; 27:1115–1121.

[184] Tucker T, Schlade-Bartusiak K, Eydoux P, Nelson TN, Brown L. Uniparental disomy: can SNP array data be used for diagnosis? *Genet Med*. 2012; Apr 26. doi: 10.1038/gim.2012.35. [Epub ahead of print].

[185] Eggermann T, Spengler S, Gogiel M, Begemann M, Elbracht M. Epigenetic and genetic diagnosis of Silver-Russell syndrome. *Expert Rev Mol Diagn*. 2012; 12:459–471.

[186] Robinson WP. Mechanisms leading to uniparental disomy and their clinical consequences. *Bioessays*. 2000; 22:452–459.

[187] Olander E, Stamberg J, Steinberg L, Wulfsberg EA. Third Prader-Willi syndrome phenotype due to maternal uniparental disomy 15 with mosaic trisomy 15. *Am J Med Genet*. 2000; 93:215–218.

[188] Yamazawa K, Ogata T, Ferguson-Smith AC. Uniparental disomy and human disease: an overview. *Am J Med Genet C Semin Med Genet*. 2010; 154C:329–334.

[189] Zilina O, Reimand T, Zjablovskaja P, Mannik K, Mannamaa M, Traat A, Puusepp-Benazzouz H, Kurg A, Ounap K. Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. *Am J Med Genet A*. 2012; 158A:254–256.

[190] Bayley N. Bayley Scales of Infant and Toddler Development. 3rd Edition (Bayley-III). 2005.

[191] Kaufman A, Kaufman N. KAIT: Kaufman Adolescents and Adult Intelligence Test. 1993.

[192] Zilina O, Reimand T, Tammur P, Tillmann V, Kurg A, Ounap K. Patient with dup(5)(q35.2-q35.3) reciprocal to the common Sotos syndrome deletion and review of the literature. *Eur J Med Genet*. 2013; 56:202–206.

[193] Zilina O, Kahre T, Talvik I, Oiglane-Shlik E, Tillmann V, Ounap K. Mosaicism for maternal uniparental disomy 15 in a boy with some clinical features of Prader-Willi syndrome. *Eur J Med Genet*. 2014; 57:279–283.

[194] Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, Bassett AS, Seller A, Holmes CC, Ragoussis J. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res*. 2007; 35:2013–2025.

[195] Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet*. 2009; 84:524–533.

[196] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001; 29:e45.

[197] Hoebeeck J, van der Luijt R, Poppe B, De Smet E, Yigit N, Claes K, Zewald R, de Jong GJ, De Paepe A, Speleman F, et al. Rapid detection of VHL exon deletions using real-time quantitative PCR. *Lab Invest*. 2005; 85:24–33.

[198] Junier T, Pagni M. Dotlet: diagonal plots in a web browser. *Bioinformatics*. 2000; 16:178–179.

[199] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002; 3:RESEARCH0034.

[200] Beaudet AL. The utility of chromosomal microarray analysis in developmental and behavioral pediatrics. *Child Dev*. 2013; 84:121–132.

[201] Sagoo GS, Butterworth AS, Sanderson S, Shaw-Smith C, Higgins JP, Burton H. 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*. 2009; 11:139–146.

[202] Battaglia A, Doccini V, Bernardini L, Novelli A, Loddo S, Capalbo A, Filippi T, Carey JC. Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *Eur J Paediatr Neurol*. 2013.

[203] Ullmann R, Turner G, Kirchhoff M, Chen W, Tonge B, Rosenberg C, Field M, Vianna-Morgante AM, Christie L, Krepischi-Santos AC, et al. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat*. 2007; 28:674–682.

[204] Burnside RD, Pasion R, Mikhail FM, Carroll AJ, Robin NH, Youngs EL, Gadi IK, Keitges E, Jaswaney VL, Papenhausen PR, et al. Microdeletion/ micro-duplication of proximal 15q11.2 between BP1 and BP2: a susceptibility region for neurological dysfunction including developmental and language delay. *Hum Genet*. 2011; 130:517–528.

[205] Girirajan S, Rosenfeld JA, Coe BP, Parikh S, Friedman N, Goldstein A, Filipink RA, McConnell JS, Angle B, Meschino WS, et al. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N Engl J Med*. 2012; 367:1321–1331.

[206] De Wolf V, Brison N, Devriendt K, Peeters H. Genetic counseling for susceptibility loci and neurodevelopmental disorders: the del15q11.2 as an example. *Am J Med Genet A*. 2013; 161A:2846–2854.

[207] Zilina O, Teek R, Tammur P, Kuuse K, Yakoreva M, Vaidla E, Molter-Vaar T, Reimand T, Kurg A, Ounap K. Chromosomal microarray analysis as a first-tier clinical diagnostic test: Estonian experience. *Mol Genet Genomic Med*. 2014; 2:166–175.

[208] Abuelo DN, Padre-Mendoza T. Cat-like cry and mental retardation owing to 7q interstitial deletion (7q22 leads to 7q32). *J Med Genet*. 1982; 19:473–476.
[209] Ayraud N, Rovinski J, Lambert JC, Galiana A. [Interstitial deletion of the long arm of chromosome 7 in a female child with leprechaunism]. *Ann Genet*. 1976; 19:265–268.
[210] Cheong ML, Tsai MS, Cortes RA, Harrison MR. Intermediate interstitial deletion of chromosome 7q detected by first-trimester Down's syndrome screening. *Fetal Diagn Ther*. 2008; 24:340–344.
[211] Fagan K, Gill A, Henry R, Wilkinson I, Carey B. A summary of 7q interstitial deletions and exclusion mapping of the gene for beta-glucuronidase. *J Med Genet*. 1989; 26:619–625.
[212] Franceschini P, Silengo MC, Davi GF, Santoro MA, Prandi G, Fabris C. Interstitial deletion of the long arm of chromosome 7 46,XX,del(7)(pter leads to q2200::q3200 leads to qter). *Hum Genet*. 1978; 44:345–348.
[213] Higginson G, Weaver DD, Magenis RE, Prescott GH, Haag C, Hepburn DJ. Interstitial deletion of the long arm of chromosome no. 7 (7q–) in an infant with multiple anomalies. *Clin Genet*. 1976; 10:307–312.
[214] Klep-de Pater JM, Bijlsma JB, Bleeker-Wagemakers EM, de France HF, de Vries-Ekkers CM. Two cases with different deletions of the long arm of chromosome 7. *J Med Genet*. 1979; 16:151–154.
[215] Martin-Pont B, Pilczer C, Dandine M, Tamboise A. De novo interstitial deletion of the long arm of chromosome 7:46,XY,del(7)(q23;q32). *Ann Genet*. 1985; 28:251–253.
[216] Montgomery TL, Wyllie J, Oley C. Ectrodactyly and glaucoma associated with a 7q21.2-q31.2 interstitial deletion. *Clin Dysmorphol*. 2000; 9:235–239.
[217] Morey MA, Higgins RR. Ectro-amelia syndrome associated with an interstitial deletion of 7q. *Am J Med Genet*. 1990; 35:95–99.
[218] Sarda P, Turleau C, Cabanis MO, Jalaguier J, de Grouchy J, Bonnet H. [Interstitial deletion in the long arm of chromosome 7]. *Ann Genet*. 1988; 31:258–261.
[219] Serup L. Interstitial deletion of the long arm of chromosome 7. *Hum Genet*. 1980; 54:19–23.
[220] Stallard R, Juberg RC. Partial monosomy 7q syndrome due to distal interstitial deletion. *Hum Genet*. 1981; 57:210–213.
[221] Young RS, Weaver DD, Kukolich MK, Heerema NA, Palmer CG, Kawira EL, Bender HA. Terminal and interstitial deletions of the long arm of chromosome 7: a review with five new cases. *Am J Med Genet*. 1984; 17:437–450.
[222] Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet*. 2004; 13 Spec No 1:R57–64.
[223] Trifonov EN. Curved DNA. *CRC Crit Rev Biochem*. 1985; 19:89–106.
[224] Thomas AC, Frost JM, Ishida M, Vargha-Khadem F, Moore GE, Stanier P. The speech gene FOXP2 is not imprinted. *J Med Genet*. 2012; 49:669–670.
[225] Camarata T, Bimber B, Kulisz A, Chew TL, Yeung J, Simon HG. LMP4 regulates Tbx5 protein subcellular localization and activity. *J Cell Biol*. 2006; 174:339–348.
[226] Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, Reymond A, Sun M, Sawa A, Gusella JF, et al. KCTD13 is a major driver of

mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature*. 2012; 485:363–367.
[227] Horsthemke B, Nazlican H, Husing J, Klein-Hitpass L, Claussen U, Michel S, Lich C, Gillessen-Kaesbach G, Buiting K. Somatic mosaicism for maternal uniparental disomy 15 in a girl with Prader-Willi syndrome: confirmation by cell cloning and identification of candidate downstream genes. *Hum Mol Genet*. 2003; 12:2723–2732.
[228] Izumi K, Santani AB, Deardorff MA, Feret HA, Tischler T, Thiel BD, Mulchandani S, Stolle CA, Spinner NB, Zackai EH, et al. Mosaic maternal uniparental disomy of chromosome 15 in Prader-Willi syndrome: utility of genome-wide SNP array. *Am J Med Genet A*. 2013; 161A:166–171.
[229] Liehr T. Cases with uniparental disomy. http://www.fish.uniklinikum-jena.de/UPD.html. 2013.
[230] Wang JC, Vaccarello-Cruz M, Ross L, Owen R, Pratt VM, Lightman K, Liu Y, Hafezi K, Cherif D, Sahoo T. Mosaic isochromosome 15q and maternal uniparental isodisomy for chromosome 15 in a patient with morbid obesity and variant PWS-like phenotype. *Am J Med Genet A*. 2013; 161A:1695–1701.
[231] Howell KB, Kornberg AJ, Harvey AS, Ryan MM, Mackay MT, Freeman JL, Rodriguez Casero MV, Collins KJ, Hayman M, Mohamed A, et al. High resolution chromosomal microarray in undiagnosed neurological disorders. *J Paediatr Child Health*. 2013.
[232] Bradley WE, Raelson JV, Dubois DY, Godin E, Fournier H, Prive C, Allard R, Pinchuk V, Lapalme M, Paulussen RJ, et al. Hotspots of large rare deletions in the human genome. *PLoS One*. 2010; 5:e9401.
[233] Kapoor A, Satishchandra P, Ratnapriya R, Reddy R, Kadandale J, Shankar SK, Anand A. An idiopathic epilepsy syndrome linked to 3q13.3-q21 and missense mutations in the extracellular calcium sensing receptor gene. *Ann Neurol*. 2008; 64:158–167.
[234] Ferrier RA, Lowry RB, Lemire EG, Stoeber GP, Howard J, Parboosingh JS. Father-to-son transmission of an X-linked gene: a case of paternal sex chromosome heterodisomy. *Am J Med Genet A*. 2009; 149A:2871–2873.
[235] de Jong A, Dondorp WJ, Macville MV, de Die-Smulders CE, van Lith JM, de Wert GM. Microarrays as a diagnostic tool in prenatal screening strategies: ethical reflection. *Hum Genet*. 2014; 133:163–172.
[236] Fiorentino F, Napoletano S, Caiazzo F, Sessa M, Bono S, Spizzichino L, Gordon A, Nuccitelli A, Rizzo G, Baldi M. Chromosomal microarray analysis as a first-line test in pregnancies with a priori low risk for the detection of submicroscopic chromosomal abnormalities. *Eur J Hum Genet*. 2013; 21:725–730.
[237] Hillman SC, Pretlove S, Coomarasamy A, McMullan DJ, Davison EV, Maher ER, Kilby MD. Additional information from array comparative genomic hybridization technology over conventional karyotyping in prenatal diagnosis: a systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2011; 37:6–14.
[238] Donnelly JC, Platt LD, Rebarber A, Zachary J, Grobman WA, Wapner RJ. Association of Copy Number Variants With Specific Ultrasonographically Detected Fetal Anomalies. *Obstet Gynecol*. 2014; Jun 4. [Epub ahead of print].

WEB RESOURCES

Database of Genomic Variants – a curated catalogue of human genomic structural variations: http://dgv.tcag.ca/

DECIPHER – Database of chromosomal imbalance and phenotype in humans using Ensembl resources: http://decipher.sanger.ac.uk/

Ensembl Genome Browser: http://www.ensembl.org

Illumina, Inc. – a company that develops and applies innovative sequencing and array technologies to the analysis of genetic variation and function: http://www.illumina.com

OMIM – Online Mendelian Inheritance in Man database: http://www.ncbi.nlm.nih.gov/omim

PubMed – a repository of peer-reviewed primary research reports in life sciences: http://www.ncbi.nlm.nih.gov/pubmed

Primer-BLAST – online tool for primer design: http://www.ncbi.nlm.nih.gov/tools/primer-blast/

qRTDesigner 1.2 – a program for designing primers for real-time quantitative PCR assays: http://bioinfo.ut.ee/gwRTqPCR/

UCSC Genome Browser: http://genome.ucsc.edu/

SUMMARY IN ESTONIAN

Vaimse ja füüsilise arengu mahajäämuse, kaasasündinud anomaaliate ja neuroloogiliste haiguste tekkepõhjuseks võivad olla kromosomaalsed aberratsioonid. Kromosomaalne mikrokiibianalüüs (ingl. k. *chromosomal microarray analysis*; KMA) on uus molekulaarne kromosoomide uurimismeetod, mis võimaldab analüüsida tervet genoomi ühe eksperimendi käigus tuvastades väga väikesi kromosomaalseid muutusi ning pakkudes seejuures kõrgemat lahutusvõimet võrreldes tavapäraste tsütogeneetiliste meetoditega. Sellest tulenevalt tõuseb oluliselt diagnostiline efektiivsus. Tänu sellele soovitatakse kasutada KMA esmase tsütogeneetilise testina kliinilises diagnostikas kromosomaalsete aberratsioonide kiireks ja täpseks tuvastamiseks füüsilise ja vaimse arengu mahajäämuse, kaasasündinud väärarengute ja/või autismi spektri häiretega patsientidel.

Käesoleva uuringu eesmärkideks oli:

1. Hinnata KMA kasutamist esmase tsütogeneetilise testina rutiinses kliinilises praktikas Eestis intellektipuude, arengu mahajäämuse, kaasasündinud väärarengute ja autismi spektri häiretega patsientidel.
2. Hinnata KMA tõhusust sünnieelses diagnostikas.
3. Hinnata kõne ja keele probleemidega seotud *FOXP2* geeni hõlmava 7q31 deletsiooni kliinilist avaldumist ning määrata kas kõne ja keele probleemide raskusaste sõltub deletsiooni vanemlikust päritolust.
4. Uurida Sotose sündroomi põhjustava deletsiooni retsiprookset 5q35.2-q35.3 duplikatsiooni kliinilisi tagajärgi.
5. Kirjeldada KMA abil mosaiikse emapoolse 15. kromosoomi uniparentaalse disoomiaga (ingl. k. *uniparental disomy; UPD*) patsienti ning esitada tööskeem vermimise häirete diagnoosimiseks.

Uuringu esimeseks eesmärgiks oli hinnata KMA kasulikkust kromosomaalsete muutuste detekteerimiseks rutiinses kliinilises praktikas. Selleks analüüsiti 2009.–2012. a. Eestis uuritud patsientide andmeid (kokku 1191). Kliiniliselt olulised leiud tuvastati 11%-l patsientidest, mis on võrreldav varasemalt publitseeritud uuringute tulemustega. Käesolevas töös leiti, et KMA kasutamine varasemalt hoolikalt uuritud kuid etioloogilise põhjuseta jäänud patsientide hulgas annab suurema diagnostilise efektiivsuse (15%), samas kui KMA kasutamine esmase tsütogeneetilise testina heterogeensemas populatsioonis tuvastab haiguse põhjuse 10%-l patsientidel. Siiski on KMA diagnostiline ulatus vähemalt 2–3 korda suurem võrreldes klassikalise karüotüpiseerimisega. Samal ajal, jäi ebaselge kliinilise tähendusega KMA leidude osakaal antud uuringugrupis suureks (41% kõigist raporteeritud leidudest, k.a. pikad homosügootsed alad (ingl. k. *long contiguous stretches of homozygosity; LCSH*); ebaselgete leidude hulk oli 24% kui *LCSH* on välja jäetud). Varem on leitud ning ka käesolev uuring kinnitab, et KMA tulemuste interpreteerimine võib muutuda aja jooksul

paremaks ning seda tänu andmebaaside ja kirjanduseandmete pidevale täienemisele. Siiski peab arvestama sellega, et ebaselge kliinilise tähendusega tulemuse raporteerimine patsiendile ja tema perekonnale võib põhjustada neile tõsise stressi ja ebakindluse tunde. Selleks et tõlgendada maksimaalset arvu KMA leidudest ning tagada patsientidele korralikku geneetilist nõustamist, on vajalik tihe koostöö tsütogeneetikute ja kliiniliste geneetikute vahel.

Teiseks on antud töös esitatud Eesti kogemus KMA rakendamise kohta sünnieelses diagnostikas. Enamikul juhtudel teostati KMA paralleelselt tavapärase karüotüpiseerimisega, sest KMA ei ole soovitatav kasutada esmase tsütogeneetilise testina sünnieelses diagnostikas, vaid pigem täiendava meetodina. KMA tuvastas kõik tasakaalustamata kromosomaalsed aberratsioonid, mis olid nähtavad ka traditsioonilisel karüotüpiseerimisel. Lisaks tuvastas KMA patoloogilisi muutusi, mis ei olnud detekteeritavad rutiinse karüotüpiseerimise meetodil, pakkudes seega uusi diagnostilisi lahendusi sünnieelse diagnostika valdkonnas. Siiski, suhteliselt väikese patsientide kohordi tõttu (60 prenataalset juhtumit) ei olnud võimalik teha kaalukaid järeldusi KMA rakendamise kohta esmase tsütogeneetilise testina sünnieelses diagnostikas. Kuigi hiljuti ilmusid kirjanduses uued juhtnöörid, kus soovitatakse valida *K*MA esmaseks testiks ebanormaalse ultraheli leiuga raseduste puhul, põhjustab KMA kasutamine sünnieelses diagnostikas siiski palju vaidlusi seoses sobiliku mikrokiibi platvormi valikuga, ebaselge kliinilise tähendusega leidudega kaasnevate probleemidega ning ka sellega, kellele oleks mõttekas seda analüüsi pakkuda ning kellele mitte.

Kolmandaks, kirjeldati antud töös kahte kõne ja keele probleemidega ning *FOXP2* geeni hõlmava 7q31 deletsiooniga perekonda. Eelnevalt oli kirjanduses pakutud, et sõltuvalt vanemlikust päritolust võib *FOXP2* lookus olla erinevalt reguleeritud. Isapoolse alleeli kahjustus võiks põhjustada raskemaid kõne ja keele probleeme, samas kui emapoolse alleeli kahjustus väljendub kergema kliinilise pildina. Kirjeldatud perekonnad andsid võimaluse hinnata potentsiaalseid vanemliku päritoluga seotuid erinevusi uuritavate perekondade liikmete fenotüübis. Detailne kliiniline kirjeldus ning kõne ja keele hindamine võimaldas järeldada, et nii isapoolse kui ka emapoolse alleeli kahjustus põhjustab tõsist ning kliiniliselt samasuguse raskusastmega kõne ja keele häiret.

Neljandaks, kirjeldati Sotose sündroomi põhjustava deletsiooni retsiprookse 5q35.2-q35.3 duplikatsiooniga patsienti, kes laiendas teadmisi retsiprooksete deletsioonidega-duplikatsioonidega seotud fenotüüpide kohta. Huvitav on see, et 5q35.2-q35.3 deletsioonide ja duplikatsioonide poolt põhjustatud kliiniline fenotüüp on nn. „peegelfenotüüp" keha parameetrite osas nagu pikkus ja pea ümbermõõt. Kuigi on teada, et 5q35.2-q35.3 regioonis paiknev *NSD1* geen võiks olla vastutav liigkasvu/väikese kasvu eest, on siiski täpne neid fenotüübilisi muutusi põhjustav molekulaarne mehhanism teadmata. Samuti ei ole teada, kas teised samas kromosomaalses piirkonnas asetsevad geenid võiksid mängida rolli kliinilise fenotüübi välja kujunemisel. Me uurisime Sotose sündroomi kriitilises regioonis ja naaberregioonides paiknevate geenide

ekspressiooni kolmel 5q35.2-q35.3 duplikatsiooniga patsiendil ning leidsime, et peaaegu kõik uuritud geenid olid üleekspresseeritud. Samas, uuringu disain ja väike patsientide arv ei võimalda teha statistiliselt olulisi järeldusi. Selge on see, et madala sagedusega genoomsete aberratsioonide puhul nende fenotüübilise efekti kirjeldamiseks ja toimemehhanismide välja selgitamiseks on vajalik koostöö erinevate keskuste vahel.

Lõpuks, kirjeldati mosaiikse emapoolse 15. kromosoomi *UPD*-ga patsienti, kes on heaks näideks SNP-põhise KMA võimest tuvastada uniparentaalset disoomiat ja mosaiiksust. Patsiendil leiti mosaiikne 15. kromosoomi iso-disoomia 55–60% perifeerse vere rakkudes, samas kui ülejäänud rakud olid normaalse kromosoomide komplektiga. Põhinedes KMA, karüotüpiseerimise ja metülatsiooni analüüsile pakuti ka *UPD* tekkimise eest vastutav mehhanism. Siiani on kirjeldatud vaid kolme Prader-Willi sündroomiga (PWS) patsienti, kellel esineb mosaiikne emapoolne isodisoomia 15 ilma trisoomsete/monosoomsete rakuliinideta. Siiski jääb ebaselgeks, kas mosaiikne *UPD* on tõesti niivõrd harv või jääb osalise PWS fenotüübi tõttu ning ka sobilike diagnostiliste vahendite puudumise või tulemuste interpretatsiooniga seotud raskuste tagajärjel suurem osa patsientidest ilma diagnoosita. Kuigi on välja töötatud terve hulk erinevaid meetodeid vermimishäirete diagnoosimiseks, ei võimalda mitte ükski neist tuvastada kõiki võimalikke mutatsioone ja epimutatsioone. Põhinedes meie tulemustele, mis saadi kasutades KMA ja kvantitatiivset metülatsiooni analüüsi, järeldasime, et nende kahe meetodi kombinatsioon sobib hästi vermimishäirete diagnoosimiseks.

Kokkuvõtteks, näitab meie kogemus, et KMA on usaldusväärne ja efektiivne vahend arenguhäiretega seotud kromosomaalsete aberratsioonide tuvastamiseks ning selle kasutamine esmase tsütogeneetilise testina kliinilises praktikas on täiesti õigustatud. Siiski on vaja meeles pidada, et KMA tulemuste interpretatsioon põhineb meie praegustel teadmistel, mis võivad aja jooksul muutuda, mistõttu on tulevikus tulemuste ülevaatamine väga oluline ja jääb raviarsti ülesandeks.

ACKNOWLEDGEMENTS

Feeling gratitude and not expressing it is like wrapping a present and not giving it.

William Ward

I would like to express my deepest gratitude to my supervisors: Professor Ants Kurg who introduced me to the field of molecular genetics, helped to survive there and feel at home, who always supported and believed in me; and Professor Katrin Õunap who let me apply practically my knowledge and modest experience by joining clinical cytogenetic laboratory and have the opportunity to see how useful could become research study results for real people, who finally encouraged me to start and complete this study, and always offered her help and support whenever needed.

The major part of this study was performed in the Department of Genetics, in Tartu University Hospital and I would like to thank my colleagues – Tiiu Ilus, Piret Ilisson, Sander Pajusalu, Mare Jürgenson, Peeter Toit, Tiina Kahre, Ülle Murumets, Neeme Tõnisson, Heete Kärmas – for creating very warm and supportive atmosphere: My special thanks belong to Kati Kuuse and Pille Tammur for their interest, firm support, sharing their knowledge in the field of cytogenetics and readiness to discuss the most complicated and interesting patients. I warmly thank all clinical geneticists – Tiia Reimand, Kai Muru, Rita Teek, Riina Žordania, Kairit Joost, Elvira Kurvinen – for acting successfully as a go-between for the patients and laboratory staff.

I wish to thank all my colleagues from the Institute of Molecular and Cell Biology, especially Katrin Männik, Margit Nõukas, Merle Külaots, Sven Parkel, and Ott Scheler. My warmest thanks go to Olga Tšuiko for our lunch-time and tea-time discussions, for her inspiring curiosity and hundreds of questions.

I would like to thank Professor Andres Metspalu and people from Estonian Genome Center, especially Reedik Mägi for his help in *in silico* world. Viljo Soo who provided the most technical support for this study deserves and has my special thanks.

I am also deeply grateful to Professor Pärt Peterson for all his valuable suggestions on reviewing the thesis and David Schryer for revising the English text of the thesis while doing this very delicately trying to keep the original ideas and style.

Finally, my most loving thanks go to my family for their understanding, encouragement and love. I am deeply grateful to my mother who in critical situations took my urgencies before her own, to my father for being there, and to my sweet Natalie who believes I am the best mother in the world and thus making me to become better than I am.

PUBLICATIONS

CURRICULUM VITAE

Name: Olga Žilina
Date of Birth: February 9th 1983
Address: Institute of Molecular and Cell Biology
Riia 23, Tartu 51010
E-mail: olga.zhilina@ut.ee

Education

Since 2007 PhD studies in molecular diagnostics, Faculty of Science and Technology, University of Tartu, Estonia
2005–2007 MSc in molecular diagnostics, Faculty of Biology and Geography, University of Tartu, Estonia
2001–2005 BSc in molecular diagnostics, Faculty of Biology and Geography, University of Tartu, Estonia

Working Experience

Since 2011 Researcher, Institute of Molecular and Cell Biology, University of Tartu, Estonia
Since 2010 Specialist, Department of Genetics, United Laboratories, Tartu University Hospital, Tartu, Estonia
2010–2011 Associate researcher, Institute of Molecular and Cell Biology, University of Tartu, Estonia
2007–2009 Associate researcher, Estonian Biocentre

Additional Professional Training

2012 Genomic Disorders 2012: The Genomics of Rare Diseases; Hinxton, United Kingdom
2011 6th Goldrain Course in Clinical Cytogenetics; Goldrain, Italy
2008 1st AnEUploidy Workshop; Geneva, Switzerland
2008 2nd International Summer School in Functional Genomics; Copenhagen, Denmark
2008 4th International Meeting on Cryptic Chromosomal Rearrangements in Mental Retardation and Autism; Sicilia, Italy
2006 EU FP6 Marie Curie Conferences and Training Courses on arrayCGH and Molecular Cytogenetics, „2nd Workshop on arrayCGH and Molecular Cytogenetics"; Gent, Belgium
2005 European School of Genetic Medicine „7th Course in Molecular Cytogenetics and DNA Arrays"; Bertinoro, Italy

Honour and Fellowships

2011	European Society of Human Genetics; scholarship for participating in the 6th Goldrain Course in Clinical Cytogenetics
2008	European Society of Human Genetics; national fellowship for young investigator
2007	II Prize for MSc Thesis, Estonian National Contest for Young Scientists at university level
2005	Scientific Committe of 7th Course in Molecular Cytogenetics and DNA Arrays; best poster presentation award

Other Scientific Activities

Since 2013	Member of the European Society of Human Genetics
Since 2007	Member of the Estonian Society of Human Genetics

Publications

Žilina O., Kahre T., Talvik I., Õiglane-Šlik E., Tillmann V., Õunap K. (2014). Mosaicism for maternal uniparental disomy 15 in a boy with some clinical features of Prader-Willi syndrome. Eur J Med Genet. Published online April 2.

Žilina O., Teek R., Tammur P., Kuuse K., Yakoreva M., Vaidla E., Mölter-Väär T., Reimand T., Kurg A., Õunap K. (2014). Chromosomal microarray analysis as a first-tier clinical diagnostic test: Estonian experience. Mol Genet Genomic Med. 2(2): 166–175.

Oiglane-Shlik E., Puusepp S., Talvik I., Vaher U., Rein R., Tammur P., Reimand T., Teek R., Zilina O., Tomberg T., Ounap K. (2014). Monosomy 1p36 – A multifaceted and still enigmatic syndrome: Four clinically diverse cases with shared white matter abnormalities. Eur J Paediatr Neurol. 18(3):338–346

Teek R., Kruustük K., Žordania R., Joost K., Kahre T., Tõnisson N., Nelis M., Zilina O., Tranebjaerg L., Reimand T., Ounap K. (2013). Hearing impairment in Estonia: an algorithm to investigate genetic causes in pediatric patients. Adv Med Sci. 58(2):419–428.

Žilina O., Reimand T., Tammur P., Tillmann V., Kurg A., Õunap K. (2013). Patient with dup(5)(q35.2-q35.3) reciprocal to the common Sotos syndrome deletion and review of the literature. Eur J Med Genet. 56(4):202–206.

Zilina O., Reimand T., Zjablovskaja P., Männik K., Männamaa M., Traat A., Puusepp-Benazzouz H., Kurg A., Ounap K. (2012). Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. Am J Med Genet A. 158A(1):254–256.

Mikelsaar R, Nelis M, Kurg A, Zilina O, Korrovits P, Rätsep R, Väli M. (2012) Balanced reciprocal translocation t(5;13)(q33;q12) and 9q31.1 microduplication in a man suffering from infertility and pollinosis. Appl Genet. 53(1):93–97

Joost K., Tammur P., Teek R., Zilina O., Peters M., Kreile M., Lace B., Zordania R., Talvik I., Ounap K. (2011). Whole Xp Deletion in a Girl with Mental Retardation, Epilepsy, and Biochemical Features of OTC Deficiency. Mol Syndromol. 1(6):311–315

Sismani C., Anastasiadou V., Kousoulidou L., Parkel S., Koumbaris G., Zilina O., Bashiardes S., Spanou E., Kurg A., Patsalis PC. (2011). 9 Mb familial duplication in chromosome band Xp22.2-22.13 associated with mental retardation, hypotonia and developmental delay, scoliosis, cardiovascular problems and mild dysmorphic facial features. Eur J Med Genet. 54(5):e510–515

Männik K., Parkel S., Palta P., Zilina O., Puusepp H., Esko T., Mägi R., Nõukas M., Veidenberg A., Nelis M., Metspalu A., Remm M., Ounap K., Kurg A. (2011). A parallel SNP array study of genomic aberrations associated with mental retardation in patients and general population in Estonia. Eur J Med Genet. 54(2):136–143.

Puusepp H., Zilina O., Teek R., Männik K., Parkel S., Kruustük K., Kuuse K., Kurg A., Ounap K. (2009). 5.9 Mb microdeletion in chromosome band 17q22-q23.2 associated with tracheo-esophageal fistula and conductive hearing loss. Eur J Med Genet. 52(1):71–74

Kousoulidou L., Männik K., Zilina O., Parkel S., Palta P., Remm M., Kurg A., Patsalis PC. (2008). Application of two different microarray-based copy-number detection methodologies – array-comparative genomic hybridization and array-multiplex amplifiable probe hybridization – with identical amplifiable target sequences. Clin Chem Lab Med. 46(5):722–724

Kousoulidou L., Männik K., Sismani C., Zilina O., Parkel S., Puusepp H., Tõnisson N., Palta P., Remm M., Kurg A., Patsalis PC. (2008). Array-MAPH: a methodology for the detection of locus copy-number changes in complex genomes. Nat Protoc. 3(5):849–865

Kousoulidou L., Parkel S., Zilina O., Palta P., Puusepp H., Remm M., Turner G., Boyle J., van Bokhoven H., de Brouwer A., Van Esch H., Froyen G., Ropers HH., Chelly J., Moraine C., Gecz J., Kurg A., Patsalis PC. (2007). Screening of 20 patients with X-linked mental retardation using chromosome X-specific array-MAPH. Eur J Med Genet. 50(6):399–410

Patsalis PC., Kousoulidou L., Männik K., Sismani C., Zilina O., Parkel S., Puusepp H., Tõnisson N., Palta P., Remm M., Kurg A. (2007). Detection of small genomic imbalances using microarray-based multiplex amplifiable probe hybridization. Eur J Hum Genet. 15(2):162–172

ELULOOKIRJELDUS

Nimi:	Olga Žilina
Sünniaeg:	9. veebruar 1983
Aadress:	Molekulaar- ja rakubioloogia instituut Riia 23, Tartu 51010
E-post:	olga.zhilina@ut.ee

Haridus

Alates 2007	doktorant, molekulaardiagnostika õppekava, loodus- ja tehnoloogia-teaduskond, Tartu ülikool, Eesti
2005–2007	MSc, molekulaardiagnostika õppekava, bioloogia-geograafia-teaduskond, Tartu ülikool, Eesti
2001–2005	BSc, molekulaardiagnostika õppekava, bioloogiageograafia-teaduskond, Tartu ülikool, Eesti

Teenistuskäik

Alates 2011	teadur, Molekulaar- ja rakubioloogia instituut, Tartu ülikool, Eesti
Alates 2010	laborispetsialist, Geneetikakeskus, Ühendlabor, Tartu ülikooli kliinikum, Tartu, Eesti
2010–2011	erakorraline teadur, Molekulaar- ja rakubioloogia instituut, Tartu ülikool, Eesti
2007–2009	erakorraline teadur, Eesti Biokeskus, Eesti

Täiedkoolitused

2012	Genomic Disorders 2012: The Genomics of Rare Diseases; Hinxton, Suurbritannia
2011	6th Goldrain Course in Clinical Cytogenetics; Goldrain, Itaalia
2008	1st AnEUploidy Workshop; Genf, Šveits
2008	2nd International Summer School in Functional Genomics; Kopenhaagen, Taani
2008	4th International Meeting on Cryptic Chromosomal Re-arrangements in Mental Retardation and Autism; Sitsiilia, Itaalia
2006	EU FP6 Marie Curie Conferences and Training Courses on arrayCGH and Molecular Cytogenetics, „2nd Workshop on arrayCGH and Molecular Cytogenetics“; Gent, Belgia
2005	European School of Genetic Medicine „7th Course in Molecular Cytogenetics and DNA Arrays“; Bertinoro, Itaalia

Tunnustused ja stipendiumid

2011	Euroopa Inimesegeneetika Ühing; stipendium täiendkoolitusel (the 6th Goldrain Course in Clinical Cytogenetics) osalemiseks
2008	Euroopa Inimesegeneetika Ühing; rahvuslik stipendium noorteadlasele
2007	Eesti Teaduste Akadeemia; magistritöö autasustatud II preemiaga üliõpilastööde konkursil
2005	7th Course in Molecular Cytogenetics and DNA Arrays teadukomitee; stendiattekanne on autasustatud I preemiaga

Muu teaduslik organisatsiooniline ja erialane tegevus

Alates 2013	Euroopa Inimesegeneetika Ühingu liige
Alates 2007	Eesti Inimesegeneetika Ühingu liige

Avaldatud teadusartiklid

Žilina O., Kahre T., Talvik I., Õiglane-Šlik E., Tillmann V., Õunap K. (2014). Mosaicism for maternal uniparental disomy 15 in a boy with some clinical features of Prader-Willi syndrome. Eur J Med Genet. Published online April 2.

Žilina O., Teek R., Tammur P., Kuuse K., Yakoreva M., Vaidla E., Mölter-Väär T., Reimand T., Kurg A., Õunap K. (2014). Chromosomal microarray analysis as a first-tier clinical diagnostic test: Estonian experience. Mol Genet Genomic Med. 2(2): 166–175.

Oiglane-Shlik E., Puusepp S., Talvik I., Vaher U., Rein R., Tammur P., Reimand T., Teek R., Zilina O., Tomberg T., Ounap K. (2014) Monosomy 1p36 – A multifaceted and still enigmatic syndrome: Four clinically diverse cases with shared white matter abnormalities. Eur J Paediatr Neurol. 18(3):338–346

Teek R., Kruustük K., Žordania R., Joost K., Kahre T., Tõnisson N., Nelis M., Zilina O., Tranebjaerg L., Reimand T., Ounap K. (2013). Hearing impairment in Estonia: an algorithm to investigate genetic causes in pediatric patients. Adv Med Sci. 58(2):419–428.

Žilina O., Reimand T., Tammur P., Tillmann V., Kurg A., Õunap K. (2013). Patient with dup(5)(q35.2-q35.3) reciprocal to the common Sotos syndrome deletion and review of the literature. Eur J Med Genet. 56(4):202–206.

Zilina O., Reimand T., Zjablovskaja P., Männik K., Männamaa M., Traat A., Puusepp-Benazzouz H., Kurg A., Ounap K. (2012). Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. Am J Med Genet A. 158A(1):254–256.

Mikelsaar R, Nelis M, Kurg A, Zilina O, Korrovits P, Rätsep R, Väli M. (2012) Balanced reciprocal translocation t(5;13)(q33;q12) and 9q31.1 microduplication in a man suffering from infertility and pollinosis. Appl Genet. 53(1):93–97

Joost K., Tammur P., Teek R., Zilina O., Peters M., Kreile M., Lace B., Zordania R., Talvik I., Ounap K. (2011) Whole Xp Deletion in a Girl with Mental Retardation, Epilepsy, and Biochemical Features of OTC Deficiency. Mol Syndromol. 1(6):311–315

Sismani C., Anastasiadou V., Kousoulidou L., Parkel S., Koumbaris G., Zilina O., Bashiardes S., Spanou E., Kurg A., Patsalis PC. (2011) 9 Mb familial duplication in chromosome band Xp22.2–22.13 associated with mental retardation, hypotonia and developmental delay, scoliosis, cardiovascular problems and mild dysmorphic facial features. Eur J Med Genet. 54(5):e510–515

Männik K., Parkel S., Palta P., Zilina O., Puusepp H., Esko T., Mägi R., Nõukas M., Veidenberg A., Nelis M., Metspalu A., Remm M., Ounap K., Kurg A. (2011). A parallel SNP array study of genomic aberrations associated with mental retardation in patients and general population in Estonia. Eur J Med Genet. 54(2):136–143.

Puusepp H., Zilina O., Teek R., Männik K., Parkel S., Kruustük K., Kuuse K., Kurg A., Ounap K. (2009) 5.9 Mb microdeletion in chromosome band 17q22-q23.2 associated with tracheo-esophageal fistula and conductive hearing loss. Eur J Med Genet. 52(1):71–74

Kousoulidou L., Männik K., Zilina O., Parkel S., Palta P., Remm M., Kurg A., Patsalis PC. (2008) Application of two different microarray-based copy-number detection methodologies – array-comparative genomic hybridization and array-multiplex amplifiable probe hybridization – with identical amplifiable target sequences. Clin Chem Lab Med. 46(5):722–724

Kousoulidou L., Männik K., Sismani C., Zilina O., Parkel S., Puusepp H., Tõnisson N., Palta P., Remm M., Kurg A., Patsalis PC. (2008) Array-MAPH: a methodology for the detection of locus copy-number changes in complex genomes. Nat Protoc. 3(5):849–865

Kousoulidou L., Parkel S., Zilina O., Palta P., Puusepp H., Remm M., Turner G., Boyle J., van Bokhoven H., de Brouwer A., Van Esch H., Froyen G., Ropers HH., Chelly J., Moraine C., Gecz J., Kurg A., Patsalis PC. (2007) Screening of 20 patients with X-linked mental retardation using chromosome X-specific array-MAPH. Eur J Med Genet. 50(6):399–410

Patsalis PC., Kousoulidou L., Männik K., Sismani C., Zilina O., Parkel S., Puusepp H., Tõnisson N., Palta P., Remm M., Kurg A. (2007) Detection of small genomic imbalances using microarray-based multiplex amplifiable probe hybridization. Eur J Hum Genet. 15(2):162–172

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand.** The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro.* Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous grassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.
41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.

42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.
61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.

63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p
82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.

84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.
102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.
103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.
105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.

106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.
107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.
108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.
109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.
110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.
111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.
112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.
113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.
114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.
115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.
116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.
117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.
118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.
119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.
120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.
121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.
122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.
123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.
125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.

126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.
127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.
128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.
129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.
132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007. 112 p.
133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.
135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.
137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.
138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.
139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.
140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.
142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.
143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.

147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2009, 112 p.
158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.
161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.
162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.

166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.
168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan.** Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe.** Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi.** Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson.** Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts.** Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis.** Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov.** Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Köster.** Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap.** Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar.** Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.
181. **Kristjan Herkül.** Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.
182. **Arto Pulk.** Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü.** Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla.** Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey.** The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.
186. **Katrin Kepp.** Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.

187. **Virve Sõber.** The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro.** The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold.** Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert.** Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu.** Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik.** ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida.* Tartu, 2010, 118 p.
193. **Siim Sõber.** Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper.** Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak.** Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo.** Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel.** Multiple faces of cell-penetrating peptides – their intracellular trafficking, stability and endosomal escape during protein transduction. Tartu, 2011, 161 p.
198. **Andres Jaanus.** Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius.** Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv.** Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk.** Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.
202. **Arno Põllumäe.** Spatio-temporal patterns of native and invasive zooplankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.
203. **Egle Tammeleht.** Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.
205. **Teele Jairus.** Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.
206. **Kessy Abarenkov.** PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.

207. **Marina Grigorova.** Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.
208. **Anu Tiitsaar.** The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.
209. **Elin Sild.** Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.
210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.
211. **Pauli Saag.** Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.
212. **Aleksei Lulla.** Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage. Tartu, 2012, 143 p.
213. **Mari Järve.** Different genetic perspectives on human history in Europe and the Caucasus: the stories told by uniparental and autosomal markers. Tartu, 2012, 119 p.
214. **Ott Scheler**. The application of tmRNA as a marker molecule in bacterial diagnostics using microarray and biosensor technology. Tartu, 2012, 93 p.
215. **Anna Balikova**. Studies on the functions of tumor-associated mucin-like leukosialin (CD43) in human cancer cells. Tartu, 2012, 129 p.
216. **Triinu Kõressaar.** Improvement of PCR primer design for detection of prokaryotic species. Tartu, 2012, 83 p.
217. **Tuul Sepp.** Hematological health state indices of greenfinches: sources of individual variation and responses to immune system manipulation. Tartu, 2012, 117 p.
218. **Rya Ero.** Modifier view of the bacterial ribosome. Tartu, 2012, 146 p.
219. **Mohammad Bahram.** Biogeography of ectomycorrhizal fungi across different spatial scales. Tartu, 2012, 165 p.
220. **Annely Lorents.** Overcoming the plasma membrane barrier: uptake of amphipathic cell-penetrating peptides induces influx of calcium ions and downstream responses. Tartu, 2012, 113 p.
221. **Katrin Männik.** Exploring the genomics of cognitive impairment: whole-genome SNP genotyping experience in Estonian patients and general population. Tartu, 2012, 171 p.
222. **Marko Prous.** Taxonomy and phylogeny of the sawfly genus *Empria* (Hymenoptera, Tenthredinidae). Tartu, 2012, 192 p.
223. **Triinu Visnapuu.** Levansucrases encoded in the genome of *Pseudomonas syringae* pv. tomato DC3000: heterologous expression, biochemical characterization, mutational analysis and spectrum of polymerization products. Tartu, 2012, 160 p.
224. **Nele Tamberg.** Studies on Semliki Forest virus replication and pathogenesis. Tartu, 2012, 109 p.

225. **Tõnu Esko.** Novel applications of SNP array data in the analysis of the genetic structure of Europeans and in genetic association studies. Tartu, 2012, 149 p.
226. **Timo Arula.** Ecology of early life-history stages of herring *Clupea harengus membras* in the northeastern Baltic Sea. Tartu, 2012, 143 p.
227. **Inga Hiiesalu.** Belowground plant diversity and coexistence patterns in grassland ecosystems. Tartu, 2012, 130 p.
228. **Kadri Koorem.** The influence of abiotic and biotic factors on small-scale plant community patterns and regeneration in boreonemoral forest. Tartu, 2012, 114 p.
229. **Liis Andresen.** Regulation of virulence in plant-pathogenic pectobacteria. Tartu, 2012, 122 p.
230. **Kaupo Kohv.** The direct and indirect effects of management on boreal forest structure and field layer vegetation. Tartu, 2012, 124 p.
231. **Mart Jüssi.** Living on an edge: landlocked seals in changing climate. Tartu, 2012, 114 p.
232. **Riina Klais.** Phytoplankton trends in the Baltic Sea. Tartu, 2012, 136 p.
233. **Rauno Veeroja.** Effects of winter weather, population density and timing of reproduction on life-history traits and population dynamics of moose (*Alces alces*) in Estonia. Tartu, 2012, 92 p.
234. **Marju Keis.** Brown bear (*Ursus arctos*) phylogeography in northern Eurasia. Tartu, 2013, 142 p.
235. **Sergei Põlme.** Biogeography and ecology of *alnus*- associated ectomycorrhizal fungi – from regional to global scale. Tartu, 2013, 90 p.
236. **Liis Uusküla.** Placental gene expression in normal and complicated pregnancy. Tartu, 2013, 173 p.
237. **Marko Lõoke.** Studies on DNA replication initiation in *Saccharomyces cerevisiae.* Tartu, 2013, 112 p.
238. **Anne Aan.** Light- and nitrogen-use and biomass allocation along productivity gradients in multilayer plant communities. Tartu, 2013, 127 p.
239. **Heidi Tamm.** Comprehending phylogenetic diversity – case studies in three groups of ascomycetes. Tartu, 2013, 136 p.
240. **Liina Kangur.** High-Pressure Spectroscopy Study of Chromophore-Binding Hydrogen Bonds in Light-Harvesting Complexes of Photosynthetic Bacteria. Tartu, 2013, 150 p.
241. **Margus Leppik.** Substrate specificity of the multisite specific pseudouridine synthase RluD. Tartu, 2013, 111 p.
242. **Lauris Kaplinski.** The application of oligonucleotide hybridization model for PCR and microarray optimization. Tartu, 2013, 103 p.
243. **Merli Pärnoja.** Patterns of macrophyte distribution and productivity in coastal ecosystems: effect of abiotic and biotic forcing. Tartu, 2013, 155 p.
244. **Tõnu Margus.** Distribution and phylogeny of the bacterial translational GTPases and the Mqsr/YgiT regulatory system. Tartu, 2013, 126 p.
245. **Pille Mänd**. Light use capacity and carbon and nitrogen budget of plants: remote assessment and physiological determinants. Tartu, 2013, 128 p.

246. **Mario Plaas.** Animal model of Wolfram Syndrome in mice: behavioural, biochemical and psychopharmacological characterization. Tartu, 2013, 144 p.
247. **Georgi Hudjašov.** Maps of mitochondrial DNA, Y-chromosome and tyrosinase variation in Eurasian and Oceanian populations. Tartu, 2013, 115 p.
248. **Mari Lepik**. Plasticity to light in herbaceous plants and its importance for community structure and diversity. Tartu, 2013, 102 p.
249. **Ede Leppik**. Diversity of lichens in semi-natural habitats of Estonia. Tartu, 2013, 151 p.
250. **Ülle Saks.** Arbuscular mycorrhizal fungal diversity patterns in boreo-nemoral forest ecosystems. Tartu, 2013, 151 p.
251. **Eneli Oitmaa**. Development of arrayed primer extension microarray assays for molecular diagnostic applications. Tartu, 2013, 147 p.
252. **Jekaterina Jutkina.** The horizontal gene pool for aromatics degradation: bacterial catabolic plasmids of the Baltic Sea aquatic system. Tartu, 2013, 121 p.
253. **Helen Vellau.** Reaction norms for size and age at maturity in insects: rules and exceptions. Tartu, 2014, 132 p.
254. **Randel Kreitsberg.** Using biomarkers in assessment of environmental contamination in fish – new perspectives. Tartu, 2014, 107 p.
255. **Krista Takkis.** Changes in plant species richness and population performance in response to habitat loss and fragmentation.Tartu, 2014, 141 p.
256. **Liina Nagirnaja.** Global and fine-scale genetic determinants of recurrent pregnancy loss. Tartu, 2014, 211 p.
257. **Triin Triisberg.** Factors influencing the re-vegetation of abandoned extracted peatlands in Estonia. Tartu, 2014, 133 p.
258. **Villu Soon.** A phylogenetic revision of the *Chrysis ignita* species group (Hymenoptera: Chrysididae) with emphasis on the northern European fauna. Tartu, 2014, 211 p.
259. **Andrei Nikonov.** RNA-Dependent RNA Polymerase Activity as a Basis for the Detection of Positive-Strand RNA Viruses by Vertebrate Host Cells. Tartu, 2014, 207 p.
260. **Eele Õunapuu-Pikas**. Spatio-temporal variability of leaf hydraulic conductance in woody plants: ecophysiological consequences. Tartu, 2014, 135 p.
261. **Marju Männiste**. Physiological ecology of greenfinches: information content of feathers in relation to immune function and behavior. Tartu, 2014, 121 p.
262. **Katre Kets.** Effects of elevated concentrations of CO_2 and O_3 on leaf photosynthetic parameters in *Populus tremuloides*: diurnal, seasonal and interannual patterns. Tartu, 2014, 115 p.
263. **Külli Lokko.** Seasonal and spatial variability of zoopsammon communities in relation to environmental parameters. Tartu, 2014, 129 p.