The human genome; you gain some, you lose some

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit van Leiden,
op gezag van de Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het Collega van Promoties
te verdedigen op donderdag 6 december 2007
klokke 15.00 uur

door

Marjolein Kriek
geboren te Leiden, in 1973

Promotiecommissie

Promotoren:	Prof. dr. M.H. Breuning Prof. dr. G-J. B. van Ommen
Co-promotor:	Dr. J.T. den Dunnen
Referent:	Prof. dr. H.H. Ropers (Max Planck Instituut te Berlijn)
Overige leden:	Dr. K. Szuhai

Voor Bram

ISBN 978-90-9022286-8

Designed by: Grafisch Bureau Christine van der Ven, Voorschoten

Cover design: Ik heb tijdens mijn promotieonderzoek gezocht naar veranderingen in het erfelijk materiaal, die het voorkomen van een verstandelijke beperking bij de mens zouden kunnen verklaren. De voorkant van dit proefschrift laat de vormgeving van een mens zien, vertaald door Petra Kaak, kunstenares bij Kunst en Vliegwerk.
Kunst & Vliegwerk verzorgt een bijzondere vorm van dagbesteding voor kunstzinnig getalenteerde mensen met een verstandelijke handicap.

Printed by: Grafische Producties, Universitair Facilitair Bedrijf, Leiden

The author of this thesis was financially supported by the Netherlands Organisation for Health Research and Development (ZON-Mw), registration number 940-37-032.

© 2007 M. Kriek, Leiden, The Netherlands
All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means, elecrtonic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the copyright owner.

Content

List of definitions

Acrocentric chromosomes:	Chromosomes lacking the short arm. The human acrocentric chromosomes are 13, 14, 15, 21, and 22.
Congenital malformation:	A physical defect present in the newborn.
Copy number:	The number of copies of a given chromosomal locus.
Copy number variation:	Alteration of a copy number of a certain DNA sequence in relation to the normal situation.
with phenotypic trait:	variation with clinical consequences.
without phenotypic trait:	variation without obvious clinical consequences (also called Polymorphic CNVs).
Deletion:	Loss of a DNA sequence.
Duplication:	An extra copy of a DNA sequence.
Duplicon:	Duplicon or segmental duplication has been defined as sequences of DNA greater than 1 Kb in size sharing a homology of at least 90 %.
False positive result:	An incorrect positive result of a test.
False negative result:	A result that appears negative but fails to reveal an alteration.
Gene:	Coding sequence.
Gene desert:	Region in the human genome that does not contain genes.
Genomic disorders:	The clinical condition that results from a dosage alteration of gene(s) located within a rearranged segment of the genome.
Mendelian inheritance:	Several inheritable traits or congenital conditions in humans are classical examples of Mendelian inheritance: Their presence is controlled by a single gene that can either be of the autosomal-dominant or -recessive type. People that inherited at least one dominant gene from either parent usually present with the dominant form of the trait. Only those that received the recessive gene from both parents present with the recessive phenotype (Wikipedia).

Mental retardation (MR) classification:	Mild MR (intelligent quotient (IQ) between 50 and 70), moderate MR (IQ between 35 and 50), severe MR (IQ between 20 and 35) and profound MR (IQ below 20).
Polymorphic CNVs:	CNVs (deletions as well as duplications) that are not related to a clinical phenotype (also called CNVs without phenotypic trait).
Phenotypic trait:	Any (abnormal) clinical feature, such as mental retardation, congenital malformations, dysmorphologies.
Translocations:	Exchange of genetic material between two different chromosomes.
Robertsonian translocations:	These translocations are produced by exchange in proximal short arms of the acrocentric chromosomes. Both centromeres are present, however, they function as one unit. This translocation is named after W.R.B. Robertson who described fusion of acrocentric chromosomes in insects.
Reciprocal translocations:	A translocation where part of one chromosome is exchanged with a part of a separate non-homologous chromosome.
Transposition:	Transfer of a segment of DNA to a new position on the same or another chromosome.
Uniparental disomy:	A euploid cell in which one of the chromosome pairs have been inherited exclusively from one parent. If two identical homologues are inherited this called isodisomy; if non-identical homologues are inherited the term heterodisomy is used. This occurs when non-disjunction during meiosis in one parent leads to formation of a disomic gamete. A trisomic zygote is formed and trisomic rescue with loss of the chromosome from the other parent occurs. UPD is of particular relevance in imprinted regions of the genome.

List of abbreviations

Bp	Base pair
BAC	Bacterial Artificial Chromosome
CGH	Comparitive Genome Hybridisation
CM	Congenital Malformation
CNVs	Copy Number Variation
COBRA	COmbined Binary RAtio
DD	Development Delay
DNA	Deoxyribonucleic acid
DOP-PCR	Degenerate Oligonucleotide Primed Polymerase Chain Reaction
FISH	Fluorescent in Situ Hybridisation
I.Q.	Intelligence Quotient
K	Kilo
Kb	Kilo base (one thousand base pairs)
LCR	Low Copy Repeat
MAPH	Multiplex Amplifiable Probe Hybridisation
Mb	Mega base (one million base pairs)
M-FISH	Multi-colour FISH
MLPA	Multiplex Ligation-dependent Probe Amplification
MR	Mental Retardation
NAHR	Non Allelic Homologous Recombination
Nt	Nucleotide
PAC	P1 derived Artificial Chromosome (PAC)
PCR	Polymerase Chain Reaction
PFGE	PulseField Gel Electrophoresis
RFLP	Restriction Fragment Length Polymorphism
SKY	Spectral Karyotyping
SNP	Single Nucleotide Polymorphism
UPD	Uniparental Disomy
VNTR	Variable Number of Tandem Repeats

Chapter I

Introduction

I-1. The plasticity of the human genome

Many authors have discussed the significance of gene and whole genome duplication in evolution (these publications are reviewed in (Taylor and Raes 2004)). Indeed, Ohno (1970) (in Evolution by gene duplication. New York: Springler-Verlag) stated that duplications of the genetic material were the most important factor driving evolution. Recently, projects using genome sequencing have shown that large scale gene duplications have contributed to the creation and expansion of gene families. Whether a duplication is passed onto future generations depends on whether the change is beneficial for survival. One example is the olfactory gene family. These (pseudo)genes create a redundancy of sequences contributing to the ability to smell, which appears to be beneficial for mammalian survival. A more recent example was published by Perry et al. (2007). They found that the copy number of the AMY1 gene is positively correlated with the amount of starch in a diet. We have also learned that the susceptibility of developing a disease is influenced by changes in CNVs. It has been shown that altered copy number of the CCL3L1 and FCGR3B genes influence susceptibility to HIV infection and systemic lupus erythematosus (SLE), respectively (Gonzalez et al. 2005; Aitman et al. 2006). These examples indicate that selection may operate on copy number variants containing sequences that are coding or regulating functions involved in survival.

A substantial proportion of (partial) gene duplications are gathered in segmental duplications (**chapter II-1**). Segmental duplications presumably originated from the duplication and subsequent transposition (and / or inversion) of genomic blocks (Eichler 2001a) from one chromosomal region to another some tens of million years ago (Bailey *et al.* 2002b; Armengol *et al.* 2003). It appears that these segmental duplications are often present at (breakpoint) loci where the human genome differs from that of the great apes (Samonte and Eichler 2002a) (Stankiewicz *et al.* 2001; Locke *et al.* 2003) and other species, such as mice (Armengol *et al.* 2003).

Besides duplications of existing sequences, another frequent form of variation in the human genome is deletion of unique sequences. In fact, it has been shown that these deletions are quite common in the human genome, with each individual having at least 30-50 deletions larger than 5 kb (Conrad *et al.* 2006). Van Ommen (2005) estimated that one in eight live births may have a *de novo* deletion. Some of these may enhance adaptation to environmental changes and might therefore be beneficial for survival. It is assumed that these deletion polymorphisms are exposed to more strict selection than Single Nucleotide Polymorphisms (SNPs), based on the fact that the X-chromosome contains less deletion polymorphisms compared to SNPs (Conrad *et al.* 2006).

In contrast to their potentially positive role in evolution, duplications and deletions (e.g. copy number variations = CNVs) (figure 1 A&B) in the human genome can also be related to inherited disease, mental retardation (MR), and congenital malformations (CM). For decades, it has been clear that numerical chromosome aberrations (e.g. trisomy 13, 18 and 21) and large CNVs have enormous influence on embryonic development and can lead to malformation syndromes or intra-uterine death. More recently, a systematic search for submicroscopic CNVs leading to MR and CM was initiated by Flint *et al.* (1995). These authors focused on the chromosome ends (also called the subtelomeres) and they found the percentage of alterations in their MR study population to be around 6%. Since that time, many different screening tools have been successfully implemented to find such cryptic (subtelomeric) CNVs (table 1). Detecting small CNVs on a genome-wide scale has only recently become possible with the development of micro-arrays. First results indicate that many CNVs are detected in patients with MR and CM (CNVs with phenotypic trait) as well as in healthy individuals (CNVs without an obvious phenotypic trait). In the most comprehensive CNV study to date no less than 12% of the human genome showed variations among healthy individuals (Redon *et al.* 2006). Consequently, our main challenge is currently to determine whether a variation is related to a phenotypic trait or not. This will remain so in the near future until the complete plasticity of the human genome has been fully mapped.

In short, copy number variations (CNVs) in the human genome are inherent in both evolutionary progression as well as the etiology of disease. The introduction of this thesis will review CNVs that appear to be neutral as well as CNVs that appear to be related to a phenotypic trait. This will be followed by a review of the many different technical approaches that can be used for detecting genomic rearrangements.

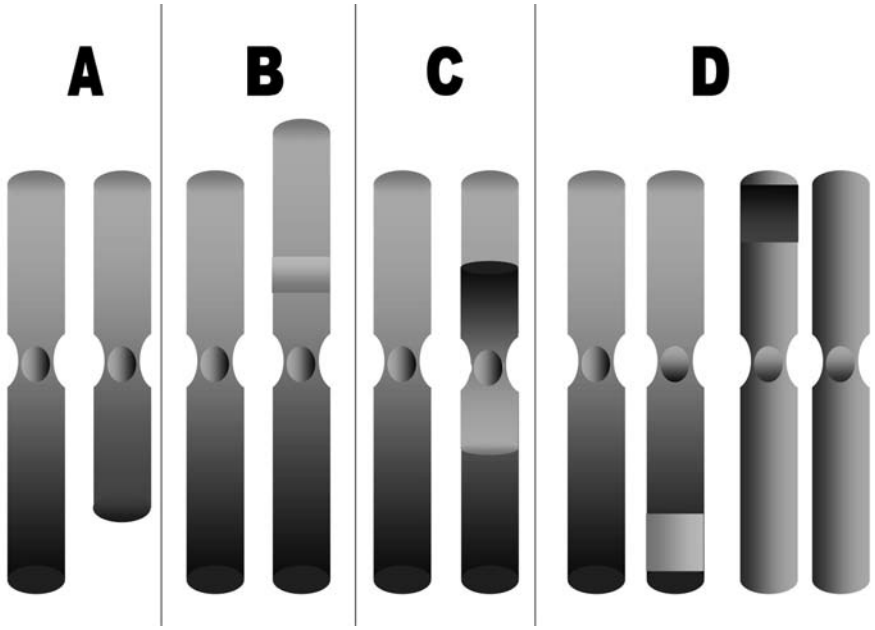
The articles **(chapter II & III)** describe several studies that have applied the rapidly evolving techniques for CNV detection to the clinical problem of unexplained MR and CM. The availability of the new diagnostic tools will greatly increase our understanding of the genetic causes of MR and CM, and might one day lead to therapeutic interventions in some cases.

I-2. CNVs with no obvious phenotypic trait

2.1. Neutral CNVs

Copy number variants have been identified since the start of the cloning era, however, the full extent of the variability and plasticity of the human genome has only recently

Figure 1. Deletion, duplication, inversion and balanced translocation.

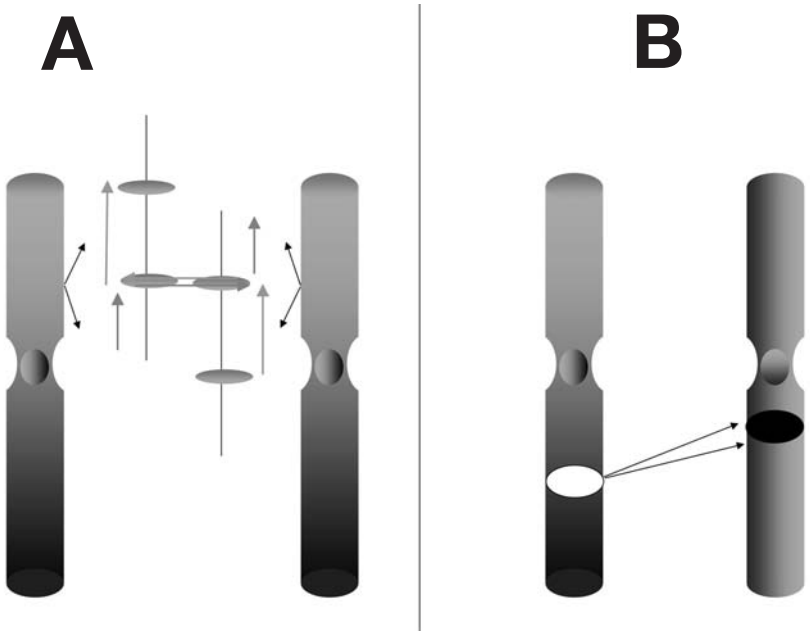


A. Part of the long arm of the right chromosome is missing. The loss of genomic material is called a deletion.
B. A part of the short arm of the chromosome is present twice (right). This extra material is called a duplication. As the duplicated region is localised within the chromosome, this duplication is called an interstitial duplication.
C. The amount of genetic material in part C of this picture is similar to the unaffected left chromosome. However, a part of the chromosome is inverted. As the centromere is localised within the invertion, this situation is called a pericentromeric inversion.
D. Again the amount of genetic material is normal, however, a part of the information of the dark grey chromosome has been transported to the light grey chromosome and vice versa. This is called a balanced translocation.

[See appendix: colour figures.]

been appreciated (Iafrate *et al*. 2004; Sebat *et al*. 2004; Fredman *et al*. 2004). Sebat *et al*. (2004) presented the first study assessing the frequency of CNVs in the healthy population using genome-wide screening tools. CNVs were shown to be frequent and, although they are present all over the human genome, loci enriched for structural rearrangements are not randomly distributed. Regions within or flanked by segmental duplications show a higher frequency of CNVs compared to regions outside these duplications. Furthermore, the genes that show enrichment in CNVs are also not random. Genes associated with immunity-, defence, cancer susceptibility, drug detoxification, signal transduction and sex hormone metabolism frequently show variations (Eichler 2006), including null-alleles. McCarroll *et al*. (2006) showed these variations to result in expression level differences, indicating that these variants are related to adaptation. On the other hand, the

Figure 2. Non-allelic homologous recombination and insertions.



A. Non allelic homologous recombination. The two alleles of a chromosome contain regions that are highly homologous (e.g. segmental duplications, low copy repeats or duplicons). The presence of these segmental duplications can result in misalignment of these regions and subsequently in non allelic homologous recombination. The green arrow shows the origin of a duplication of the region present between two highly homologous regions, whereas the red arrow indicates the origin of a deletion.
B. In this situation a part of the left chromosome is inserted in another chromosome. This is called an insertion.
[See appendix: colour figures.]

majority of deletions found thus far were located in so called gene-deserts (Conrad *et al.* 2006) and may therefore be neutral variants or have modest regulatory effects due to the presence of microRNA, noncoding RNA and other highly conserved regions.

Nearly half of all CNVs seem to be complex events, formed by more than one event (for example an inversion (figure 1C) and a deletion, or a deletion combined with a duplication) (Eichler unpublished data).

2.2 Segmental duplications

2.2.1. Characteristics of segmental duplications

Segmental duplications have been defined as sequences of DNA greater than 1 Kb in size sharing a homology of at least 90 % (She *et al.* 2006). Previous studies

indicate that at least 5% (154 Mb) of the human genome is composed of such duplications (Bailey *et al*. 2002a; Cheung *et al*. 2003b; She *et al*. 2004; Zhang *et al*. 2005), also called Low Copy Repeats (LCRs) or duplicons. Duplicons can have either a simple or a complex structure (Ji *et al*. 2000) and contain genes, pseudogenes, gene fragments, repeat gene clusters (Ford and Fried 1986) and other chromosomal segments (Eichler *et al*. 1996; Samonte and Eichler 2002b; Horvath, Schwartz, and Eichler 2000). Especially the pericentromeric regions consist of a mosaic of different genomic segments (Horvath, Schwartz, and Eichler 2000). Compared to the chimpanzee and baboon, the human genome is particularly enriched for the number and the length of mainly *Alu* repeats (Liu *et al*. 2003). Also, the degree of genome sequence identity is higher in humans compared to other vertebrates (She *et al*. 2006).

Misalignment between segmental duplications followed by Non Allelic Homologous Recombination can result in a duplication and reciprocal deletion of the sequence flanked by these duplicons (figure 2A). However, the high degree of sequence homology between segmental duplications alone is not sufficient for providing 'repetitive breakpoints events', and therefore additional conditions are needed before recombination occurs. These include minimum length of 100% homology required for recombination in human mitosis and meiosis (minimal region of homology was estimated to be 220 – 300bp and 300 – 500 bp, respectively) (Lupski *et al*. 1992; Waldman and Liskay 1988), AT-rich sequences (Peoples *et al*. 2000), for example those present on both sites of a recombination hotspot in Smith Magenis Syndrome (Bi *et al*. 2003) and enrichment of *Alu* repeats near or within the junctions present in segmental duplications (Stoppa-Lyonnet *et al*. 1990; Potocki *et al*. 2000; Bailey, Liu, and Eichler 2003).

Segmental duplications are also largely responsible for the fact that a part of the human genome sequence working draft contains gaps or is misassembled. The higher the sequence similarity the more difficult it is to distinguish and correctly assemble LCRs (Eichler 2001b).

2.2.2. Intra- and interchromosomal segmental duplications

Segmental duplications can be divided in two categories, interchromosomal and intrachromosomal. Interchromosomal segmental duplications are based on the transposition of DNA sequences towards other chromosomes, whereas intrachromosomal segmental duplications originated from a sequence that is transported to another region within the same chromosome. The prevalence of intrachromosomal segmental duplications in humans is higher than interchromosomal segmental duplications (3.97%,

113.66 Mb versus 2.37 %, 67.86 Mb)(Samonte and Eichler 2002b; Cheung *et al*. 2003a; She *et al*. 2006).

Interchromosomal segmental duplications are frequently found at pericentromeric and subtelomeric sites (Cheung *et al*. 2001). An example is the pericentromeric region of the short arm of chromosome 16, which contains four different segmental duplications that were duplicated and subsequently transposed from Xq28, 15q13, 2p11 and 14q32 (Ji *et al*. 2000) towards 16p11.

While studying the olfactory gene family, which is spread over several chromosomes, (Trask *et al*. 1998) found that there are differences in subtelomeric segmental duplications between different ethnic groups, suggesting that such rearrangements are still ongoing.

I-3. CNVs with phenotypic trait: genomic disorders

3.1. Genomic disorders

Genomic disorders were defined in 1998 (Lupski 1998) as the clinical condition, all types of phenotypic features included, that result from the dosage alteration of gene(s) located within a rearranged segment of the genome. It was estimated that about 0.7-1 / 1000 live births suffer from a genomic disorder (Ji *et al*. 2000). Different types of CNV are involved in genomic disorders, e.g whole, and partial chromosome alterations (see **section 4**). These alterations include deletions, duplications, inversions, insertions and translocations (see figure 1 and figure 2). Three clinical conditions frequently arising from such CNVs are discussed below.

3.2. Mental retardation (MR)

MR or developmental delay (DD) is defined as a significant impairment of cognitive and adaptive functions (Battaglia and Carey 2003). It is a clinically important condition as it affects about 1:30 – 1:50 people. MR can be categorised into four degrees of severity (WHO 1980, International classification of Impairments, disabilities and handicaps. Geneve: World Health Organisation, 1980): Mild MR (intelligent quotient (IQ) between 50 and 70), moderate MR (IQ between 35 and 50), severe MR (IQ between 20 and 35) and profound MR (IQ below 20).

Both genetic - and environmental factors can contribute to the origin of mental retardation. Environmental factors can involve pre- peri- and postnatal events, such as oxygen deprivation (perinatal event), infection (prenatal, postnatal), teratogenic

influences (prenatal) (Hamel 1999. X-linked MR. A clinical and molecular study (Alkmaar: Dekave)).

Genetic causes for mental retardation include (1) chromosomal causes such as aneuploidies, chromosome end rearrangements, rearrangements in regions related to microdeletion syndromes and other interstitial rearrangements, (2) complex disorders (caused by mutations in multiple genes) and (3) monogenic disorders (**section 4.2**.). A substantial number of point mutations have been identified in isolated genes that play an important role in early development (Petrij *et al*. 1995), such as mutations in the *RAI1* (Slager *et al*. 2003) causing Smith Magenis syndrome, mutations in the *CREBBP* gene (responsible for Rubinstein Taybi syndrome) and the CTG expansion of the *FMR*-gene which accounts for about 1:4000 – 1:6000 male cases of mental retardation (Fragile X syndrome) (Murray *et al*. 1996; Turner *et al*. 1996; De Vries *et al*. 1997) (**section 4.2.**).

It is known that the causes of mental retardation vary with the severity of the condition. Large CNVs are more frequently associated with severe cases. Chromosomal and genetic disorders account for 30%- 50% of moderate to severe mental retardation (I.Q.< 50); environmental insults explain a further 10%-30% (Gustavson, Holmgren, and Blomquist 1987; McDonald 1973; Elwood and Darragh 1981; Flint and Wilkie 1996). In mild mental retardation cases (I.Q. between 50 and 70), approximately equal proportions of genetic and environmental causes are diagnosed, about 10-30% each (Lamont and Dennis 1988; Bundey, Thake, and Todd 1989; Einfeld 1984).

The cause of MR remains unclear in about 40-50% of cases, indicating that, despite its high prevalence, the pathogenesis of MR is poorly understood. It is expected, however, that this rather high percentage will decline with the use of recently developed high-resolution genome analysis (see **section 6.2. and 6.3.**).

3.3. Congenital Malformation (CM)

Along with mental retardation, CNVs in the human genome may also result in a wide range of congenital malformations, such as organ and skeletal defects. These clinical features are already present at birth, before the mental retardation becomes apparent, so these entities can be the first indication of a genetic defect. The presence of more than one CM in a newborn that lacks a characteristic pattern of a specific microdeletion syndrome is an indication for genome-wide screening for CNV.

I-4. CNVs with phenotypic trait: Different types of variations

4.1. Whole chromosome variations

Since it was shown that an extra chromosome 21 causes Down syndrome (LEJEUNE, TURPIN, and GAUTIER 1959; Jacobs *et al.* 1959), it became clear that aneuploidy has significant influence on early development as well as on the intellectual capacities of an individual. Moreover, the severity of congenital malformations associated with trisomy 13 or 18 is such that only a small percentage of these fetuses will be viable with a drastically reduced life expectancy. Complete aneusomies of the remaining autosomal chromosomes have not been reported among live births, indicating that these are not compatible with life. Studies on material from spontaneous abortions support this statement (Carr 1971; Lauritsen *et al.* 1972; Boue and Boue 1977).

The fact that cells use one copy of the X chromosome while inactivating extra copies, combined with the small number of genes on the Y chromosome results in the less severe impact of sex chromosomes aneuploidies on the development of the embryo. Karyotypes such as 45,X, 47,XXX, 47,XXY, 47,XYY constitute the most common class of chromosome abnormality in humans (Hall, Hunt, and Hassold 2006).

Incomplete aneusomies of autosomal and sex chromosomes (chromosomal mosaicisms) are also known to be present in both affected and healthy individuals. The phenotypic consequence of a chromosomal mosaicism depends on the chromosome involved, the percentage of abnormal cells and the tissue(s) that contain cells with an abnormal chromosomal constitution.

Some of the whole chromosome variations originate from Robertsonian translocations in one of the parent of the affected fetuses / newborn. The frequency of Robertsonian translocations is 1:1000 (Shaffer and Lupski 2000).

4.1. Partial chromosome variations

4.1.1. Subtelomeric CNVs

The subtelomeric regions are localized proximal to the telomere proper, which consists of short repetitive sequences that cap the end of the chromosome. The subtelomeric regions from different chromosomes are highly variable, with some having a simple pattern and little similarity to other chromosome ends, whereas others contain complex and extensive patterns of homology. A good example regarding similarity of two subtelomeric regions is 4q and 10q, both encompassing repeats that share >98% sequence homology (van Overveld *et al.* 2000; van Geel *et al.* 2002). The subtelomeres are particularly dynamic regions, due to repeat-rich sequences that have a high frequency

Table 1. Overview of subtelomeric screening studies in chronological order. Based on Rooms *et al.* (2004a) with addition of more recent publications.

Reference	Method of analysis	Number of cases	Detection rate
Flint *et al.* (1995)	VNTR marker analysis	99	3%
Knight *et al.* (1999)	Multiprobe FISH	284 moderate/severe	7.4%
		182 mild	0.5%
Slavotinek *et al.* (1999)	Microsatellitemarker analysis	27	7.5%
Bonifacio *et al.* (2001)	PRINS	65	3.1%
Borgione *et al.* (2001)	Microsatellitemarker analysis	60	6.6%
Colleaux *et al.* (2001)	Microsatellitemarker analysis	29	6.9%
Fan *et al.* (2001)	Multiprobe FISH	150	4%
Riegel *et al.* (2001)	Multiprobe FISH	254	5%
Rosenberg *et al.* (2001)	Microsatellitemarker analysis	120	4.1%
Rossi *et al.* (2001)	Multiprobe FISH	200	6%
Sismani *et al.* (2001)	Multiprobe FISH / MAPH	70	1.4%
Anderlid *et al.* (2002)	Multiprobe FISH	111	9%
Baker *et al.* (2002)	Multiprobe FISH	53 isolated MR	1.9%
		197 MR and dysmorphic features/malformations	4.1%
Clarkson *et al.* (2002)	Multiprobe FISH/ SKY	50	6%
Dawson *et al.* (2002)	Multiprobe FISH	40	10%
Hélias-Rodzewicz *et al.* (2002)	Multiprobe FISH	33	9%
Hollox *et al.* (2002)	MAPH	37	13.5%
Popp *et al.* (2002)	M-TEL	30	13.3%
Rio *et al.* (2002)	Microsatellitemarker analysis	150	10%
Van Karnebeek *et al.* (2002)	Multiprobe FISH	184	0.5%
Hulley *et al.* (2003)	Multiprobe FISH	13	7.7%
Jalal *et al.* (2003)	Multiprobe FISH	372	6.8%
Bocian *et al.* (2004)	Multiprobe FISH	59 moderate-severe	10%
		24 mild	12.5%
Harada *et al.* (2004)	Array CGH	69	5.8%
Koolen *et al.* (2004)	MLPA	210	6.7%
Kriek *et al.* (2004)	MAPH	184	4.3%
Pickard *et al.* (2004)	MAPH / FISH	69 mild	1.5%
Rodriguez-Revenga *et al.* (2004)	Multiprobe FISH	8 moderate-severe	12.5%
		22 mild	4.5%
Rooms *et al.* (2004b)	Microsatellitemarker analysis	70	-
Rooms *et al.* (2004a)	MLPA	75	5.2%
Walter *et al.* (2004)	Multiprobe FISH	50	10%
Novelli *et al.* (2004)	Multiprobe FISH	92	16.3%
Li and Zhao (2004)	Multiprobe FISH	46	4.4%
Rooms *et al.* (2006)	MLPA	275	4.4%
Lam *et al.* (2006)	MLPA / multprobe FISH	20	15%
Palomares *et al.* (2006)	MLPA	50	10%
	Multiprobe FISH	50	10%

of recombination. They are also gene- rich, and the plasticity of these chromosomal regions may be one of the factors responsible for phenotypic diversity (Mefford and Trask 2002).

CNVs near the chromosome ends are a significant cause of idiopathic mental retardation (Flint *et al*. 1995; Knight *et al*. 1999; Flint and Knight 2003). Flint *et al*. (1995) demonstrated that ~6% of the patients with idiopathic mental retardation have a rearrangement in a subtelomeric region. These findings were verified by observations in many other studies. Biesecker (2002) and later Rooms *et al*. (2004a) summarized subtelomeric aneusomy screening studies using various detection methods (table 1). In our study, **(chapter II-1)** 4.3% subtelomeric alterations were found among 184 idiopathic mild to severe MR patients.

The percentage of aberrations detected varies considerably between different studies. This is due to the different criteria for the selection of patients, different techniques used, and, in smaller patient groups, by stochastic factors. It seems that the number of CNVs detected goes up with increasing complexity and severity of the clinical problems of the patients.

A proportion of the subtelomeric imbalances originate from reciprocal translocations in one of the parents. The frequency of reciprocal translocations is 1:625 (Shaffer and Lupski 2000). All chromosomes seem to participate in reciprocal translocations and most of the breakpoints are family-specific, however some breakpoints are recurrent, such as t(11;22)(q23-q11.2) and t(4;8)(p16;p23) (Giglio *et al*. 2002). These common and recurrent breakpoints originate from misalignment between interchromosomal duplicons, which can lead to crossing over between non homologous chromosomes (Kurahashi *et al*. 2000; Kurahashi *et al*. 2003).

Gribble *et al.* (2005) studied a group of patients with a phenotypic trait and who had initially been diagnosed to have a balanced translocation based on the outcome of karyotyping. The majority of these apparent balanced translocations appeared to consist of several complex rearrangements often combined with the presence of one or more imbalances. To gain more insight in different 'balanced' translocations and their consequences, Danish investigators started to collect and characterize large numbers of balanced chromosomal rearrangements (Bugge *et al.* 2000).

4.1.2. CNVs in microdeletion syndromes regions

Microdeletion syndromes result from the loss of several genes (contiguous gene syndrome) or may result from the loss of a single gene. The majority of the microdeletion related regions are localised between intrachromosomal segmental duplications. These

Table 2. Characteristics of syndromes flanked by duplicons (recombination hotspots) of which the reciprocal alteration has also been identified to have clinical consequences.

Localisation	CNV	Genomic	Size of duplicon (kb)	Size of CNV (Mb)	Freq.	References
17p12	Del	Hereditary Neuropathy with liability to Pressure Palsy	24	1.5	1:20000	Reiter *et al.* (1996); Reiter *et al.* (1998); Inoue *et al.* (2001)
	Dup	Charcot-Marie-Tooth syndrome			1:2500	Valentijn *et al.* (1992); Pentao *et al.* (1992); Lupski *et al.* 1992; Lupski *et al.* (1991)
22q11	Del	DiGeorge - / Velo-CardioFacial Syndrome	200	3	1: 4000	Shaikh *et al.* (2000); Edelmann, Pandita, and Morrow (1999)
	Dup	22q11 duplication syndrome			Probably equal	Yobb *et al.* (2005) Ensenauer *et al.* (2003)
7p11.2	Del	Smith Magenis syndrome	250 - 400	5.0	1:25000	Bi *et al.* (2003); Slager *et al.* (2003) Shaw, Bi, and Lupski (2002)
	Dup	Potocki-Lupski syndrome			Probably equal	Chen *et al.* (1997) Potocki *et al.* (2000); Bi *et al.* 2003; Potocki *et al.* (2007)
7q11.23	Del	Williams syndrome	320	1.6	1:20000-50000	Bayes *et al.* (2003); Peoples *et al.* (2000) Urban *et al.* (1996); Francke (1999)
	Dup	Duplication of the Williams Critical region			Probably equal	Somerville *et al.* (2005); Kriek *et al.* (2006)

As reciprocal duplications have only been discovered recently, the frequency cannot be determined based on literature. Based on Non Allelic Homologous Recombination one can assume that the frequency of reciprocal duplication is equal to that of the corresponding deletion, although there is no reason to assume that the consequence of a deletion or duplication would be the same. Nevertheless, it seems that the frequency of HNPP is an underestimation. In addition to the duplication of the region involved in DiGeorge/VCF syndrome, tetrasomy of this 22q11 region has also been described in Cat eye syndrome. Del = deletion, dup =duplication, Freq. = frequency, CNV = Copy Number Variation.
This table was based on table 3 of Shaffer and Lupski (2000).

homologous regions facilitate unequal crossing over, resulting in deletions as well as duplications (Chance *et al.* 1994). This indicates that the frequency of reciprocal duplications of such regions is in principle equal to that of the corresponding deletions. In general, clinical phenotypes of these duplications are milder compared to the deletion of the same region (for references see right column of table 2), and some of these

duplications might not even result in MR. In addition, duplications used to be more difficult to detect compared to deletions. This explains the lower frequency of publications regarding micro- duplications within such regions. Examples of microdeletion syndromes that are flanked by duplicons include Hereditary Neuropathy with liability to Pressure Palsy (HNPP), Williams-Beuren syndrome, DiGeorge- / Velocardiofacial syndrome, Smith Magenis syndrome (see table 2), Angelman - /Prader Willi syndrome (Miller, Dykes, and Polesky 1988; Amos-Landgraf *et al*. 1999) (see table 2). Up to now microdeletion syndromes have been recognised by their distinctive clinical phenotypes, using targeted fluorescence in situ hybridisation (FISH) to detect the deletion in patients selected by a dysmorphologist. Recently, the genome-wide array-CGH method revealed additional microdeletions among MR patients that at first sight appeared to lack salient and distinct features. A recent example of such a microdeletion is the 17q21.31 microdeletion syndrome that is associated with parental inversion of this region (Shaw-Smith *et al*. 2006; Koolen *et al*. 2006; Sharp *et al*. 2006). After identification of the deletion, dysmorphologists do see common features in a series of patients, possibly enabling the recognition of these patients in the clinic.

4.1.3. Other interstitial CNVs

Several CNVs localised outside the subtelomeres and microdeletion related regions have been identified as being involved in the etiology of MR/CM.

Bailey *et al*. (2002) described a bioinformatic approach to analyse the human genome sequence, and identified nearly two hundred potential hotspots for CNVs, e.g. regions flanked by segmental duplications (Bailey *et al*. 2002a). Some of these regions appear to be related to genomic disorders. 130 of these regions were subsequently tested for rearrangements among 47 healthy individuals using a segmental duplicon BAC microarray (Sharp *et al*. 2005). 79 of the 130 potential CNV hotspots showed no alteration among this study population, supporting the hypothesis that alterations within these regions could be related to disease. **Chapter II-2** summarizes our results of screening for CNVs of regions flanked by intrachromosomal duplicons among 105 MR/CM patients. As expected, the rearrangement frequency per unit of DNA is much higher in regions flanked by duplicons compared to regions without known duplicons nearby, supporting the statement that regions flanked by duplicons are enriched for copy number variations. Of course, pathogenic CNVs outside duplicon-flanked_regions have also been identified, for example the interstitial deletion of chromosome band 2p16p21 (Sanders *et al*. 2003; Lucci-Cordisco *et al*. 2005) (see **chapter III-4**) and the *DMD* gene (Blonden *et al*. 1991; Nobile *et al*. 2002).

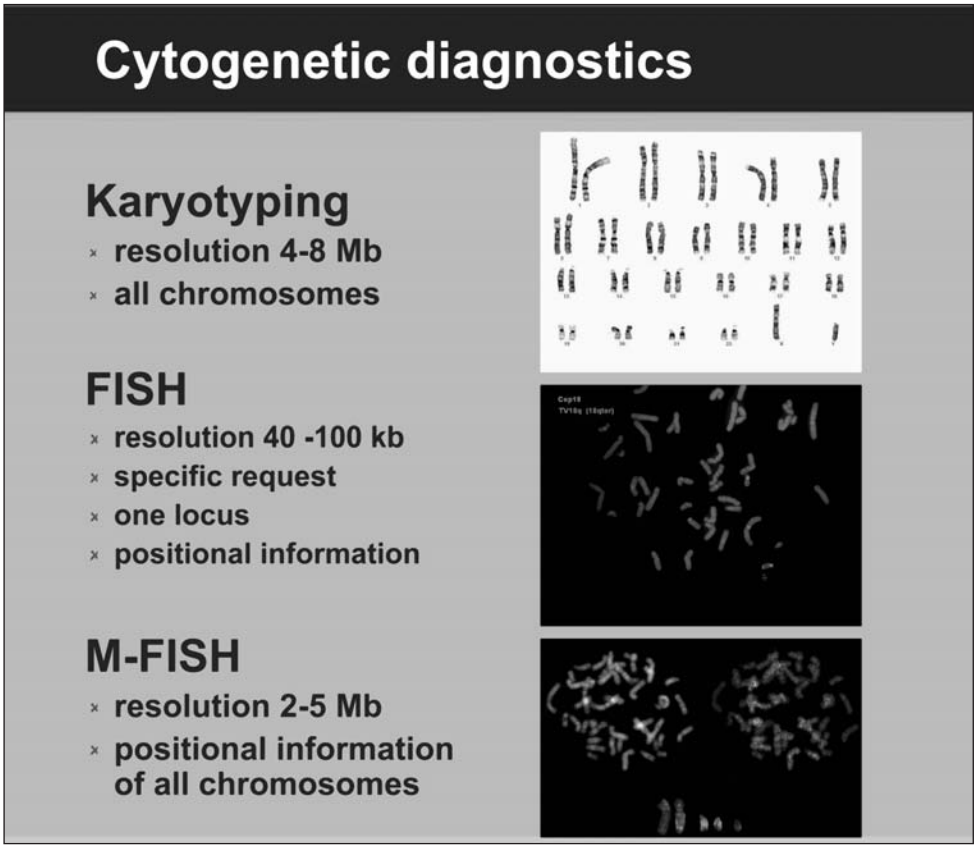
4.2. Other variations

Several microdeletion syndromes are in fact caused by the inactivation of a single gene. An example is the Rubinstein Taybi Syndrome (RTS). After two reciprocal translocations with a breakpoint in the short arm of chromosome 16 had been described in RTS patients, submicroscopic deletions were detected in six of a series of 25 patients with the syndrome (Breuning *et al*. 1993). Subsequent mutation detection using the protein truncation test identified two point mutations in the CREBBP gene in 16p (Petrij *et al*. 1995), indicating that RTS was not, as previously thought, a contiguous gene syndrome, but due to haplo-insufficiency of a single gene. Similarly, Smith Magenis syndrome was initially found to be caused by a microdeletion of chromosome band 17p11.2. Subsequently mutations in the *RAI1* gene were shown to be responsible for the vast majority of the clinical features associated with the syndrome (Slager *et al*. 2003). More recent examples of variants within a single gene that are found to related to a syndrome or a sequence include the gene for CHARGE sequence (Vissers *et al*. 2004) and the gene involved in Cornelia de Lange syndrome (Krantz *et al*. 2004). In 2006, the gene linked to Peters Plus syndrome was identified after finding two splice donor site mutations within the *B3GALTL* gene (**chapter III-2**). This year, Zweier *et al*. revealed that haplo-insufficiency of *TCF4* is responsible for the Pitt Hopkins syndrome (Zweier *et al*. 2007).

I-5. Considerations regarding pathogenicity of CNVs

The vast majority of the large CNVs related to genomic disorders are thought to be *de novo* (except for CNVs with an X-linked or autosomal recessive inheritance), as affected patients often have a severe phenotype and are unable to have offspring. However, for some microdeletion syndromes an autosomal dominant transmission has been documented (Leana-Cox *et al*. 1996; Morris, Thomas, and Greenberg 1993), emphasizing that even CNVs that are known to cause genomic disorders can demonstrate phenotypic variability. The pathogenicity of familial CNVs is often hard to interpret, as variable expression of the remaining allele and incomplete penetrance can influence the clinical consequences in different family members. An example is the phenotypic variability associated with a duplication of the DiGeorge- / Velocardiofacial syndrome region. Edelmann *et al*. (1999) described an individual with this duplication who was affected by failure to thrive, marked hypotonia, sleep apnoea and seizure-like episodes. The healthy mother and grandmother however also carried the same duplication. Ad-

Figure 3. Current standard cytogenetic diagnostic tools and their characteristics.



[See appendix: colour figures.]

ditional reports verified that this specific alteration, despite showing a very wide range of clinical features, is not a benign genomic variant (Ensenauer *et al.* 2003; Yobb *et al.* 2005). A second example includes the 1.5 Mb duplication of chromosome band 16p13.1 that has been recently found among four severe autistic male patients. The same duplication was detected among less affected and unaffected family members (Ullmann *et al.* 2007).

In general, the presence of a particular CNV in a patient as well as in family members does not exclude a causal relation with the clinical problem, since autosomal recessive, digenic, complex or multifactorial inheritance can apply. The identification of the gene responsible for Peters' plus syndrome (**chapter III-2**) is the perfect example to

underline the presence of an autosomal recessive inherited disorder. This syndrome was suspected to be an autosomal recessive disorder, although cryptic unbalanced translocations could not be excluded based on the presence of multiple spontaneous miscarriages in several families. We identified an interstitial deletion in two affected brothers that was also present in the mother and the maternal grandmother. The latest two were both suffering from breastcancer. Additional investigation of the brothers identified a mutation in the *B3GLTL* gene from the same region on the paternal allele.

A *de novo* variant is often assumed to be causative, however, since many CNVs are (neutral) polymorphisms, *de novo* variations can also be inconsequential. Van Ommen (2005) discussed the frequency of *de novo* deletions and duplications. He estimated a frequency of 1 in 8 for deletions, and 1 in 50 for duplications comprising random events in human newborns. It was noted that these are likely to be underestimates as, in addition, segmental duplicons cause recurrent non-random variations. Given, therefore, that *de novo* CNV is relatively frequent and not in all cases linked to genomic disorders, the finding of a *de novo* variation in a patient is not sufficient to conclude that this CNV is causally related to the clinical phenotype.

Recent initiatives, such as those of the Sanger Institute (www.sanger.ac.uk/PostGenomics/decipher/) and Ecaruca, to create platforms for collecting and comparing molecular cytogenetic data from many clinical genetic centers in relation to the human genome sequence, will assist in giving a better understanding of the role of CNVs in MR, CM and other genetic diseases.

I-6. Detection of CNVs

6.1. (Standard) Cytogenetic tools (figure 3)

6.1.1. Karyotyping

Analysis of chromosomes using the light microscope has been the gold standard for chromosome analysis during the past five decades. The banding technique, developed in the 1970s, enables the identification of specific chromosomes and large rearrangements (Caspersson, Lomakka, and Zech 1972; Yunis 1976). Using this technique, it became clear that chromosomes from healthy individuals are not completely similar. For each and every chromosome, microscopically visible variations not related to any phenotypic trait have been identified (Wyandt HE, Tonk VS (eds), 2004. Atlas of human chromosome heteromorphisms, Kluwer). These variants are called heteromorphisms.

Karyotyping has been implemented worldwide in a diagnostic setting, as it is very specific and reproducible in detecting large chromosomal variations among different groups of patients.

Even with optimal quality, however, it is not possible to identify structural imbalances smaller that 3-5 Mb (figure 3).

The implementation of the high-resolution banding (more than 800-band level) may not always resolve the resolution problem, as it can result in both false positive and false negative results (Kuwano *et al*. 1992; Delach *et al*. 1994; Butler 1995). An example of this was published by Francke *et al*. (1985). They described a patient suffering from Duchenne muscular dystrophy, chronic granulomatous disease associated with cytochrome b deficiency and with the McLeod phenotype in the Kell red cell antigen system and retinitis pigmentosa due to an interstitial deletion of part of band Xp21. This deletion could be identified by standard resolution chromosome banding. However, using higher resolution chromosomes, the loss of genetic material was very hard to appreciate. Flint and Knight (2003) also found a negative correlation between the resolution of the banding and the number of chromosomal alterations found. This phenomenon may be explained by the fact that high resolution banding uses chromosomes that are in the prometaphase stage. At this stage the condensation of the chromatids is incomplete, resulting in elongated chromosomes. Since the condensation process is ongoing and variable during pro-metaphase, apparent differences in length may be due to unequal condensation instead of a "real" difference caused by a gain or loss of genetic material.

6.1.2. Fluorescent in Situ Hybridisation (FISH) analysis

FISH analysis (Prooijen-Knegt *et al.* 1982; Landegent *et al.* 1985; Ried *et al.* 1990) (figure 3) is based on the hybridisation of a fluorescently labelled probe containing a sequence of several tens (cosmids) to hundreds of kilobases (Bacterial Artificial Chromosomes (BACs)/ P1 derived Artificial Chromosomes (PACs)) that is complementary to the region of interest. The fluorescently labelled sequences will bind to the genomic DNA, which is subsequently visualised under a microscope. The two types of FISH analysis commonly used in diagnostic procedures are (1) metaphase FISH, that uses cultured cells for analysis, and (2) interphase FISH, that does not require culturing of cells. The advantage of interphase FISH analysis is that it has a higher resolution, allowing the detection of small tandem duplications, whereas FISH using metaphase cells will often miss such duplications as the extra signal is overlapping the original signal. Furthermore, interphase FISH can be used for the detection of low-level mosaics as large numbers of cells can be scored. On the other hand, the advantage of metaphase

FISH analysis is that individual chromosomes are visible, providing positional information of the CNV.

Detecting CNVs using FISH analysis is only possible if the following criteria are fulfilled: (1) The CNV must be characterized by a specific phenotype, (2) this phenotype must be recognized by a specialist (for example clinical geneticist) and (3) a specific diagnostic FISH test must be available.

6.1.3. Fiber FISH

Fiber FISH refers to the analysis of extended chromatin fibers. It provides a higher resolution than conventional FISH, because the chromosomes are analysed as distinct single threads under the microscope. Fiber FISH can also be used to resolve complex rearrangements. The principal drawback of this approach is that it is technically challenging and time consuming (Wiegant *et al.* 1992; Florijn *et al.* 1995; Rosenberg *et al.* 1995; Giles *et al.* 1997; Raap *et al.* 1996).

6.1.4. Multi-probe FISH (M-FISH) and SKY (Spectral Karyotyping)

Multiple color FISH was first described in the late eighties (Nederlof *et al.* 1989; Nederlof *et al.* 1990; Dauwerse *et al.* 1992). In general, Multiprobe FISH and SKY (Schrock *et al.* 1997) provide recognition of many chromosomes simultaneously by labelling them with a distinct combination of fluorochromes (Fan *et al.* 2000; Speicher, Gwyn, and Ward 1996). By pooling cloned DNA fragments of a particular (part of a) chromosome, the FISH probe can 'paint' the chromosome or a region of interest. By combining different fluorophores in different proportions, chromosome specific colors can be generated (Tanke *et al.* 1999; Raap and Tanke 2006). This COmBined RAtio labelling or COBRA–FISH is particularly useful for the detection of balanced translocations or to determine the content of a marker chromosome. As shown in figure 3, the resolution of tools is better than that of karyotyping. COBRA-FISH was used for the screening of subtelomeres (Engels *et al.* 2003). By applying the subtelomeric COBRA-FISH method, it was possible to screen 41 subtelomeres (except for the p-arms of the acrocentric chromosomes), with BACs/PACs localised approximately 230 Kb from the telomeres, using only two hybridisations and four fluorochromes.

Knight *et al.* (1997) developed a multi-hybridisation protocol, using a slide divided into 24 small hybridisation chambers. By applying different dyes to label each chromosome arm, the slide can be used to perform FISH analysis for all subtelomeres in one assay (Flint and Knight 2003). As this approach is quite laborious and consequently the throughput is very limited, it is currently not used on a wide scale.

By applying karyotyping and (different applications of) FISH analysis, a significant number of chromosomal anomalies remain undetected. Therefore, there is a strong need for screening techniques with a higher resolution.

6.2. High resolution tools (not genome-wide)

6.2.1. History

As stated previously, the phenomenon of copy number variation has been recognised since the earliest days of human gene cloning. The first gene clusters cloned, those coding for the alpha and beta chain of haemoglobin were found to frequently undergo gross rearrangements, showing deletions as well as duplications. Some, but certainly not all, of the deletions appear to be related to crossing-over between repeat elements as described by Higgs *et al.* (1984). Herrmann, Barlow, and Lehrach (1987) were the first to identify a molecular basis for recombination across a large inverted duplication that resulted in duplicated and deleted regions. For their study, which was published in 1987, restriction fragment length polymorphisms of cloned regions combined with pulse field gel electrophoresis were applied.

Studying another gene cluster, using hybridisation analysis of labelled cosmid clone fragments, Groot *et al.* (1990) hypothesized that unequal intrachromosomal crossing-over might be a frequent event leading to multiple and variable copies of the amylase genes. This model was recently confirmed using array and Fiber FISH analysis (Iafrate *et al.* 2004).

This section will briefly describe several techniques used for the detection of CNVs.

6.2.2. Restriction fragment length polymorphisms

Restriction fragment length polymorphisms (RFLP) are detected by digestion of (amplified) DNA using endonucleases, which only cut in the presence of specific DNA sequences (the restriction sites). The restriction fragments are then separated according to length by agarose gel electrophoresis. Depending on changes within these sequences, the length of the fragments and thus the position of the corresponding gel bands differ between individuals. The result of RFLP may be enhanced by Southern blotting (see 6.2.3). Using RFLP analysis, it was possible to identify duplications or deletions of a certain region of the genome. For example, RFLP analysis was applied within the first series of randomly cloned DNA fragments for the detection of probes showing non-Mendelian segregation. Both missing and extra alleles were identified (E. Bakker, personal communications, 1983).

6.2.3. Southern blotting

For many years, Southern blot analysis followed by densitometry was the main assay that was utilized for the detection of CNVs in clinical molecular genetic laboratories. It was the first technique to analyse human DNA on a wider scale. The Southern blotting procedure (Southern 1975) could show differences in length of restriction fragments and was used to study single copy, as well as low copy repeat sequences. Quantitative analysis was also possible on a very limited scale. Presence or absence of a sequence was of course no problem, but even the difference between one or two copies of a fragment with similar length required optimal experimentation. In some cases a rearrangement within a gene could be visualised by finding a new junction fragment. Since the technique required the use of radioactive labels and is very laborious, it has become less popular and has been largely replaced by quantitative PCR- based techniques, such as Q-PCR and Multiplex Ligation dependent Probe Amplification (MLPA) (Schouten *et al.* 2002).

6.2.4. Pulse field gel electrophoresis (PFGE)

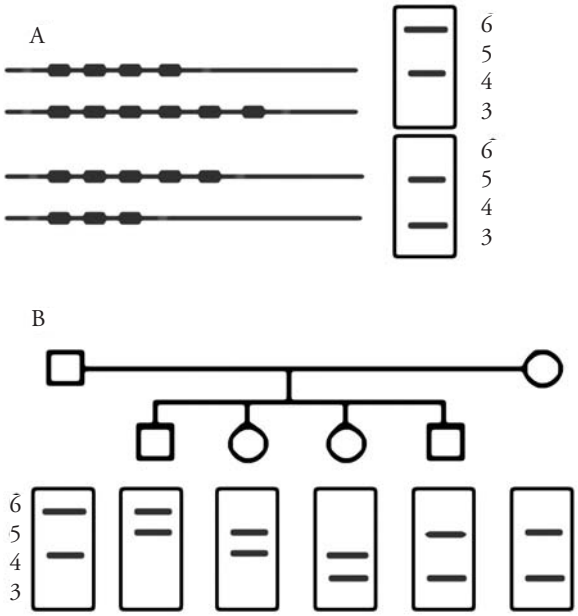
This technique (van Ommen *et al.* 1986; Den Dunnen *et al.* 1987) extends Southern blotting to include detection of very large DNA molecules (20 kb to several Mb in length) that are too large to be separated using normal agarose gel electrophoresis. It can be used to detect a rearrangement-specific junction fragment. Shearing of the genomic DNA is prevented by preservation and enzymatic digestion in solid agarose. The agarose-embedded DNA is cut by a rare-cutting restriction endonuclease and subsequently separated by an electrical current. During electrophoresis, the relative orientation of the electric field is periodically altered (Strachan and Read, Human Molecular Genetics, third edition, chapter 6.2). Fragments of different sizes will migrate at different speeds through the gel, and consequently PFGE is capable of detecting structural rearrangements.

Despite being technically challenging, is still used to study large repeat arrays e.g. FSHD (Buzhov *et al.* 2005).

6.2.5. Microsatellites for detecting CNVs

Microsatellites are sequences containing variable number of tandem repeats (hence are also known as variable number of tandem repeat markers (VNTRs). The number of repeat units for a given locus may differ between individuals, resulting in alleles of varying lengths. The differences in repeat length can be visualised either by using a nearby single copy probe on a Southern blot or by PCR-based methods. Allelic variation, the number of repeats, and allelic frequencies are available for thousands of markers across numerous

Figure 4. Identification of the parental origin of an allele.



A. Different VNTR lengths in both parents present on a specific region in the human genome.
B. One of the children has the identical combination of VNTR lengths as one of its parents. Uniparental disomy (of genetic material from the parent with identical VNTR lengths) or a deletion present at the allele inherited from the 'other' parent should be considered. Picture derived from www.geninfo.no.

[See appendix: colour figures.]

organisms. These polymorphisms can be used for the identification of CNVs by observing abnormal inheritance of parental alleles (figure 4), such as uniparental disomy. The limitation of this type of genetic marker for the detection of imbalances is that its success depends on the availability of parental DNA (Wilke, Duman, and Horst 2000).

All techniques described above have major disadvantages. They are either technically demanding, expensive, slow, require fresh samples, or have a low throughput (Heath, Day, and Humphries 2000). The major limitation is the small number of loci that can be tested in one experiment. The development of PCR based techniques, such as Multiplex Amplifiable Probe Hybridisation (MAPH) and Multiplex Ligation-dependent Probe Amplification (MLPA) allowed more widespread analysis of gene dosage.

6.2.6. Quantitative real-time Polymerase Chain Reaction (Q-PCR)

This method is independent of the availability of informative markers in the region of interest. Quantitation of input DNA is achieved by using dyes or dual-labelled probes, and a fluorescence scanner to monitor the amount of product generated during the amplification process. The method was originally designed to facilitate quantification of RNA, but it can also be used to quantify the copy number of a genomic sequence. The combination of real-time PCR and TaqMan TM fluorescent probes for the detection of CNVs has been described by Wilke, Duman, and Horst (2000) and Laurendeau *et al*. (1999). In this case, one only needs the amplification of one reference locus to measure the copy number of the test loci, instead of using different diluted DNA fragments for standardisation.

6.2.7. Towards MAPH and MLPA

In 1995, a PCR method was described which simplifies quantitative multiplex PCR (Shuber, Grondin, and Klinger 1995) where gene specific primers were tagged at the 5'end with an unrelated 20 nucleotide universal primer binding site. Based on this method, new applications of multiplex-PCR were designed such as quantitative fluorescent multiplex PCR (QFM-PCR) (Heath, Day, and Humphries 2000) that was published in the same year as Armour published another application, called Multiplex Amplifiable Probe Hybridisation, MAPH (see below). QFM-PCR, MAPH (**section 6.2.6**), MLPA (**section 6.2.7**.) are all useful, effective and reliable methods for the detection of both deletions and duplications in the same assay.

6.2.8. MAPH

MAPH was first described by Armour *et al*. (2000). MAPH is a PCR-based method for simultaneously determining the copy number of a set of up to 50 different chromosomal loci (White *et al*. 2002). The probes, usually exons from candidate genes, are individually cloned such that all can be amplified using one pair of primers. To detect copy number changes, the probes are hybridised to denatured genomic DNA that has been immobilised and cross-linked on numbered nylon filters. After stringent washing, only the probes that hybridise specifically to the complementary sequence on the genomic DNA will remain bound. These hybridised probes are recovered off the filters, quantitatively amplified using PCR and analysed. The initial publication used a radioactively labelled primer followed by separation on a slab gel. This was then exposed to a film, with the resulting bands being measured using densitometry. White *et al*. (2002) simplified the procedure by using a fluorescently labelled primer followed

by analysis using a 96 capillary sequencer. The yield, represented by peak height and area, is determined for each probe. Changes in probe yield correspond to changes in copy number of the sequence analysed, i.e. a deletion or duplication.

The first report of subtelomere screening in patients with MR using MAPH was from Sismani *et al.* (2001). In their study, a group of 70 mentally retarded individuals was screened, using multiprobe telomeric FISH assay and MAPH. One subtelomeric deletion was found and confirmed with an independent technique. It has to be mentioned, however, that not all the subtelomeric probes were informative.

It has been calculated previously (Hollox *et al.* 2002), that about 0.12% of the mentally retarded patients were reported to have false positive results (that is, MAPH analysis detected an alteration that could not be verified using an independent technique), using MAPH based screening of subtelomeres, suggesting that this technique is reliable for the detection of CNVs. Obviously, the percentage depends highly on thresholds applied in a certain study.

6.2.9. MLPA

MLPA is based on the ligation of two adjacently annealing oligonucleotides, followed by the quantitative PCR amplification of the ligated products (Schouten *et al.* 2002). The left half-probe is chemically synthesised. It consists of a unique sequence complementary to the locus of interest along with a sequence containing the primer-binding site common to all probes. The other half-probes consist of three parts. In addition to the parts present in the left half-probe, this right half-probe also contains a spacer sequence, responsible for the difference in length of the MLPA probes. As the size of the right-sided half-probe initially was designed up to 440 nt, it was not possible to synthesize this oligonucleotide. Therefore, M13 vectors were used carrying the spacer sequences. However, generating a right half-probe with a spacer requires a laborious and time consuming cloning step. Therefore, a modified protocol for designing probes was implemented (White *et al.* 2004). Using this protocol, the right half probe is also chemically synthesised followed by 5'phosphorylation. Each probe was designed to be of unique size, enabling easy differentiation. This alternative MLPA protocol significantly reduces the time necessary for MLPA probe design, however, the number of loci that can be tested by MLPA using one fluorescent dye is limited. A second (and even a third) dye can be used by designing probes with another primer binding sequence (White *et al.* 2004; Harteveld *et al.* 2005). In this way, it is possible to screen up to 60 loci in nearly 100 patients in one assay.

6.2.10. Data analysis of MLPA and MAPH

Several methods for data analysis have been described (Hollox *et al*. 2002; White *et al*. 2002) and analysis protocols are available at www.mlpa.com.

Besides analysing the result of MLPA and MAPH using either a polyacrylamide gel or through polymer-filled capillaries, both techniques can be adapted for an array- or bead based read out. This will increase the number of loci than can be tested simultaneously in one patient (Gibbons *et al.* 2006). To detect the amplified fragments, universal arrays can be designed using specific zip codes. These are spotted on the array, with the complementary sequences being incorporated into the probes. An added advantage of this approach is that the half probes used can have identical sizes, facilitating uniform amplification. Using the 3-Dimensional, Flow-Through Microarray Platform from PamGene, hybridisation time of the amplified fragments to their target sequences can be reduced to minutes. This technique has been used for the rapid detection of aneusomies, resulting in a gain in time of more than 60 hours compared to karyotyping (Kalf *et al.* in preparation).

The advantage of MAPH and MLPA compared to other techniques, including (multi-probe) FISH and array-CGH, is that the resolution of detection is limited only by the size of the probes used (100-500 bp). In addition, using specific probe design, it is even possible to detect point mutations using MLPA analysis.
Both MAPH and MLPA facilitate the parallel screening of large numbers of patients at many different loci in one experiment with rather cheap consumables.
A disadvantage of these methods is that they are not suitable for genome-wide screening.

6.3 Whole genome (high resolution) screening tools; recent genome approaches

6.3.1. Overview

Affordable, high-resolution, genome-wide approaches for DNA copy number analysis have been available for less than five years. In contrast to FISH, where small fragments of DNA are labelled and hybridised to genomic DNA (in the form of chromosome spreads), array-based approaches label the genomic DNA, which is then hybridised to small fragments of DNA.

Currently, there are two main formats, array-CGH and SNP-based arrays. Both are discussed in more detail below. For array-CGH, the probes used are (3K – 30K) genomic clones or up to 400K 60-mer oligonucleotides, with the size and number determining the resolution of analysis.

SNP arrays, containing 10K–1000K loci have recently proven to facilitate, in addition to genome-wide association studies, the detection of deletions and duplications (see section 6.3.4.). The resolution of the SNP arrays depends on the number of SNP loci present and on their coverage across the genome.

The coverage of the genome of all genome-wide mapping platforms is rapidly improving.

It should be noted that these tools can not be used to detect copy-neutral rearrangements like translocations, insertions and inversions.

6.3.2. Array-CGH using BAC clones

High-resolution comparative genomic hybridisation (CGH)-based micro-arrays (Solinas-Toldo *et al.* 1997; Pinkel *et al.* 1998; Snijders *et al.* 2001) were developed to increase the resolution of chromosome studies. The technique is based on immobilised DNA isolated from Bacterial Artificial Chromosome (BAC) clones that were amplified by either DOP-PCR (Telenius *et al.* 1992) or ligation-mediated PCR (Snijders *et al.* 2001). The amplified DNA, spotted on coated microscope slides by an arrayer, is usually present in triplicate enabling internal standardisation. Test and reference DNA are differently labelled by random priming to incorporate fluorescently labelled nucleotides, and subsequently mixed with Cot –1 DNA to block repetitive DNA sequences. After hybridisation for 16-24 hours, images of hybridised fluorochromes can be obtained. The resolution obtained with BAC-arrays depends on the genomic distance between the BACs spotted on the array and the size of the BACs (Snijders, Pinkel, and Albertson 2003).

Clinical applications of array-CGH using different subsets of the human genome have been published by several groups (Veltman *et al.* 2002; Rauen *et al.* 2002; Bruder *et al.* 2001; Rosenberg *et al.* 2006). Veltman *et al.* (2002) estimated, based on their results obtained by screening 20 patients with known cytogenetic abnormalities, that the incorrect positive result of the 3500 BAC-array is approximately 0.4%, whereas no abnormality was missed. Many papers have been published regarding findings of screening MR patients using BAC-array of ~3500 BAC DNA probes spaced at ~1 Mb density over the full genome (3K array) (table 3). De Vries *et al.* (2005), Vissers *et al.* (2005) and Koolen *et al.* (2006) presented the results of screening using a BAC array with 10 fold higher resolution (33000 BACs). BAC arrays are also widely used in cancer diagnostics (Snijders *et al.* 2003; Weiss *et al.* 2003). The genomic variation among 55 healthy individuals was also tested using array-CGH (Iafrate *et al.* 2004). This study found as many as 255 alterations that were suspected to be neutral variants.

BAC-based array-CGH has been very important for the initiation of genome-wide screening at high resolution. It has proven to be a reliable and reproducible technique. Recently, oligonucleotide-based arrays have become available. These arrays come in two types, 60-mer oligos (see section 6.3.3.) for the detection of small CNVs and shorter 25-mer oligos for SNP (see section 6.3.4.) detection.
In their latest versions, these arrays have an effective resolution below 10 kilobases. A disadvantage of array-based methods is that they are currently still rather expensive.

6.3.3. Array-CGH using long oligos

Examples of these arrays include Nimblegen and Agilent. The 60 nucleotide is longer than the sequence that is spotted on the SNP array. As a result, these oligo based arrays are not suitable for SNP analysis, however, they do give stronger signal intensity. Therefore, CNVs can be detected using solely the signal intensity.

In addition, as the location of the oligos is not limited to known SNPs, it is possible to analyse regions of the genome where no validated SNPs are available. This can be particularly important when looking at duplicated regions. The most recent Agilent micro array contains ~244.000 spots on the array.

6.3.4. SNP based arrays

The 25-mer probe arrays were originally designed to detect SNPs to be used in genome wide linkage and association studies. However, they were quickly used to estimate copy number changes by using both signal strength and allele scoring. Initial studies used the Affymetrix 10K array, which demonstrated the principle that the arrays could provide quantitative data (Herr *et al*. 2005). Subsequent work has taken advantage of higher resolution chips, currently up to 500-1000K (Komura *et al*. 2006). In practice, these arrays have an effective resolution below 10 kilobases, meaning that much smaller rearrangements can be detected compared to previous genome-wide technologies.

6.3.5. Comparing cross platform

Currently, there is no golden standard available to determine which platform, CGH-based or SNP-based, is the most accurate. It might be argued that high density SNP genotyping would be the most appropriate to implement for screening for copy number alterations, as this tool offers the simultaneous measurement of copy number changes and copy-neutral loss of heterozygosity (i.e uniparental disomy). On the other hand, SNP arrays have been selected based on criteria such as heterozygosity, being in Hardy-

Weinberg equilibrium. Although these features are important for association studies, where SNPs need to be informative, they are less critical for copy number analysis where even spacing is more important. Indeed, many regions prone to rearrangements (e.g. duplicons) are lacking or underrepresented on these arrays, as the associated SNPs did not meet the required quality criteria. This is in contrast to array-CGH in which the location of the oligonucleotides is not limited to known SNPs, and, therefore, it is possible to analyse regions of the genome where no validated SNPs are available. Indeed, the study of Redon *et al.* (2006) shows that in addition to the SNP-arrays, arrayCGH analysis is required to cover all CNV regions in the human genome, otherwise at least one third of the CNVs will be missed. New arrays of both Affymetrix and Illumina now close this gap by combining SNP- and non-SNP probes on one array.

Chapter III-4 attempts to compare different whole genome screening tools by applying them to four unrelated patients suffering from overlapping interstitial 2p deletions. Comparing cross-platform, we found that the localisation of both proximal and distal breakpoints was largely in agreement.

There have been few studies published screening MR patients with the new oligo-array platforms (table 3). Most studies described to date looked at either CNVs in healthy individuals (table 4) or the validation of techniques for detecting CNVs in patient populations. Using the 10K genechip of Affymetrix, seven known alterations with a size between 0.2-3.7Mb were not detectable due to insufficient SNP density in the regions involved (Rauch *et al.* 2004). Slater *et al.* (2005) were able to find all known alterations previously found by karyotyping, FISH or MLPA analysis using a ten-fold higher density (>110 K) SNP chip of Affymetrix, except for one duplication at the end of chromosome 9q. The same mapping tool was successfully validated by another group (Ting *et al.* 2006). The utility of the beadchip (SNP) array of Illumina, assaying 109,000 and 317,000 SNP loci, to detect chromosomal aberrations in samples bearing constitutional aberrations as well tumor samples at sub-100 kb effective resolution has also been described (Peiffer *et al.* 2006). In addition, summaries of different whole genome high resolution mapping tools have been published recently (Veltman 2006; Coe *et al.* 2007).

I-7. Scope of this thesis

The main aim of this thesis was to assess several new techniques for the detection of genomic rearrangements in patients with MR and / or CMs. In quick succession,

Table 3. A selection of studies using genome-wide screening tools to screen for CNVs in MR patients.

References	Methods of Analysis	Genome Coverage	Sample size	No. of dels. (*de novo*)	No. of duplications (*de novo*)	U.T	% Alterations (% *de novo*)
Vissers *et al.* (2003)	BAC arrays	3,500 BACs	20 MR patients	3 (2)	2 (1)	0	25% (15%)
Schoumans *et al.* (2005)	BAC array	2,600 BACs	41 MR patients + dysm. features	4 (4)	0	0	9.8% (9.8%)
Tyson *et al.* (2005)	BAC array	3,000 BACs	22 MR patients	1 (1)	2 (1)	0	14% (9%)
De Vries *et al.* (2005)	BAC array	33,000 BACs	100 MR patients	Many (7)	Many (3)	0	10% (10%)
Menten *et al.* (2006)	BAC array	3,500 BACs	140 MR patients	18 (11)	7 (3)	3	20% (10%)
Miyake *et al* (2006)	BAC array	2,173 BACs	30 MR patients	3 (1)	1 (1)	1(1*)	17% (10%)
Rosenberg *et al.* (2006)	BAC array	3,500 BACs	80 MR patients	12 (5)	6 (2)	2 (1*)	25% (10%)
Shaw-Smith *et al.* (2006)	BAC array	3,500 BACs	50 MR patients + dysm. features	7 (6)	5 (1)	0	24% (14%)
Ming *et al.* (2006)	Affymetrix gene chip	100K SNPs	10 MCA patients	2(2)	0	0	20% (20%)
Friedman *et al.* (2006)	Affymetrix gene chip	100 K SNPs	100 MR patients	8 (8)	(3) (1 was a mosaic)	0	11 (11%)
Sebat *et al.* (2007)	ROMA	85,000 oligos	195 autistic patients	12 (12)	3 (3)	0	7,7% (7,7%)

This table summarizes the eight studies screening MR patients using BAC arrays, and three studies screening a MR or autistic study population using oligo based arrays. Based on the data presented in this table, it shows that, independent of the sample size tested, the number of *de novo* alterations detected using whole genome screening tools is around 10%. It is noteworthy that although the number of loci tested using a BAC-array is increased significantly compared to the initial BAC-arrays, the number of *de novo* alterations detected remains 10%. The same holds true for the implementation of the 100K SNP array.
*: one of the parents is a carrier of a balanced translocation. Affy: SNP array designed by Affymetrix, ROMA: representational oligonucleotide microarray analysis, dels: deletion, U.T.: unbalanced translocation, dysm.: dysmorphic

MAPH, followed by MLPA, and MLPA in combination with array-CGH, have been implemented to expand the possibilities for diagnostic screening for deletions and duplications. By applying these high-resolution techniques, new regions and genes involved in the etiology of MR/CM were identified, resulting in an increased number of patients with a known cause for their developmental disorders. Currently, using the new genome-wide high(er) resolution techniques, such as the oligo based array, the number of variations detected in the human genome will increase even further. At this

Table 4. The results of screening for CNVs among healthy individuals using different whole genome screening tools.

References	Methods of Analysis	Genome Coverage	Sample size	Total No of CNVs
Iafrate *et al.* (2004)	BAC array	5,264 BACs	55 healthy individuals	255
Sebat *et al.* (2004)	Oligo based array (ROMA)	85,000 oligo nt	20 healthy individuals	221
Conrad *et al* (2006)	Mendelian errors	1,3 million genotyping assays	180 healthy individuals (3* 60)	586
Mc Carrol *et al.* (2006)	Clustered genotype & Mendelian errors (Hapmap data)	1,3 million genotyping assays	269 healthy individuals	541
Komura *et al.* (2006)	Affymetrix gene chip	500 K	270 healthy individuals	1,203
Redon *et al.* (2006)	Array-CGH & affymetrix gene chip	26,574 clones 500 K	270 healthy individuals	1,447

moment, the consequence of the detection of a CNV in an affected individual is not always clear. Therefore, the main challenge will be determining whether a variation is related to disease or one of the many neutral polymorphisms.

I-8. In summary

The following two chapters contain seven papers. Chapter II includes three studies where groups of patients were tested for CNVs. The frequency of subtelomeric alterations as well as interstitial variations in and outside duplicons were determined among different groups of mentally retarded patients. We were able to report the second patient with the reciprocal duplication of the Williams syndrome critical region and a previously undescribed duplication within the 16p13.1 region. In addition, based on our findings using parallel testing of both MLPA- and array based analysis, an alternative, cost effective approach is recommended for screening mentally retarded patients. Chapter III is comprised of four studies using small numbers of patients and a case report. The first report describes a complex rearrangement on both copies of chromosome 22. Different characteristics of the rearrangements were defined using different diagnostic tools. We found that haplo-insufficiency of the Cat eye critical region is probably not related to a clinical phenotype. The phenotypic variability in relation

to the size of the deletion of patients having the ATR-16 (α-thalassemia retardation-16) syndrome was explored in the next paper. It was concluded that in MR patients showing microcytic (= small cell) hypochromatic anemia, the presence of ATR-16 syndrome should be excluded.

Thirdly, we were able to unravel the etiology of the Peters Plus syndrome, an autosomal recessive inheritable disorder, using a genome-wide screening tool. Finally, four high resolution genome-wide mapping tools were compared using four patients with an overlapping interstitial 2p deletion.

Chapter II

Screening ‘large’ patient groups

Chapter II-1

Genomic imbalances in mental retardation

M Kriek[1,*], S J White[1,*], M C Bouma[2], H G Dauwerse[1], K B M Hansson[1], J V Nijhuis[1], B Bakker[1], G-J B van Ommen[1], J T den Dunnen[1] and M H Breuning[1]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, The Netherlands; [2]Department of Clinical Genetics, University Hospital Groningen, The Netherlands
*The first two authors contributed equally to this work

J. Med. Genet. 2004;41;249-255

Summary

Introduction: It has been estimated that cytogenetically visible rearrangements are present in ~1% of newborns. These chromosomal changes can cause a wide range of deleterious developmental effects, including mental retardation (MR). It is assumed that many other cases exist where the cause is a submicroscopic deletion or duplication. To facilitate the detection of such cases, different techniques have been developed, which have differing efficiency as to the number of loci and patients that can be tested.
Methods: We implemented multiplex amplifiable probe hybridisation (MAPH) to test areas known to be rearranged in MR patients (for example, subtelomeric/pericentromeric regions and those affected in microdeletion syndromes) and to look for new regions that might be related to MR.
Results: In this study, over 30 000 screens for duplications and deletions were carried out; 162 different loci tested in each of 188 developmentally delayed patients. The analysis resulted in the detection of 19 rearrangements, of which ~65% would not have been detected by conventional cytogenetic analysis. A significant fraction (46%) of the rearrangements found were interstitial, despite the fact that only a limited number of these loci have so far been tested.
Discussion: Our results strengthen the arguments for whole genome screening within this population, as it can be assumed that many more interstitial rearrangements would be detected. The strengths of MAPH for this analysis are the simplicity, the high throughput potential, and the high resolution of analysis. This combination should help in the future identification of the specific genes that are responsible for MR.

44

Introduction

The evolution of the human genome has resulted in mixture of large and small interspersed and tandem segmental duplications throughout the genome. Such duplications provide substrates for homologous recombination, and consequently, the intervening regions show considerable rate of rearrangement.[1–3] Many of these rearrangements occur in regions where a change in gene dosage does not affect human health. However, after the description by Lejeune of trisomy 21 in Down's syndrome,[4] and the many subsequent publications on different aneuploidies, it became clear that the genome contains many loci for which the correct copy number is critical for normal development. Change in genetic dosage of one or more genes is one of the most common causes of mental retardation (MR). Examples of known important loci include the subtelomeric regions and the areas involved in microdeletion syndromes.

The subtelomeric regions, localised proximal to the telomeres, have been found to be especially susceptible to copy number changes, owing to repeat rich sequences that show a high frequency of recombination.[1] It has been hypothesised that about 6% of the patients with idiopathic MR will have a subtelomeric rearrangement,[5] a figure confirmed in several studies that have reported a frequency of 2–9% of cryptic rearrangements in MR patients.[6,7]

The cause for MR is only established in approximately 50% of cases, limiting the efficiency of genetic counselling, detection of carriers, and prenatal diagnosis in these families. This rather low percentage of diagnosis may have several explanations. A routine cytogenetic analysis gives a minimum resolution of only 4–10 Mb. Fluorescent in situ hybridisation (FISH) largely overcomes this limitation of resolution; however, it can only be applied to simultaneously test a limited number of chromosome regions. FISH is therefore mostly used to confirm well recognised microdeletion syndromes in patients who present a suggestive phenotype. Another potential explanation is that the genome contains undiscovered loci that are involved in the aetiology of MR. New technologies, such as multiplex amplifiable probe hybridisation (MAPH),[8] multiplex ligation dependent probe amplification (MLPA),[9] and array based comparative genomic hybridisation (array CGH),[10] have recently been developed to search for such undiscovered regions. We chose to implement a high resolution, high throughput, rapid, and simple method, MAPH,[8] which allows the simultaneous screening at the exon level for copy number changes of 40–50 different chromosomal loci in up to 96 patients in one assay. Hollox *et al.*[11] previously described subtelomeric screening using MAPH of patients with a developmental delay. In our study, we screened loci known to be involved in MR (subtelo-

meric/pericentromeric regions and genes involved in microdeletion syndromes) as well as interstitial genes randomly spaced throughout the genome. A total of 30 000 gene dosage screens were performed from 188 cases with unexplained developmental delay that were each scanned for copy number changes at 162 loci. We were able to detect subtelomeric, pericentromeric, and interstitial rearrangements in a group of patients with MR and dysmorphic features and/or multiple congenital abnormalities, as well as in patients selected solely on the basis of developmental delay.

Subjects and methods

Probe design and MAPH

The probe design has been previously described,[12] using unique sequences only. The primers of the chosen sequences were designed using Prophet (http://www.basic.nwu.edu/ biotools/prophet.html), and supplied by Invitrogen Life Technologies. Products were amplified from genomic DNA by PCR and cloned into the pGEM-T easy vector (Promega). The correct insert was confirmed by sequencing with the BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems) at the Leiden Genome Technology Center, using an ABI 3700 Sequencer (Applied Biosystems).

MAPH was performed as described by White *et al.*[12] (see also Leiden Muscular Dystrophy Pages (http://www.dmd.nl/ DMD_MAPH.html)).

Study population

The DNA of 188 patients (110 males and 78 females) from the Center for Human and Clinical Genetics Leiden (a DNA diagnostic laboratory) was analysed. The patients had been seen by a clinical geneticist or a paediatrician and diagnosed with developmental delay. The study population was divided into two groups. The first group contained 123 coded patients who had been referred for fragile X screening. Before testing, information about the results of additional tests, such as karyotyping, was not known to the investigators. The second study group (n = 65) was known to have a normal karyotype and had tested negative for fragile X screening. All patients had (multiple) congenital malformations or dysmorphic features in addition to psychological developmental delay.

Data analysis

The data were analysed with GeneScan Analysis and Genotyper Software (Applied Biosystems). These programs provide information about the length, peak height,

and peak area of the DNA fragments. Peaks were not used for analysis if they were outside predefined thresholds (upper and lower limits of 12 000 and 150 units, respectively). To obtain a ratio, the height of a given peak was divided by the sum of the heights of the four nearest peaks. As it is not likely that all four probes from diverse regions of the genome are altered in one patient, adding unrelated standards was not necessary in most of the probe sets. For the chromosome 22 probe set, however, unrelated probes, containing sequences from other chromosomes, were used as references.

The median ratio for each probe within a single hybridisation (minimum number of samples 8; maximum number 12) was determined and used to calculate a normalised ratio for each patient. Within each patient, initial "normal" thresholds were set as 0.75 and 1.25. The standard deviation from the ratios within these limits was calculated, and three times this standard deviation was used as the threshold for any given patient. Any probe that was outside these limits was retested, and samples that showed an apparent copy number change in duplicate were examined further using other techniques. Samples that showed a standard deviation of .10% over probes within the normal thresholds were retested.

Verifying the MAPH results

Copy number changes detected by MAPH were verified using another technique, primarily FISH with a bacterial artificial chromosome (BAC) or cosmid probe covering the appropriate genomic region. The BACs used were designed by Flint,[13] or supplied by Vysis Abbott Laboratories (TV, Telvysion, LSI, locus specific identifiers) or selected from the RPCI human BAC library. The FISH experiments were performed following standard operating procedures as described in Dauwerse *et al.*[14] Some MAPH results were verified using MLPA.[9]

Results

Genotyping

We designed several probe sets covering both the subtelomeric/pericentromeric and interstitial regions, including genes involved in microdeletion syndromes, genes on chromosome 22, and genes spread across all chromosomes (table A, supplemental). The subtelomeric probe set is composed of probes corresponding to the 41 subtelomeric regions, preferably an exon of a gene within 1 Mb from the telomere, five genes

near the centromere on the q arm of the acrocentric chromosomes, a sequence in the pseudoautosomal region of chromosome Xq and Yq, and an exon of a Yp specific gene. The microdeletion probe set was made up of 27 probes from 21 different genes involved in microdeletion syndromes (Williams, Prader Willi, Angelman, Smith-Magenis, Sotos, 22q11, Alagille, and Wolf-Hirschhorn syndromes). The chromosome 22 probe set included 19 probes from genes on chromosome 22 with approximately 1 Mb spacing. Finally, we used two probe sets containing a total of 68 interstitial genes spread throughout the genome.

We applied these probe sets following two methods of validation. Firstly, a probe was considered to be reliable when the standard deviation over 12 unaffected samples (one hybridisation) was <15%. Secondly, where possible, we verified the unique and correct localisation of the probes using DNA from patients with known aberrations (42% of the subtelomeric probes, 70% of the microdeletion probes).

Overall, 188 patients were screened for deletions and duplications at 162 loci, resulting in the detection of 19 copy number changes. Of these, four aberrations turned out to be cytogenetically visible, namely an isochromosome 18p (karyotype 47, XY, +i(18p)), a marker chromosome (karyotype 47, XY, +mar.ish der(22)t(8;22)(q24.1;q 11.2)), a triple X female (karyotype 47, XXX) and a Turner syndrome (karyotype 45, X), because the outcome of additional investigations had not been made known to the investigators before testing. These patients and their corresponding aberrations were not included in the calculation of the percentage of rearrangements found by MAPH; however, they emphasise the usefulness of MAPH for detecting copy number changes.

In total, eight subtelomeric/pericentromeric rearrangements were found (table 1; upper part). Five of these mutations were detected in the group of MR patients with additional dysmorphic features or additional congenital malformations (5/65 = 7.7%) and the remaining three subtelomeric aneusomies were diagnosed in the group selected on the basis of developmental delay only (3/123 = 2.4%). The smallest mutation found was a deletion of 110 kb maximum present in chromosome band 7p22.3 (table 1, F; and data not shown). Seven rearrangements were interstitial mutations. These are summarised in the lower part of table 1. Where possible, the DNA of both parents of these patients was tested; 75% (9/12) were shown to be *de novo*. The duplication of 14q11.2 (table 1, O) and the 7ptel deletion (table 1, F) were also found in the parental DNA, and one of the parents of patient E was a balanced translocation carrier.

As the number of cytogenetically detectable aberrations is highly dependent on the banding resolution, the karyograms of all 15 patients with a MAPH detected rear-

rangement were re-examined. At a resolution of 500–550 bands per haploid set, the karyograms showed that two subtelomeric copy number changes should have been detected cytogenetically (table 1; A, C). The detection of a 1ptel deletion (table 1, H) was doubtful; however, the duplication of 1ptel (table 1, H) was picked up. This implies that although the presence of the copy number change was known, 63% (12/19) of these genomic changes found in this study were cytogenetically undetectable using karyotyping at a resolution of 500–550 bands.

Case descriptions

Case 1

This 15 year old girl was diagnosed with total anomalous pulmonary venous return, hearing loss in combination with a narrow external auditory meatus, and MR. Physical examination at the age of 14 years showed a short stature (–3 SD) and some facial dysmorphic features (small palpebrae, broad mouth, thin upper lip). Karyotyping at a resolution of 400 bands and FISH studies of the 22q11 region did not detect any rearrangements. MAPH study showed a *de novo* deletion of the subtelomeric region of 18q, which was confirmed by FISH using probe TV18q. The clinical features of this patient are consistent with those of the 18q syndrome phenotype.[15]

Case 2

A male patient, who had previously tested negative for Williams syndrome, was diagnosed with a *de novo* deletion of 16ptel by MAPH. FISH analysis confirmed this finding and limited the proximal breakpoint to chromosome band 16p13.3, distal to the *PKD1–TSC2* (LocusLink 5310–7249) gene cluster[16] using probe COS15A. As expected, owing to the location of the alphaglobin gene (*HBA1*; LocusLink 3039) in this region (16p13.3),[17] further investigation showed that this patient had mild anaemia (alpha thalassaemia heterozygosity) in addition to his moderate mental handicap and dysmorphic features.

Case 3

This boy was seen by a clinical specialist at the age of 2.5 years for his psychomotor retardation and joint hyper-flexibility. Physical examination showed few dysmorphic features (a tent shaped mouth), hypotonia, and hypermobility. MAPH analysis revealed a *de novo* deletion within chromosome band 17p11.2 corresponding to the Smith-Magenis syndrome (SMS) region, using a probe for the *DRG2* gene (LocusLink: 1819). The more distally located COPS3 gene (LocusLink: 8533) showed two copies

Table 1. An overview of all 15 patients (A–O) with MAPH detected subtelomeric/pericentromeric and interstitial aneusomies.

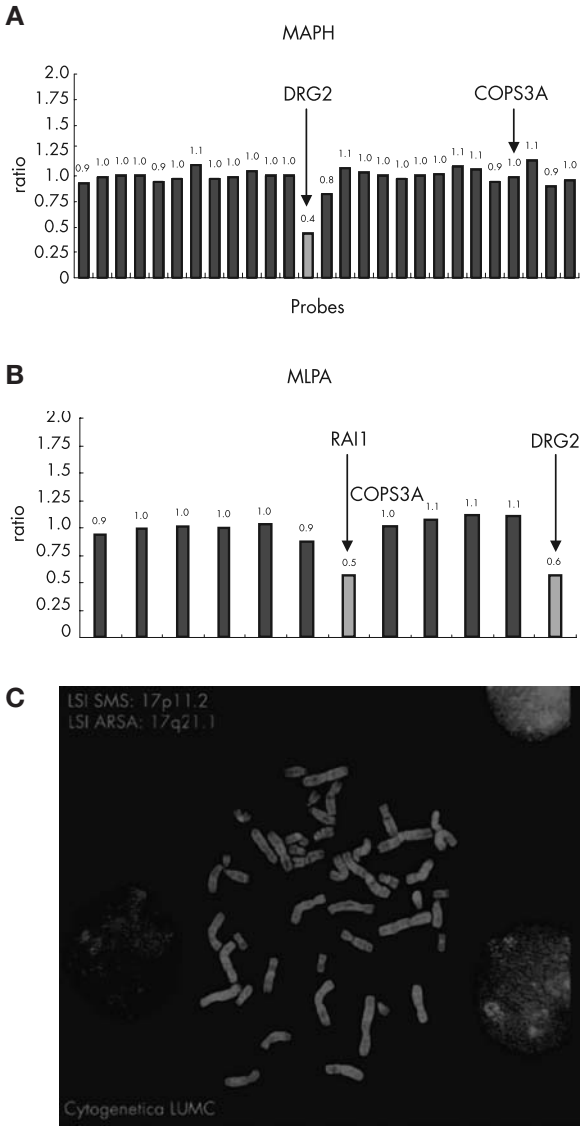
Case		Aneusomy	Group	Gender	Confirmed by	Cytogenetically visible	Clinical features	Pathogenic	References
Subtelomeric/Pericentromeric									
A	1	Deletion 18q22.1	MR++[#]	Female	FISH[¶] clone ID:TV 18q	Yes: 500–550 bands	MR, small stature, hearing loss, TAPVR,[††]mild facial dysmorphism, tapering fingers	Yes	Many: latest are[15 33]
B	2	Deletion 16p13.3	DD only[§]	Male	FISH clone ID;COS15A	No	Moderate MR, mild facial dysmorphism, mild alpha thalassemia	Yes	Many: latest is[17]
C		Deletion 6p25	DD only	Male	FISH clone ID: TV 6p	Yes: 500–550 bands	Moderate MR, iris dysplasia, excentric pupil, hypertelorism, hearing loss	Yes	[34 35]
D		Deletion pericentromeric region of chr. 22, duplication of 22q11.2	MR++	Male	FISH clone ID: RP11_ 3018K1	No	Mild MR, hearing loss, palatoschisis, cataract, microcephaly, double set of teeth	?	Kriek *et al*[†]
E		Deletion 6qtel, duplication 20qtel	MR++	Male	FISH clone ID:57H24 (6q), 81F12 (20q) MLPA**	No	MR, hypotonicity, microcephaly, brain anomalies, mild facial dysmorphism.	*	
F		Deletion 7ptel	DD only	Male	MLPA	No	Mild developmental delay in early childhood, mild facial dysmorphism	No/?	[36]
G		Duplication 1ptel	MR++	Female	FISH clone ID: 785P20, 37J18	Yes: 500–550 bands	Psychomotor developmental delay, double sided ptosis, parasis of VI cranial nerve, strabismus	?	37
H		Deletion 1ptel	MR++	Female	FISH clone ID: 465B22, 37J18	Doubtful: 500–550 bands	Psychomotor developmental delay, dysmorphic features, hirsutism, epilepsy	Yes	38

Case		Aneusomy	Group	Gender	Confirmed by	Cytogenetically visible	Clinical features	Pathogenic	References
Interstitial									
I		Duplication 17p11.2	MR++	Female	FISH clone ID: LSI-SMS	No	MR, microcephaly, retrognathia, tapering acra, hypertelorism, synophrys, epilepsy	?	39
J	3	Deletion 17p11.2	DD only	Male	FISH clone ID: LSI-SMS, MLPA	No	Psychomotor developmental delay (speech delay), infantile hypotonicity, tent shaped mouth	Yes	Many: latest is40
K	4	Deletion 4q34.1	DD only	Male	FISH clone ID: RP11-475B2	No	Mild learning disability, short stature, severe delay of bone maturation, aberrant hand shape	Yes	19
L	5	Duplication 20p12.2	DD only	Male	MLPA	No	Mild MR, psychiatric disorder	?	21
M	6	Duplication 22q11.2	MR++	Female	FISH clone ID: LSI TUPLE1	No	Severe psychomotor retardation, short stature, microcephaly, facial dysmorphism, epilepsy, brain anomalies, renal aplasia	?	41 42
N		Deletion 22q11.2	MR++	Female	FISH clone ID LSI TUPLE1	No	Developmental delay, tetralogy of Fallot, absent pulmonary valve, respiratory complications	Yes	Many: latest is43
O		Duplication 14q11.2	DD only	Male	MLPA	No	MR, mild facial dysmorphism, short hands and feet, shawl scrotum	No/?	

After the verification of these imbalances by FISH or MLPA, the karyograms of the patients were re-examined at a resolution of 500–550 bands. The results obtained are shown in the column 'cytogenically visible'. The clinical features known to be related to the rearrangement found by MAPH are highlighted. The presence or absence of a genotype–phenotype correlation is summarised under "Pathogenic".

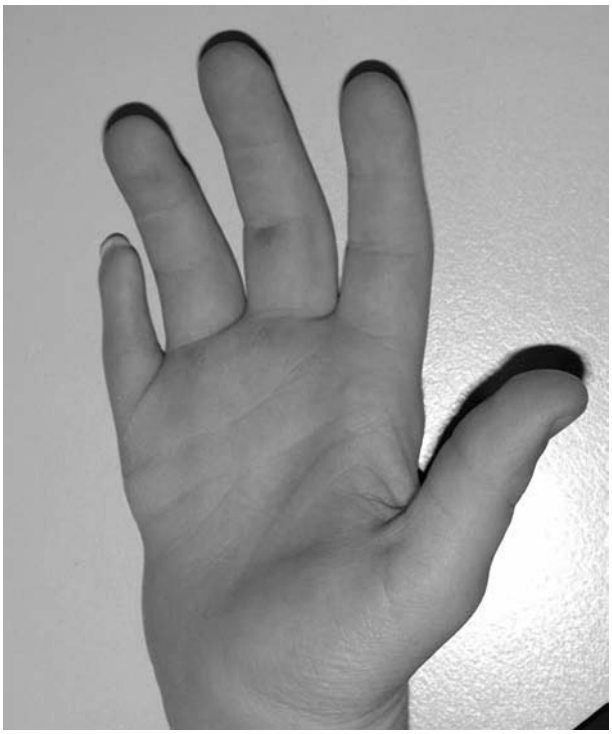
*The rearrangement is probably causative, as a sibling with a similar phenotype has the same aberration.
†Manuscript in preparation.
Group of patients #with mental retardation and additional features, §selected solely on the basis of developmental delay.
¶Fluorescent in situ hybridisation, **multiplex ligation dependent probe amplification, ††total anomalous pulmonary venous return.
No/?: one of the parents also has the aberration; however, imprinting, variable expression and low penetrance have not been excluded; TAPVR, total anomalous pulmonary venous return.
Cases 1–6 are described in more detail in the text.

Figure 1. Results of case 3.



The plots correspond to the MAPH results showing (A) a deletion of the DRG2 gene, two normal copies of COPS3A (RAI1 not present), and the MLPA results; and (B) a deletion of RAI1, a deletion of DRG2, and a normal ratio of COPS3A. (C) The additional FISH analysis using the LSI-SMS probe specific for the Smith Magenis chromosomal region shows a normal signal on the short arm of only one copy of chromosome 17. [See appendix: colour figures.]

Figure 2. The right hand of case 4 showing a short, inflexible fifth digit with a ram's horn shaped nail and hypotrophy of the hypothenar muscles.

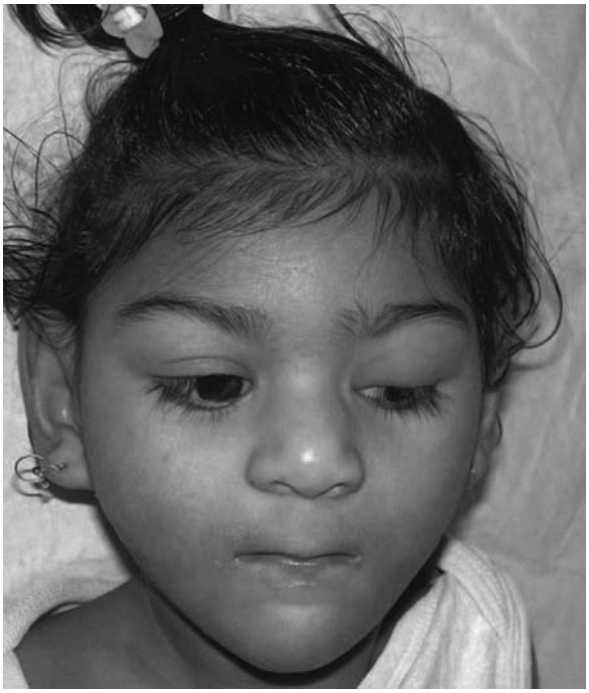


(fig 1a). Additional MLPA testing showed that the *RAI1* gene (LocusLink: 10743) was also deleted in this patient (fig 1b), and FISH analysis (probe LSI-SMS) verified the deletion of part of chromosome band 17p11.2 (fig 1c). Recently, three dominant frameshift mutations in *RAI1* have been identified in three patients with phenotypic characteristics of SMS but no cytogenetically detectable deletion of chromosome band 17p11.2.[18] The authors argue that mutations in *RAI1* are responsible for most of the characteristic features of SMS and that further variation is caused by hemizygosity of the other genes in the chromosome region.

Case 4

This male patient showed at the age of 12 years a mild learning disability, a low voice, a disproportionally short stature (height –2 SD, span –3 SD for height, sitting height –0.5 SD, head circumference –2 SD), limited elbow extension, a permanently extend-

Figure 3. Facial dysmorphism of case 6.



Note the microcephaly, ptosis of the left eye, flat philtrum, and thin upper lip. [See appendix: colour figures.]

ed, inflexible fifth digit of both hands with a ram's horn shaped nail and hypotrophy of the hypothenar muscles (fig 2), and a short broad great toe on both feet. The hand *x* ray revealed short metacarpals I and V, short distal phalange V, and a delay of bone maturation. In this patient, a *de novo* deletion of 4q34.1 was detected and confirmed by FISH (probe RP11-475B2). Analysis with a more distally located MAPH probe at chromosome band 4q35.1 showed that this latter region was still present, indicating an interstitial rearrangement. Additional FISH experiments using different BAC probes limited the deletion to a maximum of 3 Mb (data not shown).

Patients with an interstitial 4q deletion have been described with a range of features, depending on the proximal and distal breakpoints of the deletion.[19] As it is known that fifth finger anomalies and short stature are found in patients with an interstitial deletion of 4q including 4q34,[20] as well as in patients with a terminal deletion of 4q, it is possible that the genes responsible for these features are located within this region.

Case 5

This mildly retarded man, with a *de novo* duplication within chromosome band 20p12.2, containing the *Jagged1* gene (*JAG1*; LocusLin: 182), died at the age of 60 years from multiple myeloma. He had been institutionalised for over 40 years in a psychiatric hospital because of aggressive behaviour, and was diagnosed as schizophrenic. To the best of our knowledge there has been only one previous report[21] of a duplication of 20p11.21–p11.23, in four members of a family with clinical signs of Alagille syndrome. As our patient is not available for further investigation, it remains unclear whether he had such features.

Case 6

After 41 weeks of gestation, this child was born with a birth weight of 1995 g (≤2.5 SD) and a head circumference of 28.5 cm (≤2.5 SD). At the age of 25 months, her psychomotor development was severely delayed and she suffered from epilepsy. Physical examination showed growth retardation (length ≤2 SD; weight –6 SD), microcephaly (head circumference –6 SD), hypertonicity, dystonic movements, facial dysmorphisms (ptosis of the left eye, flat philtrum, thin upper lip; fig 3) ear pits, café au lait spots, and absence of the labia minora. Further investigation revealed corpus callosum hypoplasia and deformed gyri, the presence of only one kidney and mildly increased urinary glutaric acid.

Using the microdeletion probe set, a duplication of 22q11.2 was detected by MAPH, and FISH analysis in interphase nuclei confirmed this finding (LSI TUPLE1). The patient's mother did not carry the duplication, and the father was unavailable for testing. We plan to use polymorphic markers to determine the parental origin of the aberrant chromosome 22.

Discussion

Using MAPH analysis, we performed a high resolution duplication/deletion screening of 188 patients with a developmental delay; 162 loci per patient were tested, amounting to over 30 000 typings. The MAPH probes designed for this study can be broadly divided into two groups: (a) subtelomeric and pericentromeric probes (n = 48) and (b) interstitial probes (n = 114), containing sequences located in regions previously found to be rearranged in mentally retarded individuals, and genes randomly spaced through out the genome.

We detected 4.3% (8/184) subtelomeric/pericentromeric rearrangements (six de-

letions, one duplication, and one subtelomeric deletion/duplication in one patient), using 48 MAPH probes. A subdivision of subtelomeric aberrations over our two study populations agrees with the findings of Knight *et al.*[22] and Yasseen *et al.*[23] The percentage of subtelomeric mutations detected was higher in a group of MR patients with additional malformations (7.7%) than in a group selected on the basis of developmental delay only (2.5%). This supports the suggestion of De Vries *et al.* that pre-selection of patients for subtelomeric screening is worthwhile. However, pre-selection of these patients for subtelomeric rearrangements is difficult, as only two clinical features (perinatal onset growth retardation and a positive family history) differed significantly between patients with subtelomeric aneusomies and patients with idiopathic MR.[24] Our overall percentage is similar to that reported in a recent paper that summarised all previous subtelomeric publications.[7] A total of 131 subtelomeric imbalances were found using several different methods among 2582 MR patients, resulting in an overall frequency of 5.1%. A review of the corresponding clinical aspects of these subtelomeric rearrangements has been published recently.[25] After re-examining the karyogram of our patients at a banding resolution of 500–550 bands, it showed that five MAPH detected subtelomeric imbalances were not cytogenetically visible, despite the knowledge of a copy number change present. This means that the percentage of "true" submicroscopic subtelomeric/ pericentromeric findings is ~3% (5/184) in this study.

Previous reports by Sismani *et al.*[26] and Hollox *et al.*[11] had already shown the ability of MAPH to detect subtelomeric copy number changes. Hollox *et al.* found a copy number change in 5 of 37 male patients (13.5%) who had been referred for fragile X screening. The higher percentage of mutations found by this group may be due to differences in selection criteria for fragile X screening.

We also screened the subtelomeric/pericentromeric regions in eight newborns suffering multiple congenital abnormalities (MCA). Among these patients, one deletion of the subtelomeric region of chromosome 15 was detected and subsequently confirmed by FISH (data not shown).[27] To determine whether it is worthwhile to test this group for submicroscopic mutations, more newborns with MCA should be examined. The ease and relatively low cost of the MAPH technique means that such analysis is feasible. Moreover, new techniques such as MAPH/MLPA and array CGH provide the possibility of genetic diagnosis at a younger age. As the suggestive phenotype for some microdeletion syndromes emerge only later in life, this diagnosis would be very important for providing appropriate healthcare.

In addition to the reports published by Sismani *et al.*[25] and Hollox *et al.*,[11] we also examined interstitially localised genes, including genes involved in several microdele-

tion syndromes, genes on chromosome 22 (as this was the first chromosome to be completely sequenced), and genes that are spread throughout the genome and might be involved in cognitive development. Recently, Bailey *et al.*[3] argued that regions between highly similar duplications (low copy repeats) are prone to recombination and consequently, copy number changes occur at a higher frequency in these regions compared with other loci in the genome. Several of the areas described were also tested in this study, mostly corresponding to chromosomal regions involved in microdeletion syndromes. In total, seven interstitial deletions and duplications were detected, of which five were diagnosed in three different regions known to be involved in the microdeletion syndromes and flanked by segmental duplications. Three of these interstitial rearrangements detected include duplications of regions that are usually deleted (the chromosome regions of Smith Magenis (17p11.2), DiGeorge (22q11.2), and Alagille syndromes (20p12.2)). This observation supports the theory that the regions between low copy repeats can both be deleted and duplicated, and implies that the number of patients suffering from a microduplication syndrome is currently probably underestimated. The phenotype (if any) of a microduplication syndrome might, however, be less severe, and under standard diagnostic conditions, the detection of duplications is more problematical. It should be noted that in the second study group, the cases with a distinctive phenotype for a specific microdeletion syndrome were not included.

As has been the case during the development of every new technique, the genomic variations detected can be divided into the following subclasses: (a) genetic changes that are clearly pathogenic, (b) rearrangements that may or may not be causal to the patient's problem, and (c) polymorphic changes. In some cases, extensive clinical studies will be needed to determine to which category a newly detected aberration belongs. In two of our cases, we could detect the rearrangement in one of the parents (the duplication of chromosome band 14q11.2 and the deletion of 7ptel on chromosome band 7p22.3). One explanation is that these imbalances are polymorphic, and that the phenotype of the patient is not related to the copy number change. However, other explanations are possible: (a) the affected region is imprinted, and the parental origin of this region is critical in causing the deleterious phenotype;[28] (b) allelic variation in the expression of the genes may influence the phenotype;[29] and (c) low penetrance of the rearrangement–that is, a genetic defect does not always lead to a phenotypic effect. The detection of such rearrangements will increase as high resolution techniques are applied, and this will pose new problems for genetic counselling. Therefore, it is important to map these familial imbalances in further detail to allow a genotype–phenotype correlation in larger populations of individuals with the same copy number

change. In this way, the understanding of any clinical consequence of such a rearrangement should be improved.

Based on previous publications, seven rearrangements found in this study were considered to be pathogenic (table 1). In the remaining cases, the data available in literature were insufficient to support a conclusion that the aneusomy detected is related to the phenotype of the patient. It should be noted that the fact that a rearrangement is *de novo* is not in itself proof that it is causally related to the deleterious phenotype.

Several different methologies have been described to identify changes using MAPH and MLPA. These include visual comparison of traces from controls and patients,[30] the setting of arbitrary thresholds,[26] and bivariate analysis.[11] We observed that the standard deviations for each probe varied slightly between hybridisations, and could be normalised only within a single hybridisation. The standard deviation of "normal" probes within each patient was calculated, with 3 times this figure defining the threshold for a potential rearrangement, thus minimising the effect of any genuine copy number changes on the analysis. As false negative results are, by definition, mutations that were not detected, it is difficult to determine the percentage. To gain an estimate as to the actual false negative rate, we looked at a number of samples where a mutation was previously known. We tested 30 samples that had aberrations at loci corresponding to 39 of the probes used. The appropriate copy number changes were detected in all cases. Using the LaPlace formula $p = (x+1)/(n+2)$ to provide a false negative rate from our data yields an expected value of ~2.5%. This figure suggests that the true false negative rate would be, at least for the 39 probes examined, comparable to the 2% theoretically predicted by Hollox *et al.*[11] Of course, it would be desirable to test all the probes on known mutations in the future.

The number of interstitial aneusomies found in this report strengthens the arguments for genomewide screening for copy number changes in developmentally delayed patients. In most clinical laboratories, deletions and duplications are detected by FISH. This usually focuses on only one region per hybridisation, and is therefore relatively slow and expensive. Several new technologies have emerged that facilitate large scale and genomewide screening of deletion and duplication mutations. For genomewide screening, array CGH currently seems to be the most attractive, with recent publications describing screening with approximately 2000 BAC-PAC clones at an average resolution of 1.5 Mb.[31 32] This is impressive, but inherently means that 90% of the genome is not screened. In addition, probes in array CGH are 100–200 kb BAC clones, often covering more than one gene and thus able to pick up large multi-gene deletions/duplications only–that is, those >100 kb, while it is probable that a significant

proportion of deletion/duplication mutations are smaller than this. In contrast, it is possible to detect rearrangements of only 100 bp using MAPH and MLPA technology. By applying a high resolution method, however, the percentage of the genome that can be screened using the same number of probes will be much less compared with array CGH. Using MAPH/MLPA, it is not possible to screen the whole genome for copy number changes at this moment, unless a very large number of probes are included. For this reason, a different approach is required. We consider array CGH to be an excellent tool for finding large regions in the genome where genes involved in particular diseases reside. As soon as these areas have been identified, targeted and much cheaper assays can be designed, zooming in on these regions only. For these reasons, we believe that gene specific screening is ultimately more attractive. With that in mind, MAPH/MLPA have an important role in such analyses, as they are able to pick up both large and small deletions/duplications.

Acknowledgements

We would like to thank all physicians of the Center for Human and Clinical Genetics Leiden for selecting patients and gathering blood samples, Dr J P Schouten (MRC-Holland, Amsterdam) for providing the MLPA probes, the Leiden Genome Technology Center for technical assistance, Dr P Eilers and E Sterrenburg for giving statistical advice, Dr C Rosenberg for critical reading of the manuscript, Dr E Peeters for her efforts, and the patients and families for their cooperation. M Kriek is funded by ZON-MW (AGIKO fellowship 940–37–032).

References

1 Mefford HC, Trask BJ. The complex structure and dynamic evolution of human subtelomeres. *Nat Rev Genet* 2002;**3**:91–102.

2 Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 1998:417–22.

3 Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE. Recent segmental duplications in the human genome. *Science* 2002;**297**:1003–7.

4 Lejeune J, Turpin MR, Gautier M. Etude des chromosomes de neuf enfants mongoliens. *Contes Rendus Acad Sci* 1959;**248**:1721–2.

5 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**:132–40.

6 Biesecker LG. The end of the beginning of chromosome ends. *Am J Med Genet* 2002;**107**:263–6.

7 Flint J, Knight S. The use of telomere probes to investigate submicroscopic rearrangements associated with mental retardation. *Curr Opin Genet Dev* 2003;**13**:310–16.

8 Armour JA, Sismani C, Patsalis PC, Cross G. Measurement of locus copy number by hybridisation with amplifiable probes. *Nucl Acids Res* 2000;**28**:605–9.

9 Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002;**30**:e57.

10 Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998;**20**:207–11.

11 Hollox EJ, Atia T, Cross G, Parkin T, Armour JA. High throughput screening of human subtelomeric DNA for copy number changes using multiplex amplifiable probe hybridisation (MAPH). *J Med Genet* 2002;**39**:790–5.

12 White S, Kalf M, Liu Q, Villerius M, Engelsma D, Kriek M, Vollebregt E, Bakker B, van Ommen GJ, Breuning MH, Den Dunnen JT. Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet* 2002;**71**:365–74.

13 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**:132–40.

14 Dauwerse JG, Jumelet EA, Wessels JW, Saris JJ, Hagemeijer A, Beverstock GC, van Ommen GJ, Breuning MH. Extensive cross-homology between the long and the short arm of chromosome 16 may explain leukemic inversions and translocations. *Blood* 1992;**79**:1299–304.

15 Cody JD, Ghidoni PD, DuPont BR, Hale DE, Hilsenbeck SG, Stratton RF, Hoffman DS, Muller S, Schaub RL, Leach RJ, Kaye CI. Congenital anomalies and anthropometry of 42 individuals with deletions of chromosome 18q. *Am J Med Genet* 1999;**85**:455–62.

16 The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. The European Polycystic Kidney Disease Consortium. *Cell* 1994;77:881–94.

17 Wilkie AO, Buckle VJ, Harris PC, Lamb J, Barton NJ, Reeders ST, Lindenbaum RH, Nicholls RD, Barrow M, Bethlenfalvay NC. Clinical features and molecular analysis of the alpha thalassemia/mental retardation syndromes. I. Cases due to deletions involving chromosome band 16p13.3. *Am J Hum Genet* 1990;**46**:1112–26.

18 Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH. Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat Genet* 2003;**33**:466–8.

19 Lin AE, Garver KL, Diggans G, Clemens M, Wenger SL, Steele MW, Jones MC, Israel J. Interstitial and terminal deletions of the long arm of chromosome4: further delineation of phenotypes. *Am J Med Genet* 1988;**31**:533–48.

20 Tsai CH, Van Dyke DL, Feldman GL. Child with velocardiofacial syndrome and del (4)(q34.2): another critical region associated with a velocardiofacial syndrome-like phenotype *Am J Med Genet* 1999;**82**:336–9.

21 Moog U, Engelen J, Albrechts J, Hoorntje T, Hendrikse F, Schrander-Stumpel C. Alagille syndrome in a family with duplication 20p11. *Clin Dysmorphol* 1996;5:279–88.

22 Knight SJ, Regan R, Nicod A, Horsley SW, Kearney L, Homfray T, Winter RM, Bolton P, Flint J. Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* 1999;**354**:1676–81.

23 Yasseen AA, Al Musawi TA. Cytogenetics study in severely mentally retarded patients. *Saudi Med J* 2001;**22**:444–9.

24 De Vries BB, White SM, Knight SJ, Regan R, Homfray T, Young ID, Super M, McKeown C, Splitt M, Quarrell OW, Trainer AH, Niermeijer MF, Malcolm S, Flint J, Hurst JA, Winter RM. Clinical studies on submicroscopic subtelomeric rearrangements: a checklist. *J Med Genet* 2001;**38**:145–50.

25 De Vries BB, Winter R, Schinzel A, Ravenswaaij-Arts C. Telomeres: a diagnosis at the end of the chromosomes. *J Med Genet* 2003;**40**:385–98.

26 Sismani C, Armour JA, Flint J, Girgalli C, Regan R, Patsalis PC. Screening for subtelomeric chromosome abnormalities in children with idiopathic mental retardation using multiprobe telomeric FISH and the new MAPH telomeric assay. *Eur J Hum Genet* 2001;**9**:527–32.

27 Roback EW, Barakat AJ, Dev VG, Mbikay M, Chretien M, Butler MG. An infant with deletion of the distal long arm of chromosome 15 (q26.1–qter) and loss of insulin-like growth factor 1 receptor gene. *Am J Med Genet* 1991;**38**:74–9.

28 Pollack JR, Iyer VR. Characterizing the physical genome. *Nat Genet* 2002;**32**(Suppl):515–21.

29 Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. Allelic variation in human gene expression. *Science* 2002;**297**:1143.

30 Duponchel C, Di Rocco C, Cicardi M, Tosi M. Rapid detection by fluorescent multiplex PCR of exon deletions and duplications in the C1 inhibitor gene of hereditary angioedema patients. *Hum Mutat* 2001;**17**:61–70.

31 Veltman JA, Fridlyand J, Pejavar S, Olshen AB, Korkola JE, DeVries S, Carroll P, Kuo WL, Pinkel D, Albertson D, Cordon-Cardo C, Jain AN, Waldman FM. Array-based comparative genomic hybridization for genome-wide screening of DNA copy number in bladder tumors. *Cancer Res* 2003;**63**:2872–80.

32 Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, Law S, Myambo K, Palmer J, Ylstra B, Yue JP, Gray JW, Jain AN, Pinkel D, Albertson DG. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001;**29**:263–4.

33 Brkanac Z, Cody JD, Leach RJ, DuPont BR. Identification of cryptic rearrangements in patients with 18q-deletion syndrome. *Am J Hum Genet* 1998;**62**:1500–6.

34 Zurcher VL, Golden WL, Zinn AB. Distal deletion of the short arm of chromosome 6. *Am J Med Genet* 1990;**35**:261–5.

35 Palmer CG, Bader P, Slovak ML, Comings DE, Pettenati MJ. Partial deletion of chromosome 6p: delineation of the syndrome. *Am J Med Genet* 1991;**39**:155–60.

36 Chotai KA, Brueton LA, van Herwerden L, Garrett C, Hinkel GK, Schinzel A, Mueller RF, Speleman F, Winter RM. Six cases of 7p deletion: clinical, cytogenetic, and molecular studies. *Am J Med Genet* 1994;**51**:270–6.

37 Heilstedt HA, Shapira SK, Gregg AR, Shaffer LG. Molecular and clinical characterization of a patient with duplication of 1p36.3 and metopic synostosis. *Clin Genet* 1999;**56**:123–8.

38 Heilstedt HA, Ballif BC, Howard LA, Lewis RA, Stal S, Kashork CD, Bacino CA, Shapira SK, Shaffer LG. Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. *Am J Hum Genet* 2003;**72**:1200–12.

39 Potocki L, Chen KS, Koeuth T, Killian J, Iannaccone ST, Shapira SK, Kashork CD, Spikes AS, Shaffer LG, Lupski JR. DNA rearrangements on both homologues of chromosome 17 in a mildly delayed individual with a family history of autosomal dominant carpal tunnel syndrome. *Am J Hum Genet* 1999;**64**:471–8.

40 Vlangos CN, Yim DK, Elsea SH. Refinement of the Smith-Magenis syndrome critical region to approximately 950 kb and assessment of 17p11.2 deletions. Are all deletions created equally? *Mol Genet Metab* 2003;**79**:134–41.

41 Edelmann L, Pandita RK, Spiteri E, Funke B, Goldberg R, Palanisamy N, Chaganti RS, Magenis E, Shprintzen RJ, Morrow BE. A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum Mol Genet* 1999;**8**:1157–67.

42 Ensenauer RE, Adeyinka A, Flynn HC, Michels VV, Lindor NM, Dawson DB, Thorland EC, Lorentz CP, Goldstein JL, McDonald MT, Smith WE, Simon-Fayard E, Alexander AA, Kulharya AS, Ketterling RP, Clark RD, Jalal SM. Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet* 2003;**73**:1027.

43 Bartsch O, Nemeckova M, Kocarek E, Wagner A, Puchmajerova A, Poppe M, Ounap K, Goetz P. DiGeorge/velocardiofacial syndrome: FISH studies of chromosomes 22q11 and 10p14, and clinical reports on the proximal 22q11 deletion. *Am J Med Genet* 2003;**117A**:1–5.

Chapter II-2

Copy number variation in regions flanked (or unflanked) by duplicons among patients with developmental delay and/or congenital malformations; detection of reciprocal and partial Williams-Beuren duplications

Marjolein Kriek[1], Stefan J White[1], Karoly Szuhai[2], Jeroen Knijnenburg[2], Gert-Jan B van Ommen[1], Johan T den Dunnen[1] and Martijn H Breuning[1]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, The Netherlands;
[2]Department of Molecular Cell Biology, Leiden University Medical Center, The Netherlands

Eur J Hum Genet. 2006 Feb;14(2):180-9

Summary

Duplicons, that is, DNA sequences with minimum length 10 kb and a high sequence similarity, are known to cause unequal homologous recombination, leading to deletions and the reciprocal duplications. In this study, we designed a Multiplex Amplifiable Probe Hybridisation (MAPH) assay containing 63 exon-specific single-copy sequences from within a selection of the 169 regions flanked by duplicons that were identified, at a first pass, in 2001. Subsequently, we determined the frequency of chromosomal rearrangements among patients with developmental delay (DD) and/or congenital malformations (CM). In addition, we tried to identify new regions involved in DD/CM using the same assay. In 105 patients, six imbalances (5.8%) were detected and verified. Three of these were located in microdeletion-related regions, two alterations were polymorphic duplications and the effect of the last alteration is currently unknown. The same study population was tested for rearrangements in regions with no known duplicons nearby, using a set of probes derived from 58 function-selected genes. The latter screening revealed two alterations. As expected, the alteration frequency per unit of DNA is much higher in regions flanked by duplicons (fraction of the genome tested: 5.2%) compared to regions without known duplicons nearby (fraction of the genome tested: 24.5–90.2%). We were able to detect three novel rearrangements, including the previously undescribed reciprocal duplication of the Williams Beuren critical region, a subduplicon alteration within this region and a duplication on chromosome band 16p13.11. Our results support the hypothesis that regions flanked by duplicons are enriched for copy number variations.

Introduction

Many genetic disorders are caused by changes in chromosomal structure. Deletions, duplications, inversions and translocations can all lead to changes in the effective dosage of one or more genes, often with pathological consequences. Large rearrangements affecting at least 5 Mb can be seen cytogenetically, and many disorders have been recognised and characterised based solely on microscopic analysis.[1-4]

It was shown in 1992 that the region duplicated in Charcot-Marie-Tooth (CMT) was flanked by highly similar (>98%) sequences.[5] Unequal crossing over between these duplicons leads both to this duplication and the reciprocal deletion, which was later shown to cause hereditary neuropathy with liability to pressure palsies (HNPP).[6] Duplicons, also known as low copy repeats (LCRs), have since been implicated in many other disorders.[7,8] It has been estimated that 5% of the human genome is composed of such LCRs, which can be present both inter-and intrachromosomally.[9,10]

In 2002, Bailey *et al.*[11] identified 169 unique regions of at least 10 kb in size, between intrachromosomal duplicons with >95% sequence identity. These data were based on the Human Working draft of August 2001. In all, 24 of these regions were already associated with known genetic disorders. It was hypothesised that these 169 regions are likely to undergo rearrangements more frequently compared to interstitial regions outside the defined regions, due to misaligned recombination between the LCRs, creating microdeletions, microduplications and inversions of the segments involved. To assess this in more detail, we have designed a Multiplex Amplifiable Probe Hybridisation (MAPH) probe set containing 30% of these regions, including those related to microdeletion syndromes. In all, 105 unrelated patients with developmental delay (DD) and/or congenital malformations (CM) were tested using these probes. We compared the performance of this probe set with a set of probes located outside the thus far known duplicons. The second purpose of this study was to identify new regions that are frequently altered in DD patients or patients with CM using the duplicon data of 2002.

The assay using sequences flanked by duplicons resulted in the detection of six duplications, of which three were located in regions related to known disorders. Two alterations were detected by screening regions outside known duplicons. These results show that in our study population the genetic variation within duplicon-flanked regions was three times more common compared to the regions outside the duplicons. Among the rearrangements detected was the postulated, but until now unidentified, reciprocal duplication of the Williams Beuren critical region (WBCR) and a smaller subduplicon alteration within this region.

Materials and methods

Patients

The DNA of 99 DD/CM patients and six individuals with CM only (64 males and 41 females) from the Center of Human and Clinical Genetics Leiden (DNA Diagnostic Laboratory) was analysed. Prior to MAPH analysis, all patients showed a normal karyotype and, where tested, had tested negative for Fragile X syndrome. This study cohort does not include any patient presenting with typical microdeletion characteristics. These had been previously diagnosed by the cytogenetics department.

This study was approved by the Institutional Review Board of the Leiden University Medical Center, conforming to Dutch law. All subjects, or their representatives, gave informed consent for DNA studies.

Multiplex Amplifiable Probe Hybridisation

MAPH was performed as described by White *et al.*[12] Ratios were obtained by dividing the peak height of each probe by the sum of the peak heights of the four nearest probes. The probes with a normalised ratio between 0.75 and 1.25 (log(2) scale –0.42 to +0.32) were considered to be present in two copies. The probes with a ratio outside these thresholds were considered to have a copy number alteration. All samples in which an alteration was found were screened at least in duplicate.

The different probe sets used contained respectively 63 probes from genes flanked by duplicons (see Appendix A) in 51 different regions, including those involved in Smith Magenis (SMS (MIM 182290)), William Beuren (WBS (MIM 194050)), DiGeorge (DGS (MIM 188400)), Cat eye (CES (MIM 115470)), Prader Willi (PWS (MIM 176270)), Angel-man syndrome (AS (MIM 105830)) and 58 probes containing function-selected genes outside the duplicons (Appendix B).

Multiplex Ligation-dependent Probe Amplification

A modified protocol of multiplex ligation-dependent probe amplification (MLPA)[13] was performed as described by White *et al.*[14] In the current study, MLPA was performed to verify alterations obtained by MAPH analysis. The data analysis is identical with that applied for MAPH analysis. The MLPA probes used were derived from the sequences of *RAI1* (GeneID: 10743), *DRG2* (GeneID: 1819), *COPS3* (GeneID: 8533), *ELN* (GeneID: 2006), *CYLN2* (GeneID: 7461), *FKBP6* (GeneID: 8468), *TBL2* (GeneID: 26608), *FZD9* (GeneID: 8326), *GTF2IRD1* (GeneID: 84163), *GTF2I* (GeneID: 2969), *HIP1* (GeneID:3092), *AUTS2* (GeneID:26053), *CALN1* (GeneID: 83698),

NUDE1 (GeneID: 54820), *PYRR1*, defender against cell death 1 (*DAD1*) gene (GeneID: 1603) and the diacylglycerol kinase iota (*DGKI*) gene (GeneID: 9162).

Fluorescence In Situ *Hybridisation*

The FISH experiments were performed following Standard Operating Procedures.[15] An FITC-labeled FISH clone LSI-ELN (Vysis) was used for the Williams critical Region. BAC clones RP11-14N9, RP11-M13, RP11-489O1 and RP11-72I8 were used to determine the extent of the rearrangement on chromosome band 16p13.3.

Array comparative genomic hybridisation

The array comparative genomic hybridisation (array-CGH) procedures were performed as described in Knijnenburg *et al.*[16] using larger genomic insert clones retrieved from the Sanger Center (UK) (1 MB clone set). *In silico* data at the http://www.ensemble.org were used to determine the size of the duplications.

Results

Considering that duplicon-flanked regions might be preferentially involved in copy number variation, we based our MAPH probe set to detect new regions involved in DD/CM on a gene-enriched selection from the 169 regions published by Bailey *et al.*[11]

The MAPH probes were designed based on autosomal exon-specific single-copy sequence. Regions lacking known genes and/or single-copy sequence (62/169 or 37% of the defined regions) were excluded. Before the actual screening, the probe sets were validated using DNA samples derived from 50 anonymous healthy controls. Among those, we detected a pancreatic polypeptide receptor 1 (*PPYR1*) gene duplication that was verified using MLPA analysis. Probes showing inconsistent copy number variation within an individual (duplicate testing) were excluded ($n = 9$). The validated probe sets, targeting 63 unique sequences in 51 different regions (see Appendix A), were tested among a total of 105 unrelated patients (64 males, 41 females), including 99 developmentally delayed (DD) patients (25 mild DD; 74 severe DD) and six individuals with CM.

Screening these 105 patients revealed six imbalances (5.8%), all duplications (Table 1). All rearrangements were verified using MLPA, array-CGH or FISH. Three of the rearrangements were located in areas known to be involved in microdeletion syndromes, including two duplications within the WBCR on chromosome band 7q11.23

Table 1. Alterations in regions flanked by duplicons.

Case	Alteration	Chrom. Band	Gene(s) involved	Size (Mb)	*de novo*	Confirmed by
1	Duplication	7q11.23	CYLN2, ELN, FZD9, FKBP6, TBL2	1.4–1.7	No, present in father	MLPA/FISH
2	Duplication	7q11.23	FKBP6	0.3–0.4	No, present in mother	MLPA
3	Duplication	17p11.2	RAI1, DRG2, COPS3	min. 3.5[a]	Yes	MLPA/FISH/array-CGH
4	Duplication	16p13.11	NUDE1, MYH11	0.8–2.4	Yes	MLPA/FISH/array-CGH
5	Duplication	10q11.22	PPYR1	0.5–2.3	No, present in father	MLPA/array-CGH
6	Duplication	10q11.22	PPYR1	max. 1.4	Unknown[b]	MLPA/array-CGH

Summary of results obtained by screening 105 DD/CM patients using 51 unique regions flanked by duplicons. The sizes of the different alterations were determined based on results of both MAPH/MLPA and array-CGH.

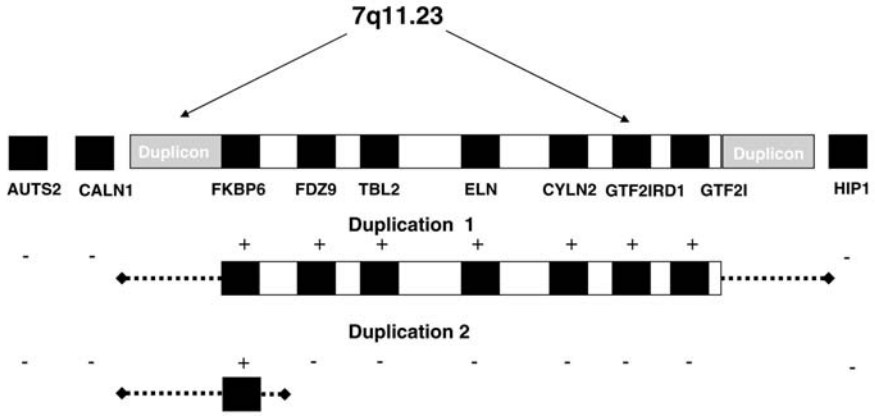
a) As the regions near the centromere of chromosome 17 are not covered by array-CGH, the centromeric breakpoint of this duplication remains unknown.

b) The mother of case 6 did not carry the duplication. The father was not available for testing.

(see case reports), and a *de novo* duplication of the Smith Magenis Critical Region (SMCR) on chromosome band 17p11.2. The two 7q11.23 duplications, detected in two unrelated patients, differed in length, as one was found using four MAPH probes (containing sequences derived from the *CYLN-2*, *ELN*, *FKBP6* and *TBL2* genes) and the other with only one of these, the FKBP6 gene (Figure 1). Additional array-CGH analysis did not detect this alteration. The exact size of the duplication is difficult to define as the BACs flanking this region (RP11-450O3, RP4-771P4) partly colocalise with segmental duplicons in this region. Additional MLPA was performed using sequences of the *GTF2I* and *GTF2IRD1* genes within the WBCR and *HIP1*, *CALN1* and *AUTS2* genes localised just outside the telomeric and centromeric sides of the segmental duplicon, respectively. This assay revealed that this duplication is the reciprocal duplication of the deletion causing Williams–Beuren syndrome.

To fine map the other duplications (case 2), additional MLPA probes were designed. Exon 4 and exon 8 (the last exon) of the *FKBP6* gene were shown to be duplicated. We were unable to test the first three exons of this gene, as they contain large repetitive sequences. The probe derived from the adjacent *FZD9* gene showed no alteration. Testing the parents of the patients showed that in each case the duplication was present in one of the parents (data not shown). There appeared to be no parent of origin effect, as the large alteration was found in the patient's father, and the small alteration in the mother of the other patient.

Figure 1. The duplications within 7q11.23 (WBCR).



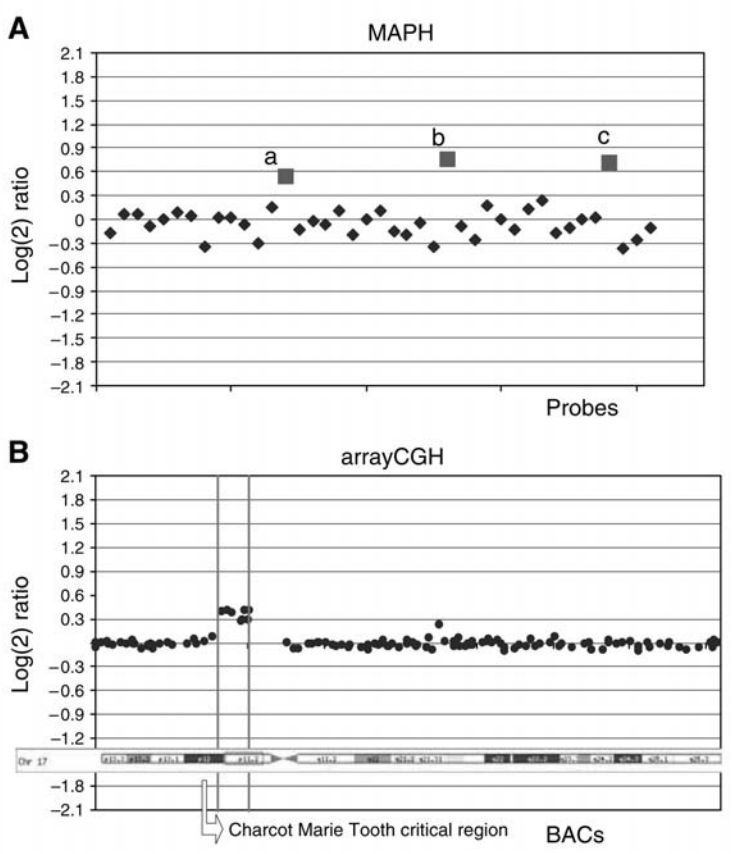
The figure shows the length of the two duplications in the WBCR, detected in unrelated patients. Duplication 1 encompasses the whole critical area flanked by two large duplicons, whereas the other duplication involves only (a part of) the *FKBP6* gene. The diamonds represent the maximum size of both duplications. The *AUTS2*, *CALN1* and *HIP1* genes localised just outside the duplicons were not altered.

The duplication of the SMCR (case 3) was detected using three probes corresponding to the *RAI1*, *DRG2* and *COPS3* gene. Array-CGH testing was performed to determine the length of the duplication on chromosome 17 (Table 1). This analysis excluded a duplication of chromosome band 17p12, which causes CMT disease (Figure 2).

Chromosome 16 contains many repeats, limiting the application of additional FISH analysis. Thus, it was not possible to determine the precise breakpoints of the imbalance in case 4, a *de novo* duplication of the *NUDE1* gene on the short arm of chromosome 16p13.11. Two BACs (RP11-489O1, CTD-2504F3) overlapping the *NUDE1* region were found amplified using array-CGH, indicating that the size of the duplication is between 0.8 and 2.4 Mb. We note that the dosage of the *MYH11* gene (Locus Link: 4629) must also be doubled as this gene is transcribed from the reverse strand of the *NUDE1* gene.

In two unrelated patients (cases 5 and 6), a duplication of a probe within the first exon of the *PPYR1* gene on chromosome 10 was identified and subsequently verified using MLPA. Using array-CGH analysis, a nonoverlapping BAC (RP11-292F22) localised 0.5 Mb telomeric from the *PPYR1* gene showed a duplication in only one of the patients, indicating a difference in the size of the regions duplicated. We were able to test both parents of the patient with the largest rearrangement (case 5); the father

Figure 2. Results obtained in case 3.



Results of the MAPH and array-CGH analysis revealing a duplication of the SMCR. (A) Log(2) ratio of MAPH probes showing a duplication of (a) the *RAI1* gene, (b) the *DRG2* gene and (c) the *COPS3* gene. The remaining probes contained sequences localised on different chromosomes. The probes with a normalised ratio between –0.42 and +0.32 (log(2) scale) were considered to be present in two copies. The probes are ordered by probe length, not on their position on the genome. (B) Array-CGH testing showed that chromosome band 17p12 is not duplicated, excluding CMT syndrome (white arrow). The BACs showing amplification included RP11 –219A15, RP11–524F11, RP11 –189D22, RP1– 162E17, CTB –1187M2, RP11–78O7, RP5–836L9 and RP11–121A13. The distal breakpoint matches the common deletion breakpoint of SMS.[18] The proximal breakpoint is unknown, as the region near the centromere is not covered by BACs.

carried the same duplication. The mother of the other patient did not show the duplication, the father was not available for testing.

To determine whether the number of alterations obtained is significantly higher compared to copy number changes of regions outside the duplicons described in 2001, we have tested the same study population for genomic variation in a set of probes from regions not known to be flanked by duplicons. These probes were targeting function-selected genes, such as genes involved in transcription, neuronal and brain maturity, with a potential function in mental development (Appendix B). This MAPH analysis comprised 58 validated probes (Appendix B) and resulted in the detection of two genetic imbalances (1.9%), including a duplication of the *DGKi* gene on chromosome band 7q33 and a deletion of the *DAD1* gene on chromosome band 14q11. Both alterations were verified by MLPA analysis. We were not able to test the parents of these patients. Despite their predicted function, these genes have not previously been causally linked to DD.

Case reports

Case 1

This male patient was born after an uneventful pregnancy. In the perinatal period, he was diagnosed with trigonocephalic synostosis of the metopic ridge. At the age of 1 year, he was examined by a clinical geneticist. He did not show any DD nor obvious dysmorphic features. Except for a mild aberrant shape of his skull (status after reconstruction), no CM were present.

The family history of this patient included, in the father with a complete cutaneous III–IV syndactyly of the hand, a II–III syndactyly of the feet, and a carcinoma *in situ* of the testis that was diagnosed after infertility screening. The family members of both the father's mother and father's father showed syndactyly. Additional MAPH analysis showed a duplication of the WBCR present in the patient as well as in the father. The parents of the patient's father did not carry the duplication. The parenthood of the father and his parents was proven using marker studies.

Case 2

In addition to synostosis of both the sutura lamboidea and the sutura coronalis, this 4-year-old male patient with a normal mental development showed facial asymmetry, a severe heart malformation including two ventricular septum defects and a (sub)valvular pulmonal stenosis and a finger-like thumb. Except for craniosynostosis, these features are related to hemifacial microsomia.

The family history does not include individuals with dysmorphic features nor CM. Additional investigation showed a normal karyotype. MAPH analysis showed a duplication of a part of the *FKBP6* gene that was also present in the unaffected mother and the unaffected maternal grandmother.

Discussion

In this study, we have assessed the frequency of chromosomal rearrangements in DD and/or CM patients. The fraction of the genome that was localised between the defined duplicons (as of 2001) and tested by at least one MAPH probe was 5.2% (see Appendix A). Within these regions, six alterations were detected. The fraction of the genome that was flanked by duplicons and not tested in this study was 4.6%, indicating that the majority of the genome fraction flanked by duplicons has been tested in this study. The total fraction of the genome that was flanked by duplicons identified at a first pass in 2001 is thus 9.8%. This percentage corresponds closely with the ~328 Mb of sequence calculated by Bailey *et al*.

The fraction of the genome unflanked by duplicons (defined in 2001) is 90.2%. However, we have only tested 58 sequences (probes) localised outside the duplicons. We would argue that this number is not representative for 90.2% of the genome. Based on the calculation shown in Appendix B, the fraction of the non-duplicon regions tested was at least 24.5%. The real percentage tested is higher, as sequences located at the chromosome ends could not be included. In short, the fraction of the genome localised outside the duplicons and tested ranges between 24.5 and 90.2%. Two alterations were found within these regions. While the sample sizes are small, the aberration frequency per unit (= percentage of the total genome) of DNA in regions flanked by duplicons was higher compared to the regions outside the duplicons, indicating that the regions between the duplicons are indeed enriched for dosage alterations. This supports the hypothesis of Bailey *et al*. that the regions within duplicons are more likely to undergo genomic alterations.

Retrospectively, we have checked all 58 genes localised outside the duplicons, as identified in 2001, using the most recent assembly of the Human Working Draft (May 2004). It appeared that 76% of these regions were still unflanked by intrachromosomal duplicons, including the regions containing *DGKi* and *DAD1* genes.

Several factors will lead to an underestimation of the true number of alterations occurring between duplicons, and some of these may also explain why we did not find any deletions. First, the regions lacking single-copy sequences were excluded in this

study. It is reasonable to assume that these regions are more likely to undergo rearrangements based on their repetitive sequence content. These were not included, as the MAPH assay was based on copy number alteration of single-copy sequences.

Second, haplo-insufficiency of certain genes might not be compatible with life, or they may give a deleterious phenotype other than DD/CM. These alterations will not be detected in our study. This holds equally for the function-selected genes. Brewer *et al.*[17] defined several regions that have never been involved in any deletion and those were thought to be potentially haplo-lethal. Of the 57 'Bailey' regions tested, 10 were located within these possible haplo-lethal regions. These regions need to be tested by higher resolution methods, as the analysis of Brewer *et al.* was based on karyotypic abnormalities. Third, a substantial proportion of DD/CM could originate from genetic aberrations other than nonallelic homologous recombination. For example, point mutations will not be detected using MAPH.

Fourth, the number of samples tested is rather small and the set of probes outside the duplicons is not random. In addition, the study cohort is already biased against rearrangements between duplicons, as any cases presenting with typical microdeletion syndrome-related features had already been diagnosed using cytogenetics tools.

Finally, it is possible that a part of the duplicons defined by Bailey *et al.* require additional conditions before the obligate 'repetitive breakpoints events' will occur, resulting in copy number changes. These additional conditions could include a minimum length of 100% homology required for recombination, AT-rich sequences present on both sites of a recombination hotspots,[18] or enrichment of Alu repeats within duplicons.[19] Further analysis needs to be performed to determine whether these conditions are present in the 'Bailey'-defined duplicons.

A more clinical question concerns whether the imbalances found are disease-causing changes or benign polymorphisms. Alterations due to misaligned nonallelic homologous recombination should result in a deletion and a reciprocal duplication. In the majority of reciprocal deletion/duplication disorders, deletions were discovered before the duplication of the regions due to the fact that the techniques applied (usually FISH) were more amenable for deletion detection. To date, several duplications in regions involved in microdeletion syndromes have been identified in addition to the known deletions.[20-23] The phenotype corresponding to the duplication is often milder than that related to the deletion. However, the copy number changes can also be associated with polymorphic variation.[24]

Due to the presence of >320 kb repeat structure on both sides of the Williams syndrome critical region, the existence of a reciprocal duplication of the Williams critical region was predicted,[25,26] however, it has not been reported before. The patient with the

reciprocal duplication of the Williams critical region was diagnosed with craniosynostosis and mild DD. The patient with the smaller duplication showed, in addition to craniosynostosis, multiple CM; however, his psychological development was normal. As the FKBP6 gene is the only gene in common and this gene is restricted to the male germ cells, it is reasonable to assume that the clinical overlap (craniosynostosis) is coincidental.

The clinical consequences of a duplication within the WBCR are currently unknown. The fact that the imbalance is present in unaffected family members does not automatically mean that this is not pathological. Incomplete penetrance or multifactorial influences might cause variability of the phenotype.

It seems reasonable to assume that the *de novo* 17p11.2 duplication is responsible for the clinical features of case 3, as it is known that a duplication of the SMS critical region is associated with clinical features resembling those observed in our patient.[23,27]

The *de novo* duplication of 16p13.11 was seen in a boy with mild DD and learning disability. Since the father had similar learning problems, the significance of the duplication is questionable and this awaits confirmation from other patients. We note, however, that *NUDE1* participates in a pathway that influences the neuronal migration during development of the central nervous system,[28] which makes it an interesting candidate gene in this region.

Sebat *et al.*[29] reported the screening of a total of 20 healthy individuals using the representational oligonucleotide microarray analysis (ROMA) technique. They found 76 unique large-scale copy number polymorphisms. Among those, five probes on chromosome band 10q11.2 encompassing the full length of the *PPYR1* gene were duplicated in one individual. This finding is in agreement with our finding of no less than four copy number changes in this gene, as it was altered in two unrelated patients (cases 5 and 6), one of their parents, as well as in a healthy control sample. In a subsequent study regarding genomic copy number differences in healthy individuals, 255 loci showing large-scale copy number variation (LCVs) were detected using array-CGH analysis.[30] The only probe that overlapped one of the 255 suspected polymorphic clones contained a *PPYR1* gene sequence. This clone (AL390716.27) was amplified in six individuals. Combining these findings in retrospect, it is possible that *PPYR1* undergoes nonpathological or incompletely penetrant copy number variation. Two of the function-selected genes were localised within the suspected polymorphic clones (*RYR3* within clone ACO11938.4; *ERN1* within clone RP1189H15). The probes derived from both genes were not altered in our study population. This may well be due to our modest sample size, since most copy number variations detected by Iafrate *et al.* were present in only one or two (healthy) individuals. This also holds

true for the clones overlapping *RYR3* and *ERN1*. In addition, a duplication seen with a single BAC clone might not encompass the entire clone length.

Recently, Sharp *et al.*[31] also found a difference with regard to duplicons-flanked regions and copy number variation, in agreement with our findings. In addition, 130 potential copy number variation hotspots flanked by duplicons were tested for rearrangements among 47 healthy individuals using a segmental duplicon BAC microarray. A total of 119 regions showed copy number alteration comprising 141 genes, including the *P25*, *P29* and *ADRBK2* genes, also present in our study. In all, 79 of the 130 copy number variation hotspots showed no alteration among this study population. It was suggested that these latter hotspots are excellent candidate regions to be associated with genetic disorders. Our study covers a fraction of these 'hotspots', which have thus been subjected to a first test for copy number alteration in relation to DD or CM. Using MAPH, we were able to identify three previously undescribed rearrangements, two duplications within WBCR and one duplication of chromosome region 16p13.11, of which the clinical relevance is uncertain at this moment. It will indeed be worthwhile to include these regions in further testing.

Acknowledgements

We thank Hans Dauwerse, Kerstin Hansson, Jeroen Nijhuis for the FISH analysis, Yvonne Hilhorst for providing clinical information, Peter de Knijff for parental marker analysis. MK is funded by Zon-Mw (AGIKO fellowship 940-37-032), SW is funded by ZonMw (nr 91204-047).

Note added in proof

While this work was under review, another patient was described (Severe expressive-language delay related to duplication of the Williams-Beuren locus, MJ Somerville *et al*. N Engl J Med 2005; 353:1694–1701, October 20, 2005) with a duplication of the WBS region. We have assessed the phenotype of our patient in the light of the reported clinical features (language deficiency but good spatial abilities). Considering the age of our patient, we could not assess the spatial abilities, but our patient did present with (moderate) language disability.

References

1 Leao JC, Bargman GJ, Neu RL, Kajii T, Gardner LI: New syndrome associated with partial deletion of short arms of chromosome No. 4. Clinical manifestations of hypospadias, beaked nose, abnormal iris, hemangioma of forehead, seizures, and other anomalies. *JAm Med Assoc* 1967; **202**: 434-437.

2 Alfi O, Donnell GN, Crandall BF, Derencsenyi A, Menon R: Deletion of the short arm of chromosome no.9 (46,9p-): a new deletion syndrome. *Ann Genet* 1973; **16**: 17-22.

3 Schinzel A, Auf der MP, Moser H: Partial deletion of long arm of chromosome 11[del(11)(q23)]: Jacobsen syndrome. Two new cases and review of the clinical findings. *J Med Genet* 1977; **14**: 438-444.

4 Greenberg F, Crowder WE, Paschall V, Colon-Linares J, Lubianski B, Ledbetter DH: Familial DiGeorge syndrome and associated partial monosomy of chromosome 22. *Hum Genet* 1984; **65**: 317-319.

5 Lupski JR, Wise CA, Kuwano A *et al*: Gene dosage is a mechanism for Charcot –Marie –Tooth disease type 1A. *Nat Genet* 1992; **1**: 29-33.

6 Chance PF, Abbas N, Lensch MW *et al.*: Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Hum Mol Genet* 1994; **3**: 223-228.

7 Emanuel BS, Shaikh TH: Segmental duplications: an 'expanding' role in genomic instability and disease. *Nat Rev Genet* 2001; **2**: 791-800.

8 Stankiewicz P, Lupski JR: Genome architecture, rearrangements and genomic disorders. *Trends Genet* 2002; **18**: 74-82.

9 Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE: Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res* 2001; 11: 1005-1017.

10 Eichler EE: Recent duplication, domain accretion and the dynamic mutation of the human genome. *Trends Genet* 2001; **17**: 661-669.

11 Bailey JA, Gu Z, Clark RA *et al.*: Recent segmental duplications in the human genome. *Science* 2002; **297**: 1003-1007.

12 White S, Kalf M, Liu Q *et al.*: Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet* 2002; **71**: 365-374.

13 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.

14 White SJ, Vink GR, Kriek M *et al.*: Two-color multiplex ligation-dependent probe amplification: detecting genomic rearrangements in hereditary multiple exostoses. *Hum Mutat* 2004; **24**: 86–92.

15 Dauwerse JG, Jumelet EA, Wessels JW *et al.*: Extensive cross-homology between the long and short arm of chromosome 16 may explain leukemic inversions and translocations. *Blood* 1992; **79**: 1299-1304.

16 Knijnenburg J, Szuhai K, Giltay J *et al.*: Insights from genomic microarrays into structural chromosome rearrangements. *Am J Med Genet A* 2005; **132**: 36-40.

17 Brewer C, Holloway S, Zawalnyski P, Schinzel A, Fitz Patrick D: A chromosomal duplication map of malformations: regions of suspected haplo-and triplolethality – and tolerance of segmental aneuploidy – in humans. *Am J Hum Genet* 1999; **64**: 1702-1708.

18 Shaw CJ, Withers MA, Lupski JR: Uncommon deletions of the Smith–Magenis syndrome region can be recurrent when alternate low-copy repeats act as homologous recombination substrates. *Am J Hum Genet* 2004; **75**: 75-81.

19 Bailey JA, Liu G, Eichler EE: An Alu transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet* 2003; **73**: 823-834.

20 Lupski JR: Charcot–Marie–Tooth disease: lessons in genetic mechanisms. *Mol Med* 1998; **4**:3-11.

21 Shaw-Smith C, Redon R, Rickman L *et al.*: Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J Med Genet* 2004; **41**: 241-248.

22 Ensenauer RE, Adeyinka A, Flynn HC *et al.*: Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet* 2003; **73**: 1027-1040.

23 Potocki L, Chen KS, Park SS *et al.*: Molecular mechanism for duplication 17p11.2 – the homologous recombination reciprocal of the Smith–Magenis microdeletion. *Nat Genet* 2000; **24**: 84-87.

24 Stankiewicz P, Lupski JR: Molecular-evolutionary mechanisms for genomic disorders. *Curr Opin Genet Dev* 2002; **12**: 312–319.

25 Bayes M, Magano LF, Rivera N, Flores R, Perez Jurado LA: Mutational mechanisms of Williams–Beuren syndrome deletions. *Am J Hum Genet* 2003; **73**: 131-151.

26 Peoples R, Franke Y, Wang Y *et al.*: A physical map, including a BAC/PAC clone contig, of the Williams –Beuren syndrome – deletion region at 7q11.23. *Am J Hum Genet* 2000; **66**: 47-68.

27 Shaw CJ, Bi W, Lupski JR: Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. *Am J Hum Genet* 2002; **71**: 1072-1081.

28 Yan X, Li F, Liang Y *et al.*: Human Nudel and NudE as regulators of cytoplasmic dynein in poleward protein transport along the mitotic spindle. *Mol Cell Biol* 2003; **23**: 1239-1250.

29 Sebat J, Lakshmi B, Troge J *et al.*: Large-scale copy number polymorphism in the human genome. *Science* 2004; **305**: 525-528.

30 Iafrate AJ, Feuk L, Rivera MN *et al.*: Detection of large-scale variation in the human genome. *Nat Genet* 2004; **36**: 949-951.

31 Sharp AJ, Locke DP, McGrath SD *et al.*: Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* 2005; 77: 78-88.

Appendix A. Table A1. An overview of 63 genes tested using MAPH analysis among 105 DD/CM patients.

Gene	Chrom. band	Description	GeneID	Location on chromosome	Regions flanked by duplicons	Distance between duplicon
P29	1p35.3	GCIP-interacting protein p29	25949	28755625	28834059–28932575	98 516
PRKAB2	1q21.1	Protein kinase, AMP-activated, beta 2	5565	173584889	171036140–177249592	62 13 452
CAPN2	1q42.11	Calpain 2, large subunit	824	259601508	258169103–259907561	1 738 458
FLJ2204	2q13	Hypothetical protein FLJ22004		117458997	117178230–117635198	456 968
NPHP1	2q13	Nephrocystin	4867	114316563	109925083–116080082	6 154 999
UMPS	3q21.2	Uridine monophosphate synthetase (orotate)	7372	141469386	140479577–143078865	2 599 288
GLUC	4p15.2	Cytosolic beta-glucosidase	2629	25264898	24981138–25427505	446 367
EVC	4p16.2	Ellis van Creveld syndrome protein	2121	6019414	4420622–10868121	6 447 499
P25	5p15.33	Brain-specific protein p25 alpha	11076	1251194	1180464–1306142	125 678
RANBP17	5q35.1	RAN binding protein 17	64901	188–189 Mb	189063686–189213481	149 795
MLN	6p21.31	Motilin	4295	37345065	34908829–40172081	5 263 252
DDC	7p12.2	Dopa decarboxylase (aromatic L-amino acid)	1644	55065653	49777336–61172562	11 395 226
GSBS	7p14.3	G-substrate	10842	34265228	31828578–37665179	5 836 601
JTV1	7p22.1	Multisynthetase complex auxiliary component	7965	6767988	6652376–7723723	1 071 347
TPST1	7q11.21	Tyrosylprotein sulphotransferase 1	8460	68961571	68899125–69000004	100 879
FKBP6[a]	7q11.23	FK506-binding protein 6	8468	70913203	70865853–71592416	726 563
TBL2[a]	7q11.23	Transducin (beta)-like 2	26608	*	70865853–71592416	see above
ELN[a]	7q11.23	Elastin	2006	*	70865853–71592416	see above
CYLN2[a]	7q11.23	Cytoplasmic linker 2	7461	*	70865853–71592416	see above
ARHGEF5	7q35	Rho guanine nucleotide exchange factor 5	7984	156140962	155961558–156151892	190 334
CENTG3	7q36.1	MRIP-1 protein	116988	163305750	161981314–166427628	4 446 314
di-RAS2	9q22.2	GTP-binding RAS-like 2	54769	99459475	99072373–99615625	543 252
PTCH	9q22.32	Patched (Drosophila) homolog	5727	107463432	105649233–108546485	2 897 252
FANCC	9q22.33	Fanconi anaemia, complementation group C	2176	107816739	105649233–108546485	See above
RSU1	10p13	ras suppressor protein 1	6251	17763985	17343558–17953656	610 098
KIAA0187	10q11.21	KIAA0187 gene product	9790	45421390	45331297–47906414	2 575 117
SDF1	10q11.21	Stromal cell-derived factor 1	6387	47064383	45331297–47906414	See above
PPYR1	10q11.22	Pancreatic polypeptide receptor 1	5540	49145372	49021238–54773984	5 752 746
SGPL1	10q22.1	Sphingosine-1-phosphate lyase 1	8879	77278049	75454903–80191770	4 736 867
TACR2	10q22.1	Tachykinin receptor 2	6865	75719214	75454903–80191770	See above
PAPSS2	10q23.31	3' phosphoadenosine 5' phosphosulphate	9060	88057489	86760491–88848332	2 087 841
FLJ22794	11q12.1	Hypothetical protein FLJ22794	63901	65208275	65086253–65169286	83 033
CD5	11q12.2	CD5 antigen (p56–62)	921	67565428	67348654–67606043	257 389
FADD	11q13.3	Fas (TNFRSF6)-associated via death domain	8772	79700968	77157096–82393752	5 236 656
ICEBERG	11q22.3	ICEBERG caspase-1 inhibitor	59082	120243026	120179983–120328323	148 340

Gene	Chrom. band	Description	GeneID	Location on chromosome	Regions flanked by duplicons	Distance between duplicon
HNT	11q25	Neurotrimin precursor	50863	151230535	151288128–151483616	195 488
CLECSF12	12p13.2	C-type lectin domain family c7, member Ca dep.	64581	10905383	10753561–11533368	779 807
CNTN1	12q12	Contactin 1	1272	45800931	45700335–47233112	1 532 777
DKFZp434B0417	12q12	Hypothetical protein DKFZp434B0417		46680635	45700335–47233112	see above
TMEM5	12q14.2	Transmembrane protein 5	10329	72955978	72479865–73099895	620 030
CKAP2	13q14.3	Cytoskeleton-associated protein 2	26586	52929838	52126921–53081328	954 407
NDN[b]	15q11.2	Necdin	4692	19787505*	17304292–19469943	2 165 651
UBE3A[c]	15q12	Ubiquitin protein ligase E3A	7337	21515963*	20279911–20507618	227 707
LTK	15q15.1	Tyrosine kinase	4058	38501171	35955283–41645045	5 689 762
NMB	15q25.2	Neuromedin B	4828	88362943	85255238–88736771	3 481 533
NADRIN	16p12.1	Neuronal protein	55114	29985255	14335149–35125392	20 790 243
NUDE1	16p13.11	Lis-1 interacting protein	54820	18763116	14335149–35125392	See above
TAT	16q22.2	Tyrosine aminotransferase	6898	85891517	82945639–89466425	6 520 786
CFDP1	16q23.1	Craniofacial development protein 1	10428	90635426	90575990–90735398	159 408
DRG2[d]	17p11.2	Developmentally regulated GTP binding protein 2	1819	19787405	15371266–27948279	12 577 013
COPS3[d]	17p11.2	*Homo sapiens* COP9 complex subunit 3	8533	19038181	15371266–27948279	See above
RAI1[d]	17p11.2	Retinoic acid induced 1	10743	19492572	15371266–27948279	See above
NF1	17q11.2	Neurofibromin	4763	32548362	31949051–33721569	1 772 518
ACACA	17q12	Acetyl-coenzyme A carboxylase alpha	31	38913111	37945776–39868543	1 922 767
ASPA	17p13.2	Aspartoacylase	443	3267932	3120079–3546982	426 903
CLTC	17q23.2	Clathrin heavy chain	1213	65270461	65066121–65736364	670 243
TBX2	17q23.2	T-box 2	6909	67106821	65739747–68308666	2 568 919
IMPA2	18p11.21	Inositol(myo)-1(or 4)-monophosphatase 2	3613	13146682	12188020–13392618	1 204 598
LIPG	18q21.1	Endothelial lipase precursor	9388	54173714	54104840–54449609	344 769
FLJ14686	19q13.12	Zinc-finger protein 382	84911	44915972	44764587–46070350	1 305 763
NOSIP	19q13.33	Nitric oxide synthase interacting protein	51070	61526455	59278949–62006526	2 727 577
SPIB	19q13.33	Spi-B transcription factor (Spi-1/PU.1 related)	6689	62489290	62019412–62726350	706 938
ECR2[e]	22q11.1	Cat eye syndrome chromosome region, candidate	27443	14900358	13950072–21770926	7 820 854
DGCR2[f]	22q11.2	DiGeorge syndrome critical region gene 2	9993	15882238	13950072–21770926	See above
ADRBK2	22q12.1	Beta adrenergic receptor kinase 2	157	22657045	13950072–21770926	See above

The probes were designed using exon-specific single-copy sequences located in regions defined at a first pass by Bailey *et al*.[11] The localisation of the sequences is based on the Human Working draft of August 2001, as the duplicon data of Bailey is based on this information. Some of the probes tested were localised within the regions related to microdeletion syndromes:
[a]Williams syndrome. [b]Prader Willi. [c]Angelman syndrome. [d]Smith Magenis syndrome. [e]Cat eye syndrome. [f]22q11del/dup syndrome. *These genes are now known to be located in the WBCR; however, in the Human Working Draft of August 2001, these genes were located outside the Williams-related duplicons. The sum of all basepairs that are localised between two homologous intrachromosomal duplicons and tested in this study is 155 556 588 bp. This resembles 5.2% of the total human genome. The sum of all basepairs localised between duplicons and not tested in this study is 4.6% (calculation not shown). The total percentage of the genome flanked by duplicons identified at a first pass in 2001 is 9.8%.

Appendix B. Table B1. An overview of 58 probes containing function-selected genes localised outside the duplicons.

Gene	Chrom. band	Description	GeneID	Location on chromosome	Interval regions outside duplicons	Distance between nearest duplicons
MATN3	2p24.1	Matrilin 3	4148	20824361	Nearest 92015946	
FACL3	2q35	Acyl-CoA synthetase long-chain family	5147	233250166	137736981–242791383	105 054 402
PDE6D	2q37.1	Phosphodiesterase	7182	242633293	137736981–242791383	See above
NR2C2	3p25.1	Nuclear receptor subfamily	8087	20380273	Nearest 61700285	–
FXR1	3q26.33	Fragile X mental retardation	6750	206812624	204266505–223161158	18 894 653
SST	3q27.3	Somatostatin	10934	214072357	204266505–223161158	See above
MORF4	4q34.1	Mortality factor 4	27295	190237096	158224519–207089932	48 865 413
ALP	4q35.1	Actinin-assoc. protein	10409	203065626	158224519–207089932	See above
BASP1	5p15.1	Brain abundant, membrane-attached signal protein	2554	20241923	1306142–20506502	19 200 360
GABRA1	5q34	GABA receptor	3720	179091963	122206614–189063686	66 857 072
JMJ	6p22.3	Jumonji	9113	17824722	Nearest 28850598	–
LATS1	6q25.1	Tumour suppressor	4697	169320238	104572925–191797029	87 224 104
NDUFA4	7p21.3	NADH dehydrogenase	9162	11791714	7723723–31828578	24 104 855
DGK1	7q33	Diacylglycerol kinase oita	6456	148071336	138492661–155481235	16 988 574
SH3GL2	9p22.2	SH3-domain GRB2-like 2	80380	19168227	Nearest 37513397	–
PDL2	9p24.1	Programmed cell death 1 ligand 2	7099	5822952	Nearest 37513397	–
TLR4	9q33.1	Toll-like receptor 4	7248	130092974	Nearest 108546485	–
TSC1	9q34.13	Tuberous sclerosis 1	6812	146741930	Nearest 108546485	–
STXBP1	9q34.13	Syntaxin-binding protein	64376	141343041	Nearest 108546485	–
PEGASUS	10q26.12	Zinc-finger protein, subfamily 1A, 5	372	135434807	Nearest 86760491	–
HCCA2	11p15.5	YY1 associated protein	55249	649519	Nearest 3676771	–
ARCN1	11q23.3	Archain 1	6734	134627697	120328323–151288128	30 959 805
SRPR	11q24.2	Signal recognition particle receptor	93661	144635923	120328323–151288128	See above
CAPPA3	12p12.3	Actin-assoc. protein	10959	20564435	12446880–38117363	25 670 483
RNP24	12q24.31	Coated vesicle membrane protein	7223	143196772	Nearest 73099895	–
TRPC4	13q14.11	Transient receptor potential cation channel	2073	36935467	22748066–52126921	29 378 855
ERCC5	13q33.1	Excision repair cross-complementing rodent repair deficiency	1948	106470199	Nearest 64718332	–
EFNB2	13q34	Ephrin-B2	1603	110956506	Nearest 64718332	–
DAD1	14q11.2	Defender against cell death 1	801	19506399	16665813–20896466	4 230 653
CALM1	14q32.11	Calmodulin 1	6263	89723561	Nearest 20979168	–

Gene	Chrom. band	Description	GeneID	Location on chromosome	Interval regions outside duplicons	Distance between nearest duplicons
RYR3	15q14	Ryanodine receptor	27023	29382527	28243975–30470202	2 226 227
FOXB1	15q22.2	Forkhead box 1	3073	58697843	42719271–72062958	29 343 687
HEXA	15q23	Hexosaminidase A	3419	71660617	42719271–72062958	See above
IDH3A	15q24.3	Isocitrate dehydrogenase	98994828	79869724	75343377–85255238	9 911 861
SV2B	15q26.1	Synaptic vesicle protein	53739899	96312975	88736771–105696099	16 959 328
PMM2	16p13.2	Phosphomannomutase	64775373	10760222	4249026–14335149	10 086 123
SIAH1	16q12.1	Cell cycle control	43136477	57195286	38606337–82945639	44 339 302
MMP2	16q12.2	Metalloproteinase; collagen cleavage	70844313	65631286	38606337–82945639	See above
TK2	16q22.1	Mitochondrial thymidine kinase	40947084	79020967	38606337–82945639	See above
MAF	16q23.1	v-maf musculoaponeurotic fibrosarcoma oncogene	25884094	94753664	90735398–104966351	14 230 953
GALNS	16q24.3	N-acetylgalactosamine-6-sulphatase precursor	26701819	107082132	Nearest 106410182	–
CYBA	16q24.3	Flavocytochrome b-558 alpha polypeptide	1535	106941263	Nearest 106410182	–
GFAP	17q21.31	Glial fibrillary acidic protein	20812670	47569101	39868543–48411175	8 542 632
ERN1	17q23.3	Endoplasmic reticulum to nucleus signalling	646932081	70098037	Nearest 70040648	–
CTAGE-1	18q11.1	Cutaneous T-cell lymphoma-associated antigen	100064693	20577869	13392618–54104840	40 712 222
CDH2	18q12.2	Cadherin	1630	28317343	13392618–54104840	See above
DCC	18q21.2	Deleted in colorectal carcinoma	839831630	57850369	54449609–71373485	16 923 876
NTE	19p13.3	Neuropathy target esterase	872510908	10104831	Nearest 11662191	–
NOTCH3	19p13.12	Notch homolog 3	109084854	19462538	16128820–27052527	10 923 707
SSTK	19p13.11	Serine/threonine protein kinase	85483983	4674766	16128820–27052527	See above
RMP	19q12	Transcription modulating factor	81938725	37837413	28221927–44764587	16 542 660
NEUD4	19q13.12	Zinc-finger; neural specific	48588193	47104806	46070350–52764717	6 694 367
NOVA2	19q13.31	Neuro-oncological ventral antigen 2	298444858	57148580	53918418–59278949	5 360 531
TFPT	19q13.34	TCF3 (E2A) fusion partner	533529844	66371798	62726350–67370837	4 644 487
PLCG1	20q12	Phospholipase C, gamma 1	51215335	41433242	28921184–48156350	19 235 166
PCP4	21q22.2	Purkinje cell protein	5121	38093562	Nearest 12292280	–

These data are based on the Human Working draft of August 2001. The sum of all basepairs that are localised between two nearest nonhomologous intrachromosomal duplicons is 734 106 358 bp. This resembles 24.5% of the total human genome. The regions at the chromosome ends are not included in this calculation, as these are not localised between two nonhomologous intrachromosomal duplicons.

Chapter II-3

Diagnosis of genetic abnormalities in developmentally delayed patients: a new strategy combining MLPA and array-CGH

Marjolein Kriek[1*], Jeroen Knijnenburg[2*], Stefan J. White[1], Carla Rosenberg[2,3], Johan T. den Dunnen[1], Gert-Jan B. van Ommen[1], Hans J. Tanke[2], Martijn H. Breuning[1], and Karoly Szuhai[2]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, The Netherlands
[2]Department Molecular Cell Biology, Leiden University Medical Center, The Netherlands
[3]Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, Brazil
*These two authors contibuted equally

Am J Med Genet A. 2007 Mar 15;143(6):610-4.

Developmental delay (DD) affects ~3% of the general population and the underlying cause remains unknown in about half of the cases. G-banded karyotyping is the most common approach for the detection of genomic alterations, however, despite its indisputable success, this tool has limited resolution, usually being unable to detect genomic changes ≤3–5 Mb. It is known that micro alterations that escape detection by classical cytogenetics contribute substantially to the etiology of DD (Flint *et al.*, 1995; Vissers *et al.*, 2003). This limitation has been partly overcome by fluorescence in situ hybridization (FISH) with a resolution of 5–500 kb, however, it has a limited possibility for multiplexing, for example, in most of the routine practice only 2–3 regions can be analyzed simultaneously. Therefore, candidate probes (especially for microdeletion syndromes) need to be selected a priori for FISH investigation, based on the patient's phenotype.

Recent technological developments, such as array-based comparative genomic hybridization (array-CGH) (Pinkel *et al.*, 1998; Antonarakis, 2001; Snijders *et al.*, 2001) and Multiplex Ligation-dependent Probe Amplification (MLPA) (Schouten *et al.*, 2002), are efficient methods for screening for copy number imbalances in multiple genomic regions simultaneously. MLPA especially has already found its way into the diagnostic laboratories for several indications (e.g., *BRCAI* gene and *NFI* gene screening); however, the standard of practice for the assessment of developmental delay does not currently include MLPA and array-CGH testing. In this article, it is argued that both techniques are extremely valuable tools for the diagnostic setting in DD patients, and the implementation of both techniques should be considered.

Data regarding the robustness of both techniques have been provided previously (Price *et al.*, 2005; Rooms *et al.*, 2005). In the case of array-CGH, thousands of sites can be simultaneously investigated in one patient, allowing partial or total coverage of the genome. The number of targets that can be screened by MLPA is limited to <60 loci per assay, however, 96 samples can be simultaneously tested at a cost less than one array-CGH hybridization. As MLPA analysis requires relatively little hands-on time (Table I), it is more suitable for the initial screening of large patient numbers.

To assess their value in clinical diagnosis, we have independently tested 58 developmentally delayed (DD) patients using both array-CGH and MLPA. This study was reviewed and approved by the Institutional Review Board of the Leiden University Medical Center, conforming to Dutch law and the World Medical Association Declaration of Helsinki. The patients had, in addition to DD, either dysmorphic features or congenital malformations or both (DD "plus" patients). All patients had a normal karyotype and, where tested (the vast majority of the patients),had tested negative for-

Table 1. A comparison of the man-hours and material required for both karyotyping and MLPA analysis.

	Karyotyping	**MLPA**
Number of samples performed per week	12	5 × 96 wells plate
Total time before result per sample	32–40 hr	8 hr[a]
Materials needed	Cell culture, reagents	DNA reagents, probe set

This table shows that MLPA is suitable for the screening of copy number variations in a large number of patients within relatively short time. Compared to karyotyping, this technique is much faster and requires less hands-on time. As it is also possible to analyze a part of a fragment run or use a DNA sequencer with less throughput capacity, it is not necessary to wait for 96 patient samples requiring MLPA testing.
[a]Recently, it was shown that MLPA analysis can be performed within 8 hr (Kalf *et al.* in preparation).

FragileX syndrome. The array-CGH results were partly reported elsewhere (Rosenberg *et al.*, 2006) without the comparative analysis with MLPA.

The array used in the study contained ~3,500 large genomic insert clones spaced at 1 Mb intervals over the genome, meaning thattheresolution ofthearrays used is 0.3–3 Mb. Array-CGH testing was performed as described by (Knijnenburg *et al.*, 2005). The clones were provided by the Wellcome Trust Sanger Institute (UK), and information regarding the full set is available at the Ensemble web site.

The MLPA probe design and assay was performed as described previously (White *et al.*, 2004). It included a set of synthetic probes designed for 71 regions known to be frequently altered in DD patients (probe sequences are available on request). This set targets 42 chromosome ends (except for the p-arms of the acrocentric chromosomes), five pericentromeric regions on the q-arm of acrocentric chromosomes (the regions tested included the first gene-specific unique sequence near the centromere on the q-arm) and 24 probes (Table II) containing microdeletion syndrome-relatedsequences. The size of the probes used was between 75 and 125 bp, and the number of sites investigated by MLPA corresponds to ~2% (71/3,500) of all regions tested by array-CGH.

Seventeen alterations were detected by array-CGH analysis, of which 14 were verified using either FISH or MLPA (14/58 = 24%). (The MLPA probes were specifically designed for confirming these alterations. They were not part of the screening set.) As far as was tested the remaining three changes could not be confirmed using FISH or MLPA.

MLPA analysis identified eight alterations, all of which were confirmed by FISH, MAPH or sequencing (8/58 = 14%). Table III provides an overview of the alterations found. The eight alterations found solely by array-CGH were all located in regions not

Table 2. Overview of the microdeletion syndrome-related probes used by MLPA screening.

Disorder	Chromosome band	Gene
Alagille syndrome	20p12.2	*JAG1*
Angelman syndrome	15q12	*UBE3A*
Cat eye syndrome	22q11.1	*CECR2*
DiGeorge syndrome	22q11.2	*DGCR2*
DiGeorge syndrome	22q11.2	*HIRA*
DiGeorge syndrome	22q11.2	*TBX1*
DiGeorge syndrome	22q11.2	*UFD1L*
DiGeorge syndrome like region	10p14	*CUGBP2*
Extostosis	8q24	*EXT1*
Jacobsen syndrome	11q25	*HNT*
Miller–Dieker syndrome	17p13.3	*LIS 1*
Mowat–Wilson syndrome	2q22	*SIP1*
Prader–Willi syndrome	15q12	*SNRPN*
RETT syndrome	Xq28	*MECP2*
Rubinstein–Taybi syndrome	16p13,3	*CBP*
Smith–Magenis syndrome	17p11.2	*RAI1*
Smith–Magenis syndrome	17p11.2	*COPS3*
Smith–Magenis syndrome	17p11.2	*DRG2*
Sotos syndrome	5q35	*NSD1*
Trichorhinophalangeal syndrome	8q23.3	*TRPS1*
William–Beuren syndrome	7q11.23	*ELN*
William–Beuren syndrome	7q11.23	*FKBP6*
Wolf–Hirschhorn syndrome	4p16.2	*MSX1*
X-linked hydrocephalus	Xq28	*L1CAM*

covered by MLPA probes. In contrast, the two alterations detected by MLPA only were too small to be detected by array-CGH analysis. One of these alterations was a point mutation near the ligation site of the MLPA probe, which disturbed the ligation and appeared as a deletion. The point mutation (that was never reported before) has been proven by bi-directional sequencing. It is a silent mutation, and it was also present in one of the parents. Therefore, it was considered to be a single nucleotide polymorphism (SNP). Although all MLPA probes have been designed outside the sequences containing known SNPs, theoretically, a low frequency SNP could be present at or near the ligation site. Therefore, it is necessary to confirm copy number variations by a second MLPA probe covering an adjacent sequence or by sequencing.

Of the eight alterations detected by MLPA, we considered six to be probably caus-

Table 3. Copy number variations detected by two techniques independently.

	Only by a-CGH	Only by MLPA	By a-CGH and MLPA	Total
Altered	11	2	6	19
Confirmed	8[a]	2[b]	6	16

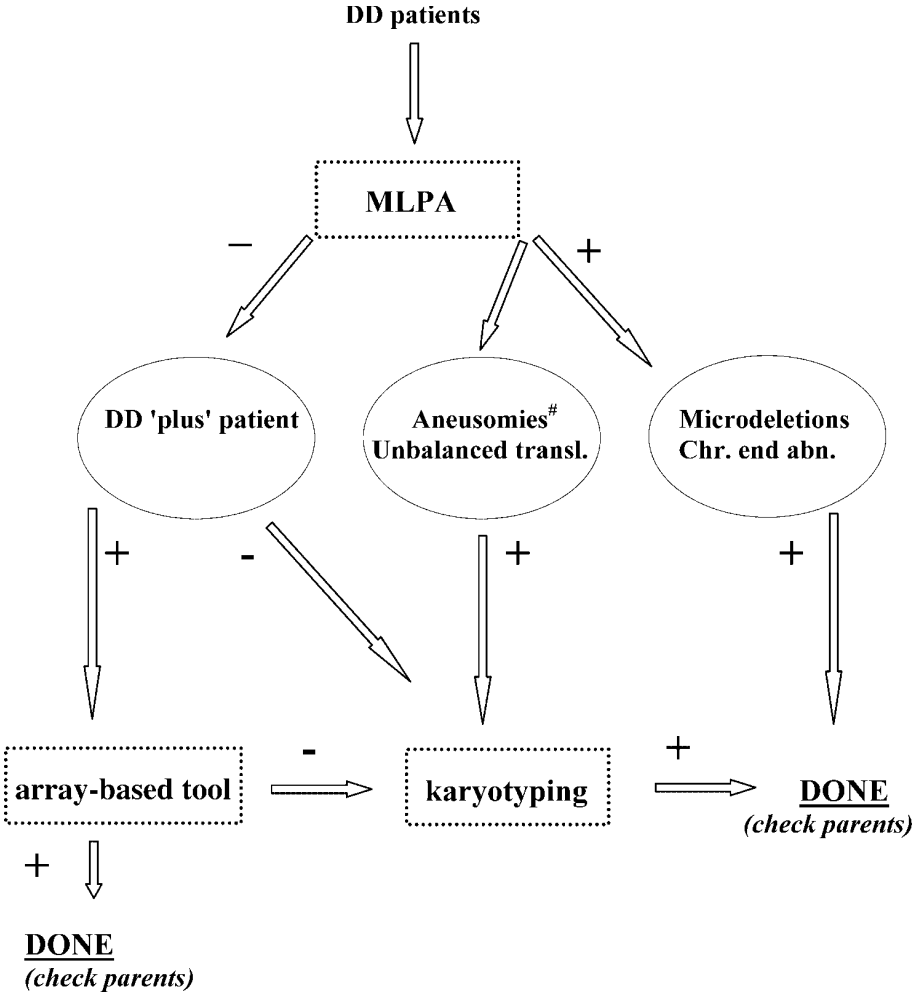
	De novo	Present in parents	Unknown	*De novo*	Present in parents	Unknown	*De novo*	Present in parents	Unknown	
Deletion	2	3	1[c]	0	1[d]	0	2[c]	0	1[c,e]	10
Duplication	1	1	0	0	1[d]	0	1[c]	0	0	4
del./dup.	0	0	0	0	0	0	0	0	1[d,e]	1
UT	0	0	0	0	0	0	1[c]	0	0	1
Confirmed total	3	4	1	0	2	0	4	0	2	16

An overview of the results obtained by screening of 58 DD patients using array-CGH and MLPA. All rearrangements were not detected by routine karyotyping.
UT, unbalanced translocation.
a These regions were not covered by MLPA analysis.
b These alterations were too small to be detected by array-CGH.
c Alterations localized at the chromosome ends.
d Alterations present in regions related to micro-deletion syndromes.
e (One of) the patient's parents were (was) unavailable for testing. The phenotype of the patient, however, resembles that described in literature. Therefore, this alteration is thought to be pathogenic.

ative as the phenotype of the patients agreed with the clinical features described in literature for those chromosome alterations. All these rearrangements were also detected by array-CGH. In two of these six cases, however, we could not confirm that the rearrangement was *de novo*. Two of the eight alterations detected by MLPA are likely to be polymorphic variants, as they are also present in unaffected family members.

Nine of the fourteen confirmed rearrangements detected by array-CGH are probably pathogenic, four alterations might be polymorphic variants as they are present in unaffected family members. The clinical consequences of the remaining alteration are currently unknown, because the patients' parents were unavailable for testing. This latest FISH confirmed array-CGH finding, which was not detected by MLPA, was located near the chromosome end of the long arm of chromosome 10. The corresponding "subtelomeric" MLPA probe in our study mapped proximal to the altered BAC. Based on the data on the human genome variation database, theregioninvolved might bepolymorphic. Moreover, the clinical features of the patient do not resemble those corresponding with previously described 10q chromosome end alterations (Waggoner *et al.*, 1999). The sizes of the reported alterations, however, are larger than the one obtained in this study.

Figure 1. Alternative diagnostic approach.



This flow chart summarizes the alternative diagnostic approach for screening developmentally delayed patient samples. In this approach, karyotyping will only be requested for a selected group of samples: (1) Samples that had tested negative for MLPA (and array-based tool in the case of DD "plus" patients). (2) Samples for which information about the location of the structural rearrangement is essential for clinical practice. These include aneusomies for which a Robertsonian translocation should be excluded (acrocentric chromosomes (#)), unbalanced translocations and some of the alterations detected by array-CGH. Chr. end abn.: chromosome end abnormality, DD "plus" patients are patients with dysmorphic features and/or congenital malformations in addition to DD. These patients are suggestive for chromosomal imbalances.

The comparison between the screening results for detecting copy number variations using the different approaches shows the reliability and specific strengths of both techniques. In summary, using ~2% of the loci tested by array-CGH, MLPA detected 50% (8/16) of all alterations. Three potentially pathogenic alterations were not detected using MLPA, as they were localized outside the regions tested.

Based on the outcome of this parallel screening and costs considerations, we suggest the following strategy for diagnostic purposes: when a patient presents with DD of unclear etiology and the G-banding karyotype is normal, the first screening will use MLPA for the commonly altered regions in DD patients (currently, chromosome ends and microdeletion syndrome-related regions). Subsequently, when MLPA is negative and the patient's phenotype is suggestive of a chromosome abnormality, array-CGH follows.

Alternatively, the order of testing could be reversed. MLPA using subtelomeric probes is capable of detecting trisomies as well as the vast majority of the unbalanced translocations, both of which comprise a substantial part of the alterations diagnosedusing cytogenetic tools. Table I shows that MLPA requires less manpower (hence is cheaper) and is considerably faster compared to karyotyping, and thus, it seems more effective to use MLPA as an initial screening tool. In addition to the time-and cost-effectiveness, MLPA has a much higher resolution for detecting copy number variations compared to karyotyping, and therefore, this technique is capable of detecting copy number variations that remain undiagnosed using this cytogenetic tool. Applying MLPA testing first will even be more effective when a MLPA probe set encompassing the most frequent microdeletion related regions is added. In a diagnostic setting, it is preferable to have at least two MLPA probes per regions of interest (instead of one as was used in this study) to limit false positive and false negative results as much as possible. Implementing microdeletion syndrome-related regions and two probes per region will increase the costs related to MLPA screening, however, this will also reduce the necessity of performing FISH for the detection of microdeletion syndromes, and the need for additional confirmation tests (with the exception of sequencing, see above).

It is obvious that balanced translocations and inversions will not be detected using this or other molecular techniques (unless they are specifically designed to detect breakpoints). Also, for a proportion of the samples with a positive outcome using the initial MLPA screening, subsequent karyotyping is essential for localization of these structural rearrangements. These include, for example, aneusomies for which Robertsonian translocations have to be excluded. Based on these arguments, karyotyping will

maintain its essential role in a diagnostic process, however it will only be implemented for selected samples.

After MLPA testing, additional array-CGH can be performed for patients with a clinical phenotype suggestive for chromosomal alterations. Although this will increase the cost, it will also increase the number of copy number variations detected.

Array-based techniques are evolving rapidly. Several reports have described the results of testing developmentally delayed patients tested using a 3,000-clone array (Vissers *et al.*, 2003; Tyson *et al.*, 2005; Menten *et al.*, 2006; Rosenberg *et al.*, 2006; Shaw-Smith *et al.*, 2006). In addition, de Vries *et al.* (2005) used an array with 32,000 clones for the detection of copy number variations. Recently, SNP-based arrays have successfully been used to detect genome-wide copy number variations (Friedman *et al.*, 2006). These type of arrays have an even higher resolution than the array used in de Vries *et al.* Future comparative studies will help to determine which array platform is the most appropriate to implement.

In short, the alternative diagnostic approach would include MLPA for DD samples, with subsequent array-based testing (for DD "plus" patients that had tested negative for MLPA). Karyotyping could then be used to locate structural rearrangements for selected cases and for samples that showed no alteration using MLPA (and array-CGH) (Fig. 1). In this way, the screening of DD samples will be more effective in relation to the probability of finding a disease-causing rearrangement, which will improve the basis for counseling.

Acknowledgements

M.K. is funded by Zon-Mw (AGIKO fellowship 940-37-032), S.W. is funded by ZonMw (nr 912-04047). C.R. is funded by CNPq. This work was partially supported by the "Doelmatigheid" grant from the Leiden University Medical Center (2002/2003). We thank Claudia Ruivenkamp for critically reading the manuscript and implementing the cytogenetics information and we thank the physicians for providing clinical information and patient material.

References

Antonarakis SE. 2001. BACking up the promises. *Nat Genet* **27**:230–232.

De Vries BB, Pfundt R, Leisink M, Koolen DA, Vissers LE, Janssen IM, Reijmersdal S, Nillesen WM, Huys EH, Leeuw N, Smeets D, Sistermans EA, Feuth T, Ravenswaaij-Arts CM, van Kessel AG, Schoenmakers EF, Brunner HG, Veltman JA. 2005. Diagnostic genome profiling in mental retardation. *Am J Hum Genet* **77**:606–616.

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

Friedman JM, Baross A, Delaney AD, Ally A, Arbour L, Asano J, Bailey DK, Barber S, Birch P, Brown-John M, Cao M, Chan S, Charest DL, Farnoud N, Fernandes N, Flibotte S, Go A, Gibson WT, Holt RA, Jones SJ, Kennedy GC, Krzywinski M, Langlois S, Li HI, McGillivray BC, Nayar T, Pugh TJ, Rajcan-Separovic E, Schein JE, Schnerch A, Siddiqui A, Van Allen MI, Wilson G, Yong SL, Zahir F, Eydoux P, Marra MA. 2006. Oligonucleotide microarray analysis of genomic imbalance in children with mentalretardation. *Am J Hum Genet* **79**:500–513.

Knijnenburg J, Szuhai K, Giltay J, Molenaar L, Sloos W, Poot M, Tanke HJ, Rosenberg C. 2005. Insights from genomic microarrays into structural chromosome rearrangements. *Am J Med Genet* Part A **132A**:36–40.

Menten B, Maas N, Thienpont B, Buysse K, Vandesompele J, Melotte C, de Ravel T, Van Vooren S, Balikova I, Backx L, Janssens S, De Paepe A, De Moor B, Moreau Y, Marynen P, Fryns JP, Mortier G, Devriendt K, Speleman F, Vermeesch JR. 2006.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 publishedreports. *J Med Genet* **43**:625–633.

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. 1998. High resolution analysis of DNA copy number-variation using comparative genomic hybridizati onto microarrays. *Nat Genet* **20**:207–211.

Price TS, Regan R, Mott R, Hedman A, Honey B, Daniels RJ, Smith L, Greenfield A, Tiganescu A, Buckle V, Ventress N, Ayyub H, Salhan A, Pedraza-Diaz S, Broxholme J, Ragoussis J, Higgs DR, Flint J, Knight SJ. 2005. SW-ARRAY: A dynamic programming solution for the identification of copy-number changes in genomic DNA using array comparative genome hybridization data. *Nucleic Acids Res* **33**:3455–3464.

Rooms L, Reyniers E, Kooy RF. 2005. Subtelomeric rearrangements in the mentally retarded: A comparison of detection methods. *Hum Mutat* **25**:513–524.

Rosenberg C, Knijnenburg J, Bakker E, Vianna-Morgante AM, Sloos W, Otto PA, Kriek M, Hansson K, Krepischi-Santos ACV, Fiegler H, Carter NP, Bijlsma EK, Van Haeringen A, Szuhai K, Tanke HJ. 2006. Array-CGH detection of micro rearrangements inmentallyretarded individuals: Clinical significanceof imbalances present both in affected children and normal parents. *J Med Genet* **43**:180–186.

Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* **30**:e57.

Shaw-Smith C, Pittman AM, Willatt L, Martin H, Rickman L, Gribble S, Curley R, Cumming S, Dunn C, Kalaitzopoulos D, Porter K, Prigmore E, Krepischi-Santos AC, Varela MC, Koiffmann CP, Lees AJ, Rosenberg C, Firth HV, de Silva R, Carter NP. 2006. Microdeletion encompassing

MAPT at chromosome 17q21.3 is associated with developmental delay andlearning disability. *Nat Genet* **38**:1032–1037.

Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, Law S, MyamboK, PalmerJ, YlstraB,YueJP, GrayJW, JainAN,Pinkel D, Albertson DG. 2001. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* **29**:263–264.

Tyson C, Harvard C, Locker R, Friedman JM, Langlois S, Lewis ME, Van Allen M, Somerville M, Arbour L, Clarke L, McGilivray B, Yong SL, Siegel-Bartel J, Rajcan-Separovic E. 2005. Submicroscopic deletions and duplications in individuals with intellectual disability detected by array-CGH. *Am J Med Genet* Part A **139A**:173–185.

Vissers LE, De Vries BB, Osoegawa K, Janssen IM, Feuth T, Choy CO, Straatman H, van der Vliet W, Huys EH, Van Rijk A, Smeets D, Ravenswaaij-Arts CM, Knoers NV, Van der Burgt I, De Jong PJ, Brunner HG, van Kessel AG, Schoenmakers EF, Veltman JA. 2003. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* **73**:1261–1270.

Waggoner DJ, Chow CK, Dowton SB, Watson MS. 1999. Partial monosomy of distal 10q: Three new cases and a review. *Am J Med Genet* **86**:1–5.

White SJ, Vink GR, Kriek M, Wuyts W, Schouten J, Bakker B, Breuning MH, Dunnen JT. 2004.Two-color multiplex ligation dependent probe amplification: Detecting genomic rearrangements in hereditary multiple exostoses. *Hum Mutat* **24**: 86–92.

Chapter III

Case report based findings

Chapter III-1

A complex rearrangement on chromosome 22 affecting both homologues; haplo-insufficiency of the Cat eye syndrome region may have no clinical relevance

Marjolein Kriek[1], Karoly Szuhai[2], Sarina Kant[1], Stefan White[1], Hans Dauwerse[1], Heike Fiegler[3], Nigel P. Carter[3], Jeroen Knijnenburg[2], Johan den Dunnen[1], Hans Tanke[2], Martijn Breuning[1], Carla Rosenberg[2,4]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, The Netherlands
[2]Dept. Molecular Cell Biology, Leiden University Medical Center, The Netherlands
[3]Wellcome Trust Sanger Institute, Wellcome trust Genome Campus, Hinxton, Cambridge, UK.
[4]Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, Brazil.

Hum Genet. 2006 Aug;120(1):77-84.

Abstract

The presence of highly homologous sequences, known as low copy repeats, predisposes for unequal recombination within the 22q11 region. This can lead to genomic imbalances associated with several known genetic disorders. We report here a developmentally delayed patient carrying different rearrangements on both chromosome 22 homologues, including a previously unreported rearrangement within the 22q11 region. One homologue carries a deletion of the proximal part of chromosome band 22q11. To our knowledge, a 'pure' deletion of this region has not been described previously. Four copies of this 22q11 region, however, are associated with Cat eye syndrome (CES). While the phenotypic impact of this deletion is unclear, familial investigation revealed five normal relatives carrying this deletion, suggesting that haplo-insufficiency of the CES region has little clinical relevance. The other chromosome 22 homologue carries a duplication of the Velocardiofacial/DiGeorge syndrome (VCFS/DGS) region. In addition, a previously undescribed deletion of 22q12.1, located in a relatively gene-poor region, was identified. As the clinical features of patients suffering from a duplication of the VCFS/DGS region have proven to be extremely variable, it is impossible to postulate as to the contribution of the 22q12.1 deletion to the phenotype of the patient. Additional patients with a deletion within this region are needed to establish the consequences of this copy number alteration. This study highlights the value of using different genomic approaches to unravel chromosomal alterations in order to study their phenotypic impact.

Introduction

The 22q11 region contains highly homologous regions known as low copy repeat (LCR) sequences. Despite the difference in size and organisation of these repeats, the overall sequence identity is 97–98% (Shaikh *et al*. 2000). It has been demonstrated that the presence of these LCRs can initiate misaligned (non-) allelic homologous recombination of the region flanked by these duplicons, resulting in a deletion and an obligate reciprocal duplication (McDermid and Morrow 2002; Bailey *et al*. 2002). As a result, 22q11 is associated with different genomic disorders (Table 1). The 22q11 related disorders display a wide variety of clinical features, with no obvious correlation between the size of the genomic imbalance and the severity of the clinical characteristics.

The most common genetic disorder in this region is the Velocardiofacial syndrome/DiGeorge syndrome (VCFS/DGS) [MIM # 192430; MIM #188400], affecting 1:4,000–6,000 individuals. This syndrome is caused by haplo-insufficiency of the 22q11.2 region. Over 90% of patients suffering from VCFS/DGS have a 3 Mb deletion between LCR22-2 and LCR22-4 (Fig. 1), that share a particularly high degree of homology. The rest of the patients have a smaller rearrangement (1.5 Mb) with breakpoints located in LCR22-2 and LCR22-3a. Fluorescent in situ hybridisation (FISH) analysis is 100% accurate in detecting VCFS. However, the DGS phenotype can also be caused by other genetic (e.g. 10p13 deletion) or non-genetic causes (Robin and Sprintzen 2005). Most of the affected organs (thymus, (para)thyroid gland, outflow area of the heart) in VCFS/ DGS (Table 1) are derived from the third, fourth and sixth branchial arch in early development. Recently, it became apparent that VCFS/DGS are due to developmental deficiency of the endodermal pharyngeal pouches and the pharyngeal mesoderm, rather than (migration) defects of the neural crest cells (Graham 2003). As the *TBX1* gene is strongly expressed in the branchial arches, it is assumed that mutations in this gene are responsible for the majority of the features of VCFS/DGS (Jerome and Papaioannou 2001; Lindsay *et al*. 2001; Mercher *et al*. 2001).

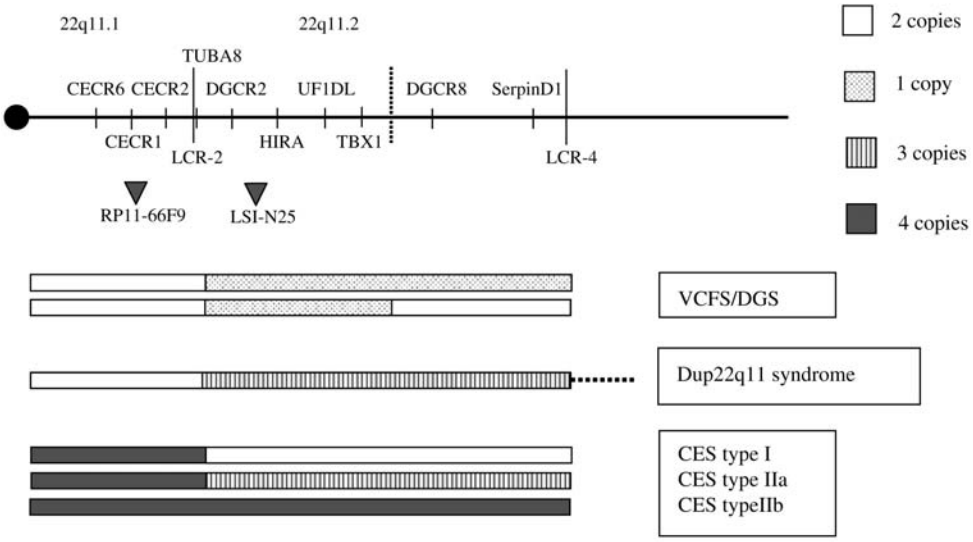
In 1999, the first report of the reciprocal duplication of the VCFS/DGS region was published. The phenotypic variability associated with the duplication of the VCFS/ DGS region emerged as a healthy mother and grandmother had the same duplication as the affected individual (Edelmann *et al*. 1999). Ensenauer *et al*. (2003) summarised the clinical characteristics of 13 patients with 22q11 duplications of variable sizes (3, 4, 6 Mb). More recently, the clinical features of another seven

Table 1. Overview of different 22q11 related syndromes.

Name of syndrome	Rearrangement	Clinical features
Velocardiofacial syndrome/DiGeorge syndrome	Deletion of the 22q11.2 region	DD, facial dysmorphisms (micrognathia, short philtrum and ear anomalies), cleft palate, cardiac outflow tract defects, Tetralogy of Fallot, nasal speech, hypocalcemia, thymic hypoplasia and behavioural problems (especially schizophrenia).
22q11.2 duplication syndrome	Duplication of the 22q11.2 region	Extremely variable. Clinical features of these patients could show similarities with those described in VCFS/DGS [DD (± motor delay), poor growth, dysmorphic features, velopharyngeal malformation ± cleft palate, urogenital malformations, hearing loss) However, dysmorphic features and behavioural problems not related to the VCFS/DGS spectrum have also been described (see text).
Cat eye syndrome	Quadruplication of the 22q11.1 region	Ocular coloboma, downslanting palpebral fissures, preauricular tags and/or pits, anal atresia with fistula, frequent occurrence of congenital heart and renal malformations and normal to near-normal mental development.

DD developmental delay

Figure 1 Overview of three 22q11 related syndromes in relation to the location of the different MAPH and BAC probes (RP11-66F9, N25) used in this study.



The size of the majority of the deletions within the VCFS/DGS regions is 3 Mb. The remaining deletions of this region encompass 1.5 Mb. The distal breakpoint of the duplications of the VCFS/DGS region is not always localised within LCR-4 (indicated by a *dotted line*) (Ensenauer *et al.* 2003). Different types of CES are depicted. This figure is based on Fig. 2 of McDermid and Morrow *et al.* (2002).

22q11 duplication syndrome patients were described, showing a very wide range of clinical variability. Furthermore, the first triplication of 22q11.2 was described (Yobb *et al*. 2005).

The Cat eye syndrome (CES [MIM #115470]) has three different subtypes: CES type I, CES type IIa and CES type IIb (Fig. 1). The endpoint of CES type I colocalises with LCR-2 and consists of two extra copies of the CES region only. Patients with CES type IIa have four copies of the CES region combined with three copies of the VCFS/DGS region. CES type IIb consists of four copies of both the CES region and the VCFS/ DGS region. The endpoint of both CES type IIa and IIB is mapped to LCR-4 (McDermid and Morrow 2002).

So far, a deletion of the Cat eye critical region has never been reported.

In this report, we describe the clinical features of a patient with complex chromosome 22 rearrangements, including a previously undescribed familial deletion of CES region in one homologue and, a duplication of VCFS/DGS region of the other homologue, in addition to a deletion of 22q12.1. These imbalances were characterised using different techniques: multiplex amplifiable probe hybridisation (MAPH), multiplex ligation-dependent amplification (MLPA), fluorescence in situ hybridisation (FISH), array-based comparative genomic hybridisation (array-CGH).

Clinical report

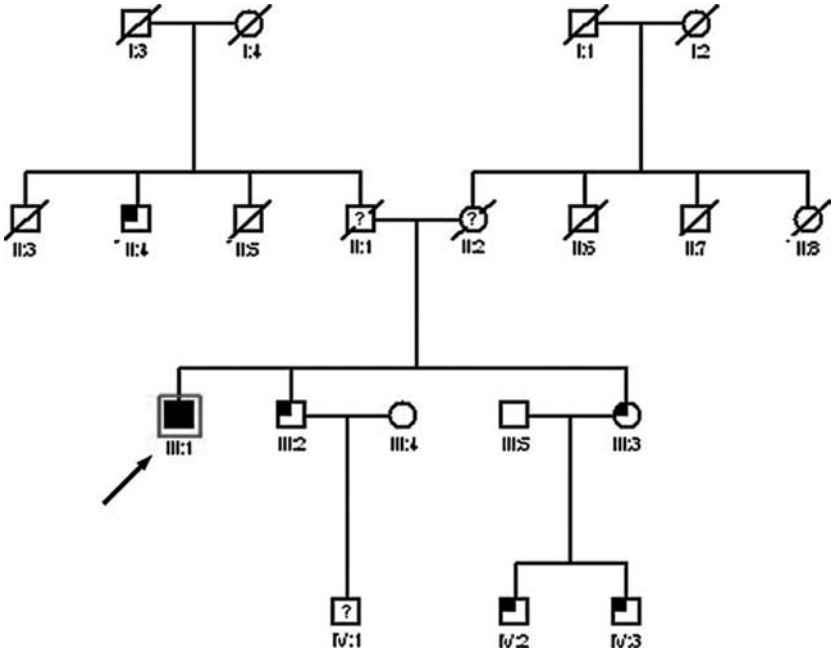
The male patient was born by forceps delivery after an uneventful pregnancy. At birth, a cleft palate was diagnosed and he was reported to have a double set of teeth. The cleft palate was corrected by surgical treatment at the age of two and five. He attended special education because of hearing loss and moderate mental retardation. From his early adolescence onwards, he has been living in a support home. His further medical history included cataract and myopia.

At the time of evaluation in the clinical genetics department, this patient was 52 years old (Fig. 2). Physical examination showed hypertonia with wooden movements. His speech was slow and difficult to comprehend and he tended to avoid eye contact. He had a normal height of 172 cm (–1.5 SD), microcephaly (head circumference 51.2 cm: –3.8 SD), round face with hypotonic expression, proptosis of the eyes, prominent simple ears, earpits on both sides and short fifth fingers. His heart tones were normal and no murmur was diagnosed. His medical record shows no history of cardiac problems.

Figure 2 Picture of the proband.

Note the microcephaly, myotonic facial expression, the proptosis of the eyes and the prominent simple ears. [See appendix: colour figures.]

Figure 3. The pedigree of proband III-1.



A square symbol and an arrow mark the proband. The symbol (■) represents all five unaffected family members with a deletion of the CES region.

Familial history

The pedigree of the family is shown in Fig. 3. Familial history included two siblings with children and grandchildren, all of them healthy. The index patient's father died at the age of 81 years of unknown causes. His mother died due to a cardiovascular accident at the age of 79 years. The overall familial history showed no other individuals with developmental delay, nor any other genetic disorders.

Additional investigation

Additional investigation showed a normal male karyotype and a normal number of CGG repeats of the *FMR1* gene. FISH analysis was performed for the detection of a deletion of chromosome band 22q11.2 (using TUPLE1 probe) and for the detection of a deletion of 4p16.3 (Wolf–Hirschhorn syndrome) (using LSI-WHS probe). No rearrangements were detected. DNA testing for myotonic dystrophy type 1 showed normal CTG repeats on both alleles.

Materials and methods

Patients

This study was approved by the Institutional Review Board of the Leiden University Medical Center, conforming to Dutch law. All subjects or their representatives gave informed consent for DNA studies.

MAPH and MLPA

Multiplex amplifiable probe hybridisation was performed as described by White *et al.* (2002). The probe set used contained 19 probes from genes on chromosome 22 with approximately 1 Mb spacing, and ten additional genes in the 22q11 region.

A modified protocol of MLPA (Schouten *et al.* 2002) was performed as described by White *et al.* (2004).

Array-comparative genomic hybridisation (array-CGH)

The array-CGH procedures were performed as described (Knijnenburg *et al.* 2004). Briefly, slides containing triplicates of ~3,500 BAC DNA probes spaced at ~1 Mb density over the full genome were produced in the Leiden Technology Center (LGTC). The BAC set used to produce these arrays was received from the Welcome Trust Sanger Institute (UK), and information regarding the full set is available in the genome browser, Ensembl (http://www.ensembl.org/).

Tiling path array

The chromosome 22 tile path array and its hybridisation and analysis were performed as described by Woodfine *et al.* (2004).

Fluorescence in situ hybridisation (FISH)

The FISH experiments were performed by standard procedures (Dauwerse *et al.* 1992).

The CES region specific BAC RP11-66F9 were visualised using Alexa594 (green). For the identification of chromosome 22, the telomere specific BACs LSI-ARSA and RP11-3018K1 (22q13) (Flint and Knight 2003) was used and visualised using FITC (green). The VCFS/ DGS region was tested by N25 Probe (Vysis). This probe consists of N25 in red (SpectrumOrange).

Results

MAPH and MLPA

Multiplex amplifiable probe hybridisation analysis of the index patients DNA revealed a deletion of probes within the CES region, (*CECR1* (GeneID: 51816), *CECR2* (GeneID: 27443), *CECR6* (GeneID: 27439) sequence) and a duplication of the probes containing sequences within VCFS/DGS region (*DGCR2* (GeneID: 9993), *DGCR8* (GeneID: 54487), *TBX1* (GeneID: 6899), *UF1DL* (GeneID: 7353), *HIRA* (GeneID: 7290), *SERPIND1* (GeneID: 3053). The *TUBA8* gene (GeneID: 51807), localised between CECR and DGCR, showed a normal copy number (Fig. 4).

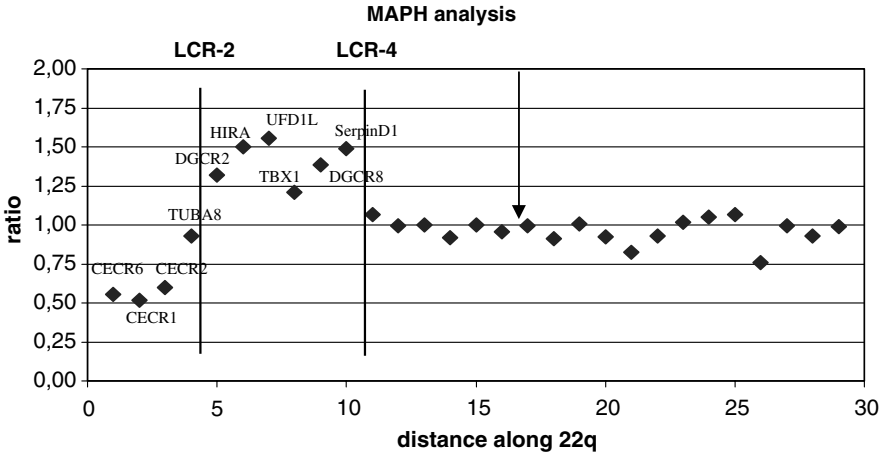
It was not possible to test the patient's parents; however, we were able to test several other healthy family members. The results are summarised in Fig. 3 and show that both siblings, two of their children and a brother of the patients' father carried the same deletion in the Cat eye region as the index patient. The duplication of 22q11.2 was absent in all family members tested.

After verification of these findings with MLPA using sequences of *CECR2* gene and *DGCR2* gene, the characteristics of the genetic rearrangements of index patient were refined by different techniques.

Array-CGH

Array-CGH using a 3500 BAC array was initially carried out to define the length of each of the two rearrangements. However, this analysis revealed a third chromo-

Figure 4. MAPH analysis of chromosome 22.



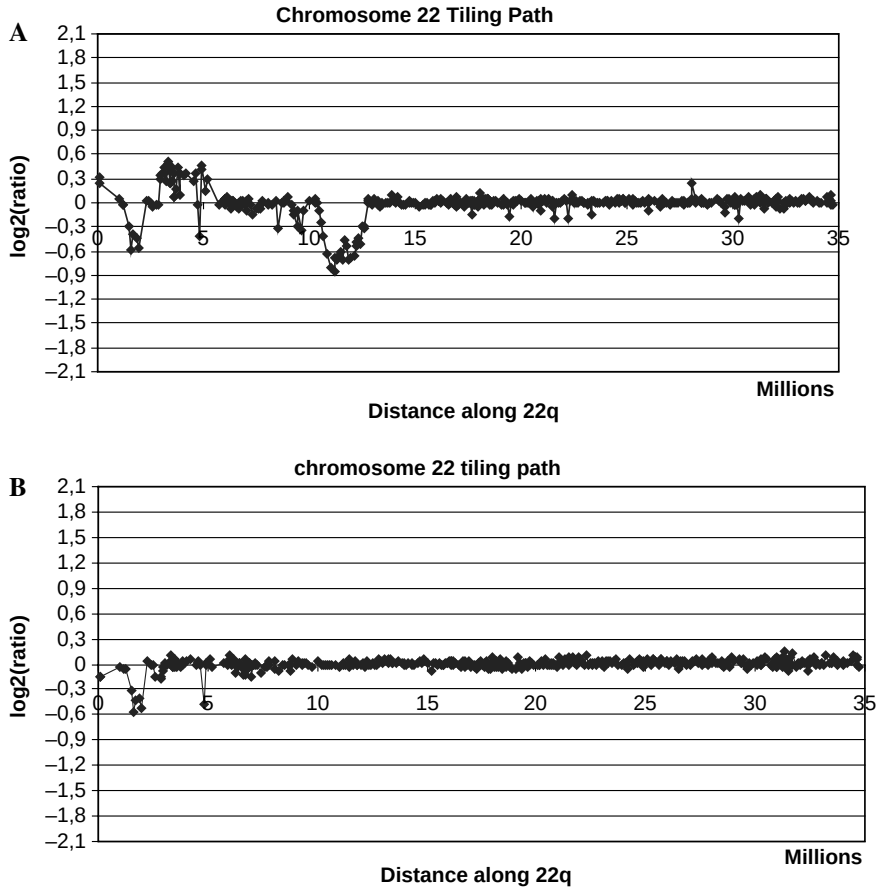
MAPH plot of chromosome 22 revealing a deletion of the CES region and a duplication of the VCFS/DGS region in the proband. A probe covering 22q12.1 was not included. The *arrow* indicates the locus of the 22q12.1 deletion.

some 22 alteration, namely a deletion of 22q12.1. The deleted area was localised about 25 Mb distal from the 22q11.2 region, between BACs CTA-57G9 and CTB-48E9. As this deletion was present in a relatively gene-poor region, the chromosome 22-MAPH-probe set did not contain a probe in this region. The duplication of 22q11.2 and the deletion of chromosome 22q12.1 were not present in the healthy brother of the index patient.

Chromosome-22-tiling-path array

To map the breakpoints of the alterations at a higher level of resolution, the patient's DNA and that of his brother were analysed on a chromosome-22-tiling-path array, as shown in Fig. 5. The sizes of the deletion and the duplication are 1.5 and 4.1 Mb, respectively. The transition of the deletion and the duplication within the 22q11 region maps to the LCR22-2. As the size of the duplication is larger than 3 Mb, the distal breakpoint of the alteration is not localised within LCR22-4, being localised more distally. The distal deletion on chromosome band 22q12.1 encompasses 2.3 Mb on chromosome band 22q12.1 and is not flanked by intrachromosomal LCRs. This region, however, is flanked by sequences that share high homology with sequences localised on other chromosomes.

Figure 5. Chromosome 22 tiling path array.



A Tiling path array analysis of the proband revealed the deletion and duplication of 22q11 subregions and a distal deletion of chromosome band 22q12.1. The sizes of the rearrangements are 1.5, 4.1 and 2.3 Mb, respectively. One BAC within the VCFS/DGS region shows an aberrant ratio. The cause of this aberration is currently unknown. B Tiling path array analysis of the healthy brother of the proband confirmed the presence of the proximal 22q11 deletion but the absence of the duplication of 22q11 and the deletion of 22q12.1.

Additional familial investigation using MLPA showed that none of the family members with a 22q11 deletion carried the distal deletion.

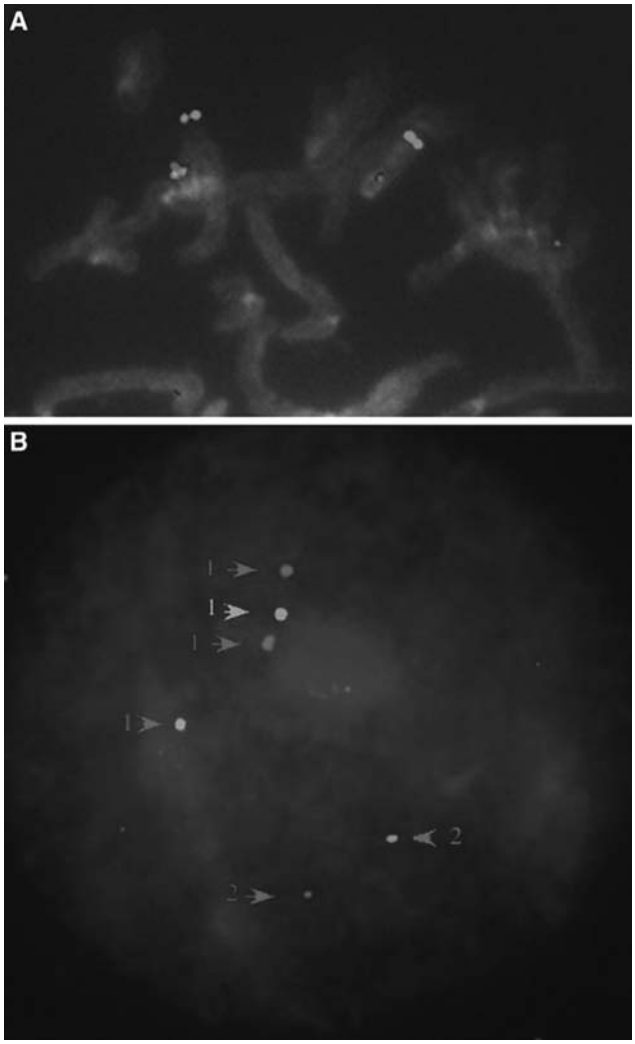
Fluorescence in situ hybridisation (FISH)

Based on FISH studies on both metaphase and interphase nuclei using FISH probes RP11-3018K1 and LSI-ARSA (both corresponding to the subtelomeric region of chromosome 22q), N25 (corresponding to the VCFS/ DGS region) and RP11-66F9 (corresponding to the CES region), it was concluded that the deletion of the CES region and the duplication of the VCFS/DGS region are localised on different homologues of chromosome 22 (Fig. 6).

Discussion

The complex rearrangement described here is, to our knowledge, the first report of a 'pure' deletion (e.g. not as a part of an unbalanced translocation) located in the CES region. The question is whether this rearrangement is related to a specific phenotype. Haplo-insufficiency of the CES region was found among five healthy relatives of the index patient. In addition, the family study indicates that the patient's father was an obligate carrier of the deletion of the CES region, as one of his brothers was a carrier of this deletion. A recent publication (Banting *et al*. 2005) showed that the vast majority of mice heterozygous for *CECR2* gene mutations were normal and capable of reproduction, whereas mice homozygous for mutations in the *CECR2* gene (correspondingly located on chromosome band 22q11.1 in the human genome) had a high penetrance of exencephaly. They established that CECR2 plays a role in neurulation during embryogenesis. These results suggest that, although the *CECR2* gene is essential during early development, a 50% decrease of gene dosage might not be associated with an aberrant phenotype. Although this data involves only one gene within the CES region, it is in accordance with our findings that a deletion in this region has no (obvious) phenotype and might therefore be present in the healthy population. In fact, the lack of clinical phenotype would explain the absence of reports on this deletion. Another reason for the lack of reported deletions of the CES region is that there is no commercial FISH probe available for this region, so it cannot be found 'coincidentally' as the duplication of the VCFS/DGS region was detected (Edelmann *et al*. 1999). Furthermore, ascertainment bias might account for this deletion. People with a mild phenotype will not be tested using high resolution-or whole genome techniques.

Figure 6. FISH analysis of chromosome 22.



A A partial metaphase of the patient, hybridised with the telomere probe RP11-3018K1 (*green*; chromosome region 22q13), N25 (*red*; VCFS/DGS region) and RP11-66F9 (*green*; CES region). On the right chromosome, green signals of RP11-3018K1 (telomeric side of chromosome 22) and a red signal N25 corresponding to the VCFS/DGS region are present; however, the signal of RP11-66F9 is lacking, indicating a deletion of the CES region. On the left chromosome, in addition to the green signals of RP11-3018K1, a red signal corresponding to the VCFS/DGS regions and a green signal corresponding to the CES region are both present. These latest two signals are partly overlapping. On this chromosome, the signal of N25 is stronger than the signal on the right chromosome, suggesting a duplication of the VCFS/DGS region. These findings are confirmed by the result of the interphase nucleus depicted in part b of this figure. B The different chromosomes 22 are marked 1 and 2. The signal of LSI ARSA, corresponding to the telomeric side of chromosome 22, is indicated with a *blue arrow*. The *red arrow* indicates the N25 signal (corresponding to the VCFS/DGS region), which is duplicated in chromosome 22 nr.1 (two red signals). The *green arrow* indicates the signal of RP11-66F9 (corresponding to the CES region). This signal is missing on chromosome 22 nr.2, demonstrating the deletion of the CES region. [See appendix: colour figures.]

Patients with duplication 22q11.2 syndrome show a wide variety of clinical features ranging from unaffected to severely affected individuals (Edelmann *et al*. 1999; Kriek *et al*. 2004; Yobb *et al*. 2005). Despite this, Ensenauer *et al*. (2003) described six clinical features that are frequent among 22q11.2 patients. Five of these (cognitive deficit, poor growth, dysmorphic features, cleft palate and hearing loss) were present in our index patient. The most characteristic dysmorphic features for the duplication 22q11.2 syndrome, however, such as superior placement of eyebrows, widely spaced eyes and downslanting of the eyes, were absent in our patient. Furthermore, our patient has several features (myotonic facial expression, proptosis of the eyes and a double set of teeth) that have not been described previously in other dup22q11.2 patients. Notably, the patients described by Ensenauer *et al*. (2003) show an ascertainment bias towards VCFS/DGS related features. All 653 patients included in this study were previously referred for 22q11 deletion screening using FISH on metaphase nuclei. In 2005, the clinical characteristics of another seven patients showing a duplication of the VCFS/DGS region were summarised (Yobb *et al*. 2005). This group of patient has a partial ascertainment bias for VCFS/DGS related features. Five were detected using FISH for 22q11 deletion screening, two were found by screening a cohort of 275 samples that was referred for fragile X screening. The clinical features of the latest two patients did not show similarity with VCFS/DGS spectrum. This last paper highlights the extreme variability of this alteration.

It is known that genetic factors localised outside the 22q11 region contribute to the variable clinical manifestations of 22q11 related alterations. It appeared that Fibroblast Growth Factor 8 (FGF8) mutant mice show close resemblance to the phenotype of del22q11.2 syndrome patients (Frank *et al*. 2002). Therefore, the *FGF8* gene, localised in the ectoderm and endoderm of the developing pharyngeal arches, might contribute to the 22q11 features. Stalmans *et al*. (2003) argued, based on mouse experiments, that the vascular endothelial growth factor gene (*VEGF* gene) modifies the expression of the VCFS/DGS syndrome, especially the cardiovascular birth defects. These, or other, as yet unidentified, modifiers localised outside the 22q11 region could also contribute to the phenotype of 22q11 duplication cases. Phenotypic variability due to the presence of a so far unknown modifier of a rearrangement might also play a role in to the phenotype of the patients with a deletion of the CES region.

In short, the clinical features described can, in theory, be caused by the unique combination of the three copy number changes on chromosome 22. However, as the deletion of the CES region probably has no clinical consequences, there is no previous MR-related literature regarding the deletion of 22q12.1 and the contribution of 22q11

rearrangements could be altered by other factors, it is not possible to determine the isolated influence of the different genetic imbalances.

To date, only a few cases with a duplication of the VCFS/DGS region have been described. It is probable that the majority of these duplications have not been detected so far due to a combination of phenotypic diversity (mentioned above) and the difficulty of diagnosis. A good example of the second argument is the clinical report described here; our patient was tested for a possible deletion in the VCFS/DGS region using FISH on the metaphase cells and the duplication present in the same region could not be seen, as two signals were overlapping. To overcome these problems, one has to focus on applying techniques in a routine diagnostic setting that are capable of detecting both duplications and deletions, within the same assay. In this way, it will be possible to increase the number of patients with a genetic diagnosis and, in parallel, learn more about possible causes of clinical features as we have demonstrated in this study of 22q11 rearrangement.

Recent initiatives such as those of the Sanger Institute (www.sanger.ac.uk/PostGenomics/decipher/) to create platforms for compiling molecular cytogenetic data from clinical genetic studies will hopefully provide a base for understanding the role of different DNA copy number alterations in genetic diseases. Collecting and understanding larger sets of data generated by different genomic approaches, as described here, will improve our ability to determine which copy number alterations contribute to abnormal phenotypes, and eventually result in a more consistent application of these techniques for genetic counseling.

References

Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE (2002) Recent segmental duplications in the human genome. *Science* **297**:1003–1007

Banting GS, Barak O, Ames TM, Burnham AC, Kardel MD, Cooch NS, Davidson CE, Godbout R, McDermid HE, Shiekhattar R (2005) CECR2, a protein involved in neurulation, forms a novel chromatin remodeling complex with SNF2L. *Hum Mol Genet* **14**:513–524

Dauwerse JG, Jumelet EA, Wessels JW, Saris JJ, Hagemeijer A, Beverstock GC, Van Ommen GJB, Breuning MH (1992) Extensive cross-homology between the long and short arm of chromosome 16 may explain leukemic inversions and translocations. *Blood* **79**:1299–1304

Edelmann L, Pandita RK, Spiteri E, Funke B, Goldberg R, Palanisamy N, Chaganti RS, Magenis E, Shprintzen RJ, Morrow BE (1999) A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum Mol Genet* **8**:1157–1167

Ensenauer RE, Adeyinka A, Flynn HC, Michels VV, Lindor NM, Dawson DB, Thorland EC, Lorentz CP, Goldstein JL, McDonald MT, Smith WE, Simon-Fayard E, Alexander AA, Kulharya AS, Ketterling RP, Clark RD, Jalal SM (2003) Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet* **73**:1027–1040

Flint J, Knight S (2003) The use of telomere probes to investigate submicroscopic rearrangements associated with mental retardation. *Curr Opin Genet Dev* **13**:310–316

Frank DU, Fotheringham LK, Brewer JA, Muglia LJ, Tristani-Firouzi M, Capecchi MR, Moon AM (2002) An Fgf8 mouse mutant phenocopies human 22q11 deletion syndrome. *Development* **129**:4591–4603

Graham A(2003) Development of the Pharyngeal Arches. *Am J Med Genet* **119A**:251–256

Jerome LA, Papaioannou VE (2001) DiGeorge syndrome phenotype in mice mutant for the T-Box gene, *TBX1*. *Nat Genet* **27**:286–291

Knijnenburg J, Szuhai K, Giltay J, Molenaar L, Sloos W, Poot M, Tanke HJ, Rosenberg C (2004) Insights from genomic micro-arrays into structural chromosome rearrangements. *Am J Med Genet* **132**(1):36–40

Kriek M, White SJ, Bouma MC, Dauwerse HG, Hansson KB, Nijhuis JV, Bakker B, van Ommen GJ, Den Dunnen JT, Breuning MH (2004) Genomic imbalances in mental retardation. *J Med Genet* **41**:249–255

Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ, Bradley A, Baldini A (2001) TBX1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* **410**:97–101

McDermid HE, Morrow BE (2002) Genomic disorders on 22q11. Am J Hum Genet 70:1077–1088

Merscher S, Funke B, Epstein JA, Heyer J, Puech A, Lu MM, Xavier RJ Demay MB, Russell RG, Factor S, Tokooya K, Jore BS, Lopez M, Pandita RK, Lia M, Carrion D, Xu H, Schorle H, Kobler JB, Scambler P, Wynshaw-Boris A, Skoultchi AI, Morrow BE, Kucherlapati R. (2001) TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* **104**:619–629

Robin NH, Shprintzen RJ (2005) Defining the clinical spectrum of deletion 22q11.2. *J Pediatr* **147**: 90–96

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

Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, Emanuel BS (2000) Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* **9**:489–501

Stalmans I, Lambrechts D, De Smet F, Jansen S, Wang J, Maity S, Kneer P *et al* (2003) VEGF: a modifier of the del22q11 (DiGeorge) syndrome? *Nat Med* **9**:173–182

White S, Kalf M, Liu Q, Villerius M, Engelsma D, Kriek M, Vollebregt E, Bakker B, van Ommen GJ, Breuning MH, Den Dunnen JT (2002) Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet* **71**:365–374

White SJ, Vink GR, Kriek M, Wuyts W, Schouten J, Bakker B, Breuning MH, Dunnen JT (2004) Two-color multiplex ligation-dependent probe amplification: detecting genomic rearrangements in hereditary multiple exostoses. *Hum Mutat* **24**:86–92

Woodfine K, Fiegler H, Beare DM, Collins JE, McCann OT, Young BD, Debernardi S, Mott R, Dunham I, Carter NP (2004) Replication timing of the human genome. *Hum Mol Genet* **13**:191–202

Yobb TM, Somerville MJ, Willatt L, Firth HV, Harrison K, Mackenzie J, Gallo N, Morrow BE, Shaffer LG, Babcock M, Chernos J, Bernier F, Sprysak K, Christiansen J, Haase S, Elyas B, Lilley M, Bamforth S, McDermid HE (2005) Microduplication and triplication of 22q11.2: a highly variable syndrome. *Am J Hum Genet* **76**:865–876

Chapter III-2

Peters Plus Syndrome Is Caused by Mutations in *B3GALTL*, a Putative Glycosyltransferase

Saskia A. J. Lesnik Oberstein[1], Marjolein Kriek[1], Stefan J. White[1], Margot E. Kalf[1], Karoly Szuhai[2], Johan T. den Dunnen[1], Martijn H. Breuning[1], and Raoul C. M. Hennekam[3]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
[2]Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, The Netherlands;
[3]Clinical and Molecular Genetics Unit, Institute of Child Health, London (R.C.M.H.); and Department of Pediatrics, Academic Medical Center, Amsterdam (R.C.M.H.)

Am J Hum Genet. 2006 Aug; 79(3):562-6.

Peters Plus syndrome is an autosomal recessive disorder characterized by anterior eye-chamber abnormalities, disproportionate short stature, and developmental delay. After detection of a microdeletion by array-based comparative genomic hybridization, we identified biallelic truncating mutations in the b1,3-galactosyltransferase–like gene (*B3GALTL*) in all 20 tested patients, showing that Peters Plus is a monogenic, primarily single-mutation syndrome. This finding is expected to put Peters Plus syndrome on the growing list of congenital malformation syndromes caused by glycosylation defects.

Peters Plus syndrome (MIM 261540) is an autosomal recessive disorder characterized by a variety of anterior eye-chamber defects, of which the Peters anomaly occurs most frequently.[1] Other major symptoms are a disproportionate short stature, developmental delay, characteristic craniofacial features, and cleft lip and/or palate.[1]

To detect potential microrearrangements affecting the disease locus, we performed genomewide 1-Mb resolution array-based comparative genomic hybridization[2] on genomic DNA of two brothers and four isolated patients who all received the clinical diagnosis of Peters Plus syndrome. In both brothers, two adjacent BAC clones (RP11-95N14 and RP11-37E23) were found to be present in a single copy, representing an ~1.5-Mb interstitial deletion on chromosome 13 (q12.3q13.1). MLPA (multiplex ligation-dependent probe amplification) analysis was used to confirm the deletion and to better define its extent. The deletion was confirmed in both brothers and their mother and spans six genes (*HSPH1, B3GALTL, LGR8, LOC196545, FRY,* and the first 13 exons of the *BRCA2* gene). Two of these, *LGR8* and *BRCA2,* are associated with human disease. Mutations in *LGR8* cause testicular maldescent[3]; since both brothers had cryptorchidism, this may be related to their LGR8 haploinsufficiency. *BRCA2* mutations are associated with hereditary breast and ovarian cancer, and large genomic rearrangements are known to contribute to ~2% of the *BRCA2* mutation spectrum.[4,5] The brothers' family history was positive for breast cancer in at least two deceased female relatives, in whom we established the presence of the deletion by interphase FISH on tumor material. Thus, this deletion constitutes a novel large *BRCA2* rearrangement associated with familial breast cancer.

Since none of the six genes was an obvious candidate gene for Peters Plus syndrome, we sequenced the genes' exons and flanking sequences in one of the affected brothers. A point mutation (c.1020+1G→A) was detected in the β1,3-galactosyltransferase–like gene (HUGO Gene Nomenclature Committee symbol *B3GALTL*) within the donor splice site of exon 8. The same mutation was also present in the other brother and as a single copy in the father. We subsequently performed targeted

sequencing analysis for the presence of the c.1020+1G→A mutation in an additional 18 patients with Peters Plus from 15 families. Fourteen patients were Dutch whites, and the other patients were Turkish, British, Arab, or Indian. All had the salient features of Peters Plus syndrome (table 1). We detected a homozygous c.1020+1G→A mutation in 16 of the 18 patients. In the remaining two patients (Dutch siblings), only a single c.1020+1G→A mutation was present (on the maternal allele). On sequencing the remainder of the gene, we detected a point mutation in intron 5 of *B3GALTL* (c.437+5G→A) on the paternal allele. Of the 11 available parent sets, all were heterozygous for the mutation detected in their affected offspring. We then excluded the presence of the c.1020+1G→A and c.437+5G→A mutations in 455 chromosomes of healthy Dutch individuals, by melting-curve analysis with specifically designed primer sequences (LightScanner HR96 [Idaho Technology]). Also, we investigated whether c.1020+1G→A could be a founder mutation, by analyzing known intragenic *B3GALTL* SNPs in 18 of the homozygous patients. Seven patients (Italian, Turkish, English, and four Dutch) showed heterozygosity for at least one of the three informative SNPs (*rs9315120, rs877103,* and *rs877104* [dbSNP]), which indicates that it is most likely a recurrent mutation, although some of the Dutch patients may have a common ancestor. The mutation is at the site of a potentially methylated CpG dinucleotide, which could explain its recurrence.[6]

A deleterious effect of the c.1020+1G→A mutation on transcription is certain, since it alters a donor splice site that is predicted to produce a skip of exon 8 and an out-of-frame mRNA product. We verified this by RT-PCR on patient material (fig. 1*D*). The c.437+5G→A mutation changes a highly conserved nucleotide and is predicted to affect splicing (Berkeley *Drosophila* Genome Project). To confirm this, we performed an RT-PCR on RNA isolated from lymphocytes from a patient with Peters Plus syndrome (c.1020+1G→A_{mat}/c.437+5G→A_{pat}). The patient's cDNA showed a skipped band, lacking exon 5, that results in an out-of-frame product. Notably, the expression of this band is much higher than that of the faint wild-type (WT) band, which is the product of the allele carrying the c.1020+1G→A mutation in exon 8 (fig. 1*E*). An explanation may be that the transcript lacking exon 8 is unstable. This theory is compatible with the fact that the individual who is heterozygous for the c.1020+1G→A mutation (fig. 1*D* [Het]), also shows a low expression of this product.

B3GALTL contains 15 exons and spans 132 kb of genomic DNA. It is transcribed in a wide range of human tissues (dbEST Web site), in the form of two transcripts (of 4.2 kb and 3.4 kb), and there is evidence of strong tissue or cell type-specific

Table 1. Clinical Characteristics of Individuals with Peters Plus Syndrome and Mutations of *B3GALTL*.

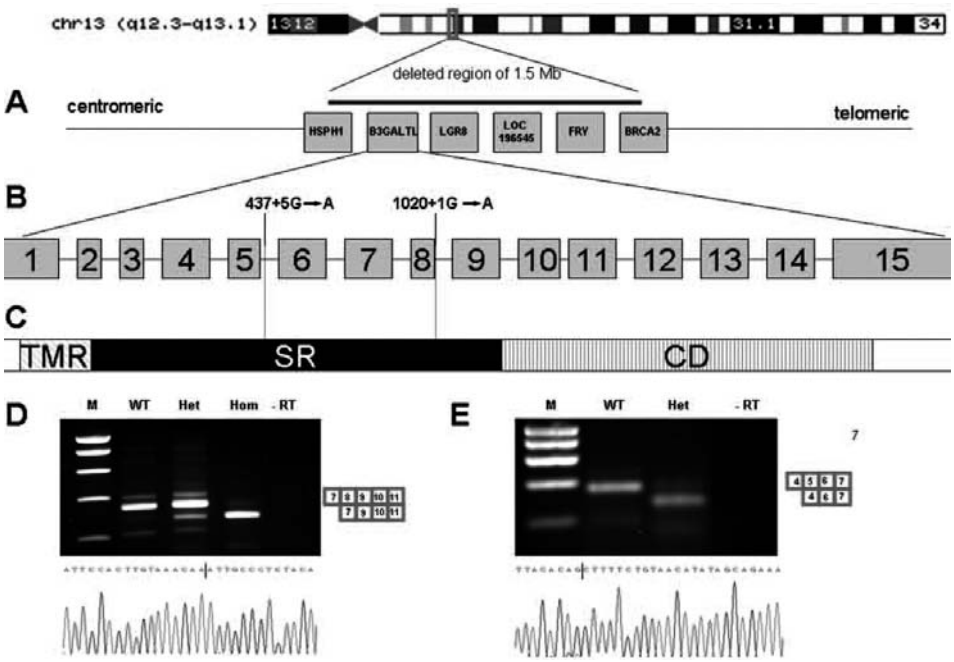
Individual	Sex	Peters Anomaly	Anterior Eye-Chamber Anomaly	Disproportionate Short Stature[a]	Cleft Lip and/or Palate	Developmental Delay	Heart Anomaly	Renal Anomaly	Ethnic Origin	Mutation
1100.1	F	–	+	+	–	+	–	+	Dutch	Homozygous 10201G→A
1100.2[b]	M	–	+	+	–	U	–	–	Dutch	Homozygous 10201G→A
1200.1	F	+	+	+	–	+	–	–	Dutch	Homozygous 10201G→A
1200.2	F	+	+	+	–	+	–	–	Dutch	Homozygous 10201G→A
1201.5	F	+	+	+	L	+	–	–	Dutch	$1020+1G\rightarrow A_{mat}/4375G\rightarrow A_{pat}$
1201.6	M	+	+	+	–	+	–	–	Dutch	$1020+1G\rightarrow A_{mat}/4375G\rightarrow A_{pat}$
1300.1	F	+	+	+	L/P	+	–	–	Dutch	Homozygous 1020+1G→A
1400.2	M	–	+	+	L/P	+	–	–	Dutch	Homozygous 1020+1G→A
1500.1	M	+	+	+	BL/P	+	–	–	Turkish	Homozygous 1020+1G→A
1600.1	M	+	+	+	P	+	+	–	Dutch	$1020+1G\rightarrow A_{pat}/del_{mat}$
1600.2	M	U	+	+	L/P	+	+	–	Dutch	$1020+1G\rightarrow A_{pat}/del_{mat}$
1700.1	F	–	+	+	BL/P	+	+	–	Dutch	Homozygous 1020+1G→A
1800.1	M	+	+	+	–	+	–	–	Dutch	Homozygous 1020+1G→A
1900.1	F	+	+	+	–	–	–	–	Dutch	Homozygous 1020+1G→A
1900.2	M	+	+	+	–	–	–	–	Dutch	Homozygous 1020+1G→A
2000.1	F	+	+	+	L	+	+	–	Dutch	Homozygous 1020+1G→A
2100.1	M	+	+	+	–	–	–	–	Dutch	Homozygous 1020+1G→A
2200.1	M	+	+	+	BL/P	+	–	+	English	Homozygous 1020+1G→A
2400.1	F	+	+	+	–	–	+	–	Arab	Homozygous 1020+1G→A
2500.1	M	+	+	+	–	+	U	U	Indian	Homozygous 1020+1G→A

NOTE.—L p cleft lip; P p cleft palate; L/P p unilateral cleft lip and palate; BL/P p bilateral cleft lip and palate; U p unknown.

a <3rd Percentile.

b Deceased in neonatal period.

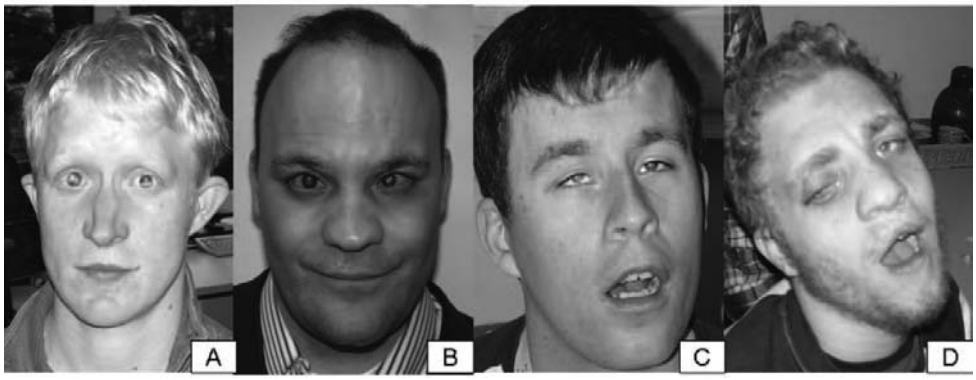
Figure 1. Overview of the location of the mutations in the *B3GALTL* gene and the results of the RT-PCR of RNA isolated from fibroblasts.



A, Genes present in the 1.5-Mb deletion found in two brothers with Peters Plus syndrome. *B,* 15 exons of the *B3GALTL* gene, with the localization of the mutations. *C,* B3GALTL protein, which consists of a transmembrane region (TMR), a stem region (SR), and a catalytic domain (CD). Both mutations (c.1020 1GrA and c.437 5GrA) are located in the stem region. *D,* Result of the nested RT-PCR of exons 7–11 of the *BGALTL* gene, with RNA derived from myoblasts (WT), RNA from fibroblasts of a father heterozygous for the c.1020 1GrA mutation (Het), and RNA from fibroblasts of his affected son with c.1020 1GrApat/delmat (Hom). The patient shows a smaller band compared with the WT band, which indicates a skip of exon 8. Sequence analysis of this band is shown. The vertical line indicates the end of exon 7 and the beginning of exon 9. The RT-PCR of the father shows, in addition to the WT band, a skipped product with much less intensity. *E,* Result of the RT-PCR encompassing exons 4–7 of the *BGALTL* gene, with RNA derived from lymphocytes of a control individual (WT) and a patient with a c.1020 1GrAmat/c.437 5GrApat genotype (Het). In addition to a faint WT band, the patient shows a smaller product that lacks exon 5. The sequence analysis of this smaller band confirms the skip of exon 5. [See appendix: colour figures.]

regulation.[7] Transcription has been shown to terminate at three different alternative polyA-addition sites, all in exon 15.[7] The B3GALTL protein spans 498 aa and contains a short N-terminal tail, a trans-membrane region (aa 5–28), a so-called stem region (aa 29–260), and a C-terminal catalytic domain (aa 261–498).[7] On the basis of the sequence of its catalytic domain, the protein most closely resembles proteins from the GT31 family of beta-3 glycosyltransferases (CAZy [CarbohydrateActive enZymes Web site]). Both the c.1020+1G→A and the c.437+5G→A mutations in

Figure 2. Facial features of four patients with Peters Plus syndrome.



Patients A and C are homozygous for the c.1020+1G→A mutation. Patient B has the c.1020+1G→A_{mat}/c.437+5G→A_{pat} genotype, and patient D has the c.1020+1G→A_{pat}/del_{mat} genotype. Note the Peters anomaly of the eyes, the long face, and the Cupid's bow shape of the upper lip in all patients. Patients B and D have a repaired cleft lip and/or palate. Patient A is female; the rest are male. [See appendix: colour figures.]

B3GALTL are predicted to lead to a truncated product lacking the catalytic domain, since they are located in the putative stem region of the protein (fig. 1*C*).[7] Thus, since all patients we analyzed have homozygous severely truncating mutations, it is expected that they have, effectively, full knockout mutations and lack any significant B3GALTL activity. Given this genetic homogeneity, there is a strikingly variable cognitive phenotype. Even within the group homozygous for the c.1020+1G→A mutation, patients range from having normal secondary education to severe cognitive impairment, which suggests that other factors modulate the phenotype. The brothers with the deletion of one of their alleles (c.1020+1G→A_{pat}/del_{mat}) have severe cognitive impairment that is within the range of Peters Plus syndrome, and they have no structural malformations outside the Peters Plus spectrum. This indicates that hemizygosity for the genes *HSPH1, LOC196545,* and *FRY,* which have hitherto not been associated with human congenital malformations, did not produce a detectable phenotype. Figure 2 illustrates the facial phenotypes of four patients with Peters Plus syndrome.

B3GALTL is a putative glycosyltransferase that has not been previously associated with human disease or congenital malformations but has recently been shown to be over-expressed in thyroid oncocytic tumors.[8] So far, we have not been able to verify a glycosylation defect in patients with Peters Plus syndrome; serum transferrin isoelectric-focusing studies in six of the current patients had normal results. We also

studied profiles of enzymatically released N-glycans by matrix-assisted laser-desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) and high-pH anion-exchange chromatography (HPAEC) with electrochemical detection. No obvious differences in overall N-glycosylation of serum proteins were observed (results not shown). However, these results do not exclude a glycosylation defect,[9] and we are initiating further (functional) studies.

There are several hundred glycosyltransferases, predicted to be active in humans, that are involved in the posttranslational modification of proteins by the addition of specific oligosaccharide side chains (glycans), to form glycoproteins. Congenital disorders of glycosylation are due to defects in the synthesis of the glycan moiety of glycoproteins or other glycoconjugates.[10] Mutations in a number of glycosyltransferases have been associated with congenital malformation syndromes.[10] Pending confirmation of the glycosylation defect, Peters Plus syndrome can most likely be added to this growing list. Anterior eye-chamber defects, such as Peters eye anomaly and glaucoma, are also described in Walker-Warburg syndrome and muscle-eyebrain disease,[10,11] which suggests that adequate glycosylation plays a critical role in the formation of the anterior eye chamber.[11,12] Interestingly, at least one Peters Plus- affected family in the present study has a documented history of glaucoma in confirmed mutation carriers. This raises the question of whether haploinsufficiency of – and possibly variations in – *B3GALTL* increases glaucoma susceptibility, which warrants further research. Finally, the present study emphasizes the value of genomewide array analysis in establishing the genetic basis of autosomal recessive disorders.

Acknowledgments

We thank the patients and their families for their generous cooperation, and we thank the following clinicians for referral of patients: J. van der Smagt (The Netherlands), I. C. Verma (India), L. Basel-Vanagaite (Israel), D. Bartholdi (Switzerland), and L. Wilson (United Kingdom). We also thank H. C. Hokke and A. M. Deelder (Biomolecular Mass Spectrometry Unit, Leiden), for glycosylation analysis; B. J. Poorthuis, for performing isoelectric-focusing studies; J. Knijnenburg and R. Vossen (Leiden Genome Technology Centre), for technical assistance; and A. Aartsma-Rus, for expert advice regarding the RT-PCR.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Berkeley *Drosophila* Genome Project, http://www.fruitfly.org/seq _tools/splice.html (for the Splice Site Prediction by Neural Network)

Carbohydrate-Active enZymes (CAZy), http://194.214.212.50/ CAZY/fam/GT31.html

dbEST, http://www.ncbi.nlm.nih.gov/dbEST/ (for the Expressed Sequence Tags database)

dbSNP, http://www.ncbi.nlm.nih.gov/SNP/ (for SNP identification numbers *rs9315120, rs877103,* and *rs877104*)

HUGO Gene Nomenclature Committee, http://www.gene.ucl.ac .uk/nomenclature/ (for *B3GALTL*)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi .nlm.nih.gov/Omim/ (for Peters Plus syndrome)

References

1. Wenniger-Prick LJJM, Hennekam RCM (2002) The Peters' plus syndrome: a review. *Ann Genet* **45**:97–103.
2. Knijnenburg J, Szuhai K, Giltay J, Molenaar L, Sloos W, Poot M, Tanke HJ, Rosenberg C (2005) Insights from genomic microarrays into structural chromosome rearrangements. *Am J Med Genet* A **132**:36–40.
3. Ferlin A, Simonato M, Bartoloni L, Rizzo G, Bettella A, Dottorini T, Dallapiccola B, Foresta C (2003) The INSL3-LGR8/ GREAT ligand-receptor pair in human cryptorchidism. *J Clin Endocrinol Metab* **88**:4273–4279.
4. Tournier I, Paillerets BB, Sobol H, Stoppa-Lyonnet D, Lidereau R, Barrois M, Mazoyer S, Coulet F, Hardouin A, Chompret A, Lortholary A, Chappuis P, Bourdon V, Bonadona V, Maugard C, Gilbert B, Nogues C, Frebourg T, Tosi M (2004) Significant contribution of germline *BRCA2* rearrangements in male breast cancer families. *Cancer Res* **64**:8143–8147.
5. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, Roach KC, Mandell J, Lee MK, Ciernikova S, Foretova L, Soucek P, King MC (2006) Spectrum of mutations in *BRCA1, BRCA2, CHEK2,* and *TP53* in families at high risk of breast cancer. *JAMA* **295**:1379–1388.
6. Zhao Z, Zhang F (2006) Sequence context analysis of 8.2 million single nucleotide polymorphisms in the human genome. *Gene* **366**:316–324.
7. Heinonen TYK, Pasternack L, Lindfors K, Breton C, Gastinel LN, Maki M, Kainulainen H (2003) A novel human glycosyltransferase: primary structure and characterization of the gene and transcripts. *Biochem Biophys Res Commun* **309**: 166–174.
8. Jacques C, Baris O, Prunier-Mirebeau D, Savagner F, Rodien P, Rohmer V, Franc B, Guyetant S, Malthiery Y, Reynier P (2005) Two-step differential expression analysis reveals a new set of genes involved in thyroid oncocytic tumors. *J Clin Endocrinol Metab* **90**:2314–2320.
9. Freeze HH (2006) Genetic defects in the human glycome. *Nat Rev Genet* **7**:537–551.
10. Jaeken J (2003) Komrower lecture: congenital disorders of glycosylation (CDG): it's all in it! *J Inherit Metab Dis* **26**:99– 118.

11. van Reeuwijk J, Janssen M, van den EC, Beltran-Valero de Bernabe D, Sabatelli P, Merlini L, Boon M, Scheffer H, Brockington M, Muntoni F, Huynen MA, Verrips A, Walsh CA, Barth PG, Brunner HG, van Bokhoven H (2005) *POMT2* mutations cause a-dystroglycan hypoglycosylation and Walker-Warburg syndrome. *J Med Genet* **42**:907–912.
12. Diskin S, Kumar J, Cao Z, Schuman JS, Gilmartin T, Head SR, Panjwani N (2006) Detection of differentially expressed glycogenes in trabecular meshwork of eyes with primary open-angle glaucoma. *Invest Ophthalmol Vis Sci* **47**:1491–1499.

Erratum

In the September 2006 issue of the *Journal,* in the article entitled "Peters Plus Syndrome Is Caused by Mutations in *B3GALTL,* a Putative Glycosyltransferase" by Lesnik Oberstein et al. (79:562–566), because of the use of an incorrect reference sequence, the annotation of the mutations in the article is incorrect. On the basis of a coding DNA reference sequence (GenBank accession number NM_194318.2), the exon 5 splice-site mutation should be described as c.347+5G→A (not c.437+5G→A), and the exon 8 splice-site mutation as c.660+1G→A (not c.1020+1G→A). All variations identified in the *B3GALTL* gene have been collected in a new locus-specific sequence-variation database. The database has been registered at the Human Genome Variation Society and can be found at http:// chromium.liacs.nl/lovd/search.php?select_dbpB3GALTL. The authors regret the errors.

Chapter III-3

Telomeric deletions of 16p causing alpha-thalassemia and mental retardation characterized by multiplex ligation-dependent probe amplification

Cornelis L.Harteveld[1], Marjolein Kriek[1], Emilia K. Bijlsma[1], Zoran Erjavec[2], Deepak Balak[1], Marion Phylipsen[1], Astrid Voskamp[1], Emmanora di Capua[1], Stefan White[1] and Piero C. Giordano[1]

[1]Center of Human and Clinical Genetics, Leiden University Medical Center, The Netherlands
[2]Delfzicht Ziekenhuis, Delfzijl, The Netherlands

Human Genet. 2007 Jun 28; [Epub ahead of print]

Abstract

Alpha thalassemia retardation associated with chromosome16 (ATR-16 syndrome) is defined as a contiguous gene syndrome resulting from haploinsufficiency of the α-globin gene cluster and genes involved in mental retardation (MR). To date, only few cases have been described which result from pure monosomy for a deletion of 16p. In most of these cases the deletion was identified by densitometric analysis of Southern blot results or by Fluorescent In Situ Hybridization analysis, and these alterations have not been mapped in detail. In this study, we have fine mapped deletions causing α-thalassemia within 2 Mb from the telomere of 16p by multiplex ligation-dependent probe amplification (MLPA). We have developed a rapid and simple test for high resolution mapping of rearrangements involving the tip of the short arm of chromosome 16 by incorporating 62 MLPA probes spaced approximately 10–200 kb over a region of 2 Mb from the telomere. One deletion of approximately 900 kb without MR was identified in addition to three *de novo* deletions varying between 1.5 and 2 Mb causing ATR-16 in three patients having mild MR and α-thalassemia. Two were found by chance to be ATR-16 because they were included in a study to search for telomeric loss in MR and not by hematological analysis. This would plead for more alertness when a persistent microcytic hypochromic anemia at normal ferritin levels is observed as suggestive for the ATR-16 syndrome. The region on chromosome 16p for which haploinsufficiency leads to the dysmorphic features and MR typical for ATR-16, has been narrowed down to a 800 kb region localized between 0.9 and 1.7 Mb from the telomere.

Introduction

Genomic deletions involving the α-globin gene cluster on chromosome 16p13.3 are the most common molecular cause of α-thalassemia (approximately 80–90% of cases) (Bernini and Harteveld 1998; Higgs 1993). Due to selective advantage, α-thalassemia carrier frequencies are high in areas endemic for malaria tropica. Less frequently α-thalassemia is found in North-European Caucasians. Even more rare are mental retardation (MR) syndromes in which the occurrence of α-thalassemia is merely a marker of the genomic defect. These syndromes are known as α-thalassemia MR syndromes ATR-X and ATR-16 (depending on the respective chromosomes involved) (Weatherall *et al.* 1981; Wilkie *et al.* 1990a, b).

ATR-X involves mutations of the X-linked *ATRX* gene, coding for helicase-2, a putative global transcriptional regulator (Cardoso *et al.* 2000, 1998; Gibbons and Higgs 2000; Gibbons *et al.* 1992; Villard *et al.* 1997; Wilkie *et al.* 1991; Wilkie 1990b; Yntema *et al.* 2002). ATR-16 is characterized by the occurrence of large deletions involving the α-globin gene cluster on chromosome 16p and is most likely a contiguous gene syndrome (Daniels *et al.* 2001; Flint *et al.* 1996; Horsley *et al.* 2001; Lamb *et al.* 1993; Wilkie *et al.* 1991). At present molecular tests commonly used to identify deletion types of α-thalassemia and ATR-16 are gap-PCR, Southern blot or fluorescent in situ hybridization (FISH) analysis (Daniels *et al.* 2001; Flint 1996; Gallego *et al.* 2005; Lindor *et al.* 1997). However, the applicability of these techniques is limited to known deletions, may involve radio-activity, is dependent upon the hybridization probes available and may require time consuming and laborious cell culture to generate metaphase chromosome spreads.

Recently, we have developed an multiplex ligation-dependent probe amplification (MLPA) based assay to perform high resolution screening for unknown rearrangements on chromosome 11p15.4 and in a 700 kb telomeric region of the short arm of chromosome 16 (16p13.3) causing β-and α-thalassemia, respectively. During the examination of 38 putative α-thalassemia carriers, we identified a single patient showing a telomeric deletion without MR, for which the 3¢breakpoint could not be determined (- $-^{GZ}$) (Harteveld *et al.* 2005). During this study we extended the MLPA assay to investigate a region of approximately 2 Mb involved in α-thalassemia and MR. A second patient was brought to our attention because of a persistent microcytic hypochromic anemia without iron depletion. The patient showed MR and the parents were normal. Two patients were detected using multiplex amplifiable probe hybridization (MAPH) for the screening of genomic imbalances in the subtelomeric region among

mentally retarded patients. Only the telomeric probe associated with the telomere of 16p appeared to be deleted, excluding unbalanced translocation in these patients. The results of this screening were verified by MLPA and FISH analysis. One of these patients has been previously described (P.V.) (Kriek *et al.* 2004).

Materials and methods

Appropriate informed consent was obtained from all human subjects studied.

Clinical report

Case G.Z.

Shortly after birth the patient had surgery for pylorus stenosis. He was regularly seen until the age of 4 because of persisting microcytic hypochromic anemia with normal iron levels. There was no developmental delay nor any other abnormalities related to the ATR-16 syndrome. The propositus was investigated for the first time at the age of 5 at the hematological and biochemical level together with his parents because of a suspected α-thalassemia. The father presented with normal hematological parameters, the mother and the patient both showed hematological abnormalities and an unbalanced hemoglobin chain synthesis ratio typical of α^0-thalassemia carriership. At that time DNA analysis was not feasible. The patient was re-investigated in 2002 at the age of 30, when microcytic hypochromic anemia at normal ferritin was still present. The seven most common α-thalassemia deletions ($-\alpha^{3.7}$,$-\alpha^{4.2}$,$-(\alpha)^{20.5}$,$--^{MedI}$,$--^{SEA}$, $--^{THAI}$,$--^{FIL}$) were excluded at the molecular level by multiplex PCR (Chong *et al.* 2000; Liu *et al.* 2000). Both mother and son had a normal school education and there was no indication for MR in these family members.

Case H.N.

H.N. is the third child of healthy, non-consanguineous parents. He was born after an uncomplicated pregnancy and delivery, with a birth weight of 3.1 kg. Directly after birth a short period of asphyxia was recorded, and was attributed to meconium-stained amniotic fluid (Apgar scores 3/6/9 after resp. 1, 5 and 10 min; pH cordblood 7.13). He made a quick recovery with an oxygen mask. He had a clubfoot on the left, for which he was operated on at the age of 1 year (lengthening of achilles tendon).

In infancy, he suffered from recurrent chest infections and asthma. His motor development was delayed: sitting at the age of 10 months, crawling at the age of 18 months

and walking at the age of 23 months. He spoke his first words at the age of 18 months. At the age of 2 years and 6 months he was referred to the pediatric neurologist because of hypotonia, psychomotor retardation and speech delay. At the age of 3 years and 8 months a severe delay in active language ability was reported. Subsequent testing showed mild MR (SON-IQ 58).

We first examined him at the age of 4 years and 8 months. He was able to construct short sentences (three words) and went to a special school with an individual teaching program. By that time, mild anemia had become apparent. Physical examination showed the following: height 104 cm (–1 SDS), weight 17.5 kg (+1 SDS), head circumference 53 cm (0 SDS). He had nasal speech. His facial features showed downslanted palpebral fissures, mild hypertelorism, a broad nasal tip, small posteriorly rotated ears, a short neck with webbing, and a low trident posterior hairline. Apart from pectus carinatum, an operated clubfoot on the left, and a flat foot on the right, no other anomalies were noted (Fig. 1a).

Additional investigations

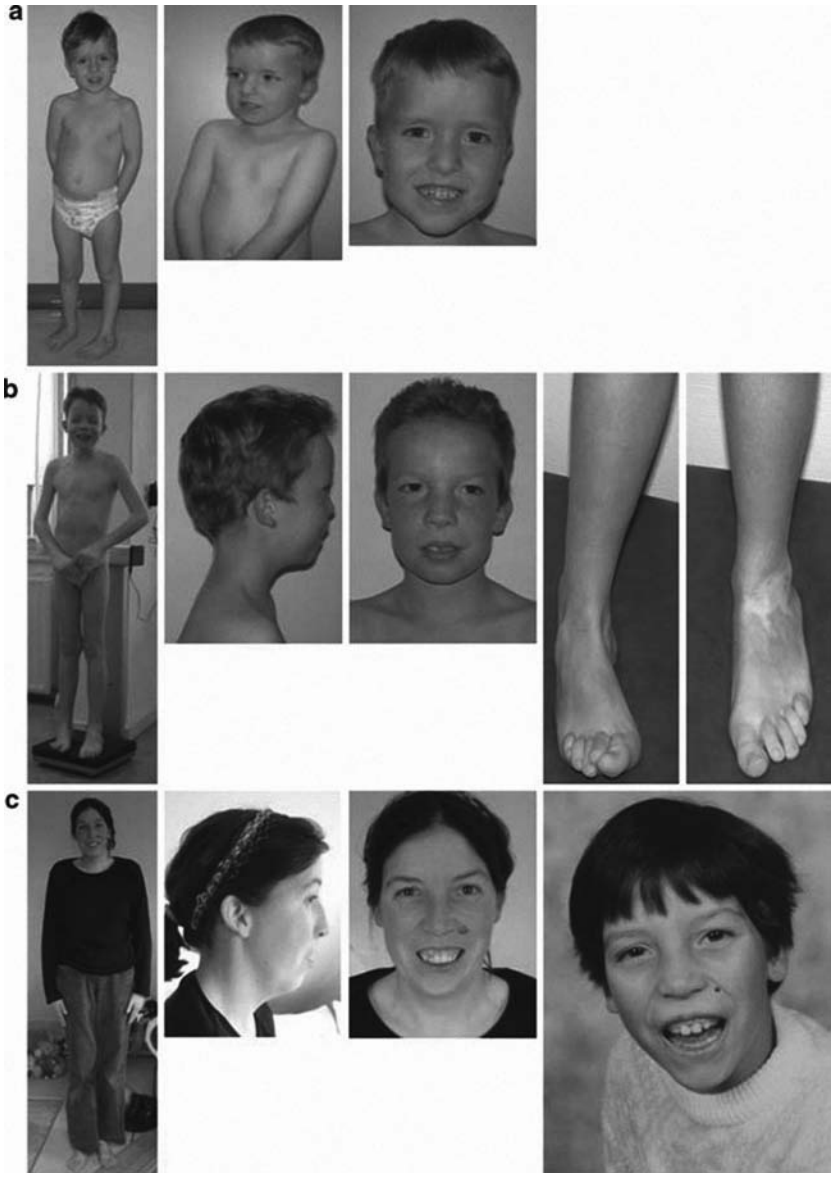
Conventional chromosome analysis showed a normal male karyotype. Metabolic screen was negative. MRI of the brain showed an arachnoidal cyst in the right temporal lobe. As some features were consistent with Noonan syndrome, *PTPN11* mutation analysis was performed, which resulted negative.

Case P.V.

This male patient was born after an uneventful pregnancy and delivery. He is the youngest of a family of three children to non-consanguineous parents. His sister died 3 days postpartum due to severe complications at delivery resulting in asphyxia. There was no family history of either developmental delay or congenital malformations. In early childhood, the patient had several episodes of pneumonia, and was diagnosed with asthma at a later stage. Neuropsychological testing was performed at the age of 2, because of developmental delay and because his overall behavior was far from consistent with his age. The patient started walking at the age of 30 months. He started to actively use language at the age of 5. He suffered from recurrent epileptic seizures at the age of seven that were treated successfully with Depakine. In addition, mild anemia was detected.

A physical examination at age 11, revealed a very outgoing boy. His height was 144.5 cm (–1 SDS), his weight 29 kg (–2 SDS) and head circumference 52.3 cm (–1 SDS). He showed mild dysmorphic features including high forehead, some periorbital

Figure 1. Three unrelated patients.



H.N. (a), P.V. (b) and F.T. (c) showing a mild mental retardation (IQ 50–60), a severe delay in active language ability, some typical facial features like downslanted palpebral fissures, mild hypertelorism, a broad nasal tip and small ears and a short neck with webbing, which is most pronounced in a and b. Patient H.N. and P.V. both show pectus carinatum. This was also observed for patient F.T. (not shown). H.N. also has an operated clubfoot on the left, while patient P.V.'s right foot is turned inside, the other foot showing a cafe´-au-lait spot. c Patient F.T. has a short neck and small ears. On the outer right a photograph is shown of the patient at age 11. The karyotype was normal in all patients and hematological analysis showed a persistent microcytic hypochromic anemia without iron deficiency [See appendix: colour figures.]

fullness, microphthalmia, telecanthus, broad nasal bridge, posterior rotated ears, a flat, rather long philtrum, full lips and micrognatia (Fig. 1b.). He had an extra incisor. His trunk showed a mild pectus excavatum and two café au lait spots. Auscultation of the heart was normal. Hyperlaxity of the joints was observed.
Conventional chromosome analysis at a resolution of 500 bands showed a normal male karyotype. Due to his behavior and some of the facial dysmorphism, FISH analysis for the Williams Syndrome Critical region was performed. No deletion was detected on chromosome band 7q11.23. No metabolic abnormalities were observed.

Case F.T.
F.T. was referred at the age of 30 years because of mild and persistent microcytic hypochromic anemia. She was born at term after an uneventful pregnancy. Her birth weight was 2,600 g. APGAR score was reported as 'low' after 1 min but recovered to ten after 5 min (no direct data, home delivery). In infancy, it was noted that her development was retarded in comparison to other children (e.g. first words after 2 years of age), but the parents declined further investigation at the time. Her behavior was reported as shy and dependant. She attended special education at the age of 6 and she is now employed in a program for people with a mental handicap. She lives in a support home. Furthermore, she plays the drums and enjoys horseback riding. She has no problems with her general health. Hearing and vision were normal. Family history was unremarkable, she had one healthy sister.

Physical examination at the age of 31 showed a height of 162 cm (–1.3 SD), head circumference 52.5 cm (–1.7 SD), elongated face with a flat midface and a prominent nasal bridge. The palpebral fissures of her left eye showed slight upslant. Her ears were rather small, and had a slight question mark configuration. In addition, she had some irregularity of teeth, marked micrognathia and retrognatia (Fig. 1c). On her palatum durum, two small bulbous lumps were present and she had hypertrophic gums. Her neck was short. A pectus excavatum was present. Internal screen was normal. Her hands are rather short with slight tapering of the fingers and bilateral fifth finger clinodactyly. On the feet, both halluces showed laterial deviation, the second digits were short and the fourth digits showed medial deviation. Her joint were rather stiff. She did not have any skin pigment aberrations. She has scarce body hair with normal scalp hair.

Hematological analysis
Blood samples of patients and parents were collected in vacutainers with EDTA as anticoagulant. Hematological analysis was carried out according to standard methods

(Dacie and Lewis 1991). The red cell indices were measured with a standard cell counter (Micros 60; ABX Diagnostics, Montpellier, France). A Brilliant Cresyl Blue staining was performed on the blood smear to identify HbH inclusion bodies according to a standard method (Dacie and Lewis 1991). Globin chain synthesis was performed for patient GZ and his parents according to standard procedures (Giordano *et al.* 1999). DNA was isolated according to standard procedures (Miller *et al.* 1988).

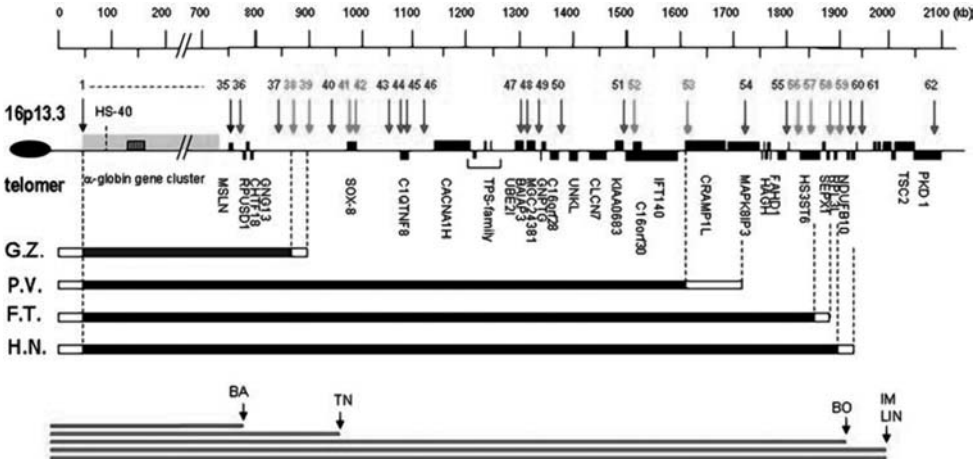
Design of the MLPA assay

In total, 62 MLPA probe pairs were designed covering a region of 2 Mb from the telomer of 16p13.3 to the *PKD*-gene. Of these, 35 were previously reported to detect (unknown) α-thalassemia deletions and rearrangements in a 700 kb region from the telomer of 16p to the *MSLN* gene (Harteveld *et al.* 2005). An additional 17 probe pairs were designed to extend the region covered by MLPA from the *MSLN* gene towards the *PKD* gene to screen for even larger rearrangements (Table 1, Fig. 2). Each primer pair contained common ends corresponding to either a HEX-labeled amplification primer (detection in green), a FAM labeled primer (detection in blue) or a ROX labeled primer (detection in red) to be analyzed simultaneously in the same fragment run on the ABI3730. Finally, ten primer pairs were designed for high resolution mapping of the breakpoints after initial screening for large deletions in three colors. The data were collected and ratios between normalized peak heights of the patient and the normal controls were presented in a single scatter plot for each patient.

The oligonucleotides were ordered from Illumina, Inc. (San Diego, CA, USA), synthesized in a salt-free environment (50 nmol scale) and used without further purification. For each probe pair the downstream primer was 5'phosphorylated to allow ligation.

The MLPA reactions were performed as described by Schouten *et al.* (2002) and White *et al.* (2004) in brief, approximately 200 ng of genomic DNA in a final volume of 5 µl was heated for 5 min at 98°C. After cooling to room temperature, 1.5 µl of the probe mix and 1.5 µl SALSA hybridization buffer (MRC-Holland, Amsterdam, The Netherlands) were added to each sample, followed by heat denaturation (2 min at 95°C), hybridization (16 h at 60°C). Ligation was performed by adding 32 µl of ligation mix at 54°C for 10 min, the reaction was terminated by 5 min incubation at 95°C. PCR amplification was carried out for 33 cycles in a final volume of 25 µl, adding the 5'ROX-labeled M13-Forward and M13-Reverse primer set to a final concentration of 100 nM. The second common primer set designed to fine map the deletion breakpoint region between two MLPA probes of the first set, were amplified by

Figure 2 Schematic presentation of short arm of chromosome 16 (16p13.3), showing a 2 Mb region from the telomere containing the α-globin gene cluster up to the *TSC* and *PKD* genes.



The *arrows* and *numbers* represent the location of the probes. The deletions found by MLPA are shown as *bars* below the figure. Large deletions previously described are indicated as *red bars* [See appendix: colour figures.]

adding the 5'HEX-labeled MAPH-Forward and MAPH-Reverse primers to a final concentration of 100 nM. Products were separated by capillary electrophoresis on the ABI 3730 (Applied Biosystems) and data analyzed as described previously (Harteveld *et al*. 2005).

Results

Hematological analysis

All patients presented with a microcytic hypochromic anemia without iron deficiency (Table 2). Two were brought to our attention because of a suspected α-thalassemia (G.Z. and F.T.) after routine hematological investigation. The other two (H.N. and P.V.) were investigated at the hematological level after identifying the telomeric loss of 16p as the only chromosomal abnormality causative for the observed MR. The patient G.Z. without MR and his parents were investigated at the hematological and biochemical level. He and his mother showed microcytic hypochromic parameters and an unbalanced α/β-globin chain synthesis ratio indicative for α^0-thalassemia carrier-ship

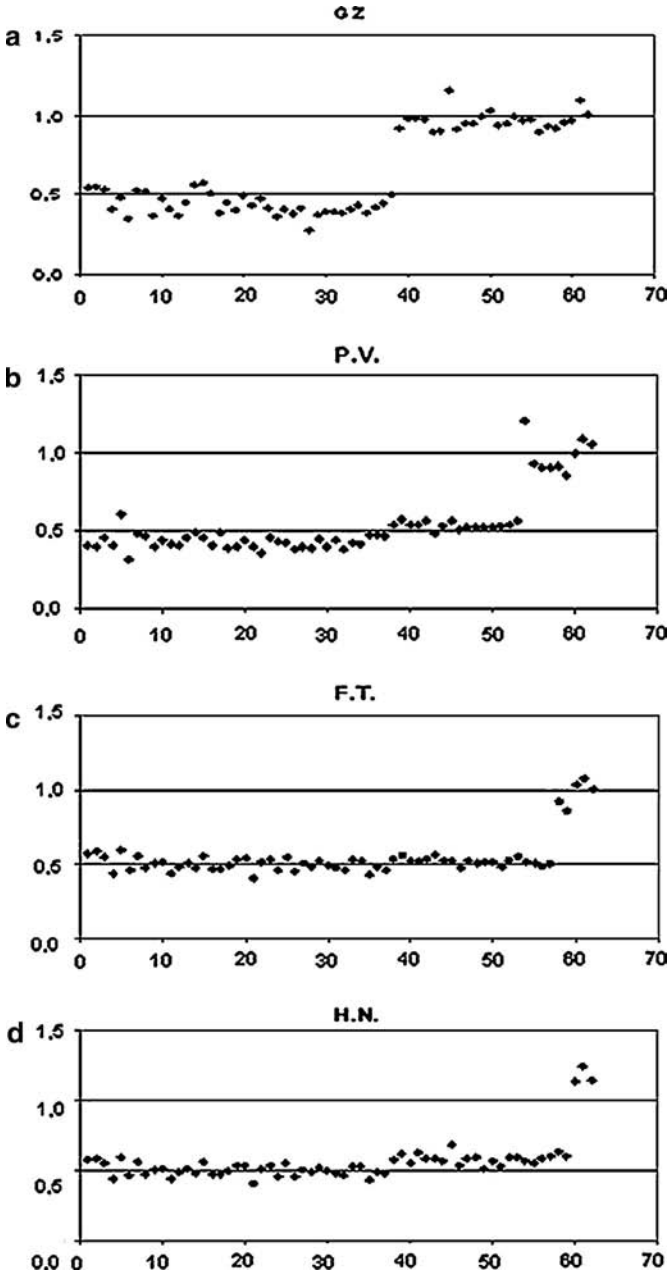
Table 1. Names and sequences of probes used for MLPA located between 750 kb and 2 Mb from the telomere of chromosome 16p.

No.	Name	Upstream hybridising sequence	Downstream hybridising sequence	Positions[a]
36	hs335h7	**GGCGATTAAGTTGGGTAAC**GAAGAGAGC TAGGAAAGGTCTGGGTGGCC	AGACCTGTCCCTTGGAAGCCCCAC T**GCTGTTTCCTGTGTGAAC**	774,902–774,930
37	hs.58362	**GGCGATTAAGTTGGGTAAC**GAAGGGCAAA CGTTGCTGAGCCG	CCGAGGGGCTGGGTCTTCGCCTTTATTTC **GCTGTTTCCTGTGTGAAC**	844,471–844,493
38	FLJ12681(2)	**GGCCGCGGGAATTCGATT**GAAGGGCACCA GGAGCTCGATGAAGTGGTTGCTGAGC	GTCTCGAAGCGATGGAACCACCAGGGTGAG T**CACTAGTGAATTCGCGGC**	869,631–869,698
39	FLJ12681(1)	**GGCCGCGGGAATTCGATT**CACCACGCAGA AGAGCTCCACTTACTGCTC	CAAGCATGATCCTGAAGATCAGCCACCGGA **CACTAGTGAATTCGCGGC**	900,907–900,966
40	hs.58362'	**GGCGATTAAGTTGGGTAAC**CTGGCTGCATCTCGGCCT GGCAGGCTCCCCTCACTTCCATACTCATTTGAGAGCCT	CACTTGAGGTTCGTAGCTCCTGAC**GCTGTT TCCTGTGTGAAC**	959,724–959,779
41	Sox-8	**GGCCGCGGGAATTCGATT**GCTCAAAGCC AAGCCGCATGT	GAAGCGGCCCATGAACGCATTCATG **CACTAGTGA ATTCGCGGC**	972,208–972,253
42	Sox-8(2)	**GGCCGCGGGAATTCGATT**GGGCCTCAG TTCTAGACGAGTCAT	ACCTGATTCACCTGCACTGCTTCCCC **CACTAGTGAA TTCGCGGC**	976,591–976,640
43	hs394h11	**GGCGATTAAGTTGGGTAAC**GTCCTTTGAC GAGGAGTTTGCGCCTC	ATTCAAGTCCCCTACCTGCATCCCTGGCG**GGCTGTT TCCTGTGTGAAC**	1,053,683–1,053,708
44	hs349e11'	**GGCGATTAAGTTGGGTAAC**CAATTTGCTTAA CGTGATTCCCGGCCAAGCTAAACATGACT	AATCGTGTTTACCCGGTGATCCCGC**GCTGTTT CCTGTGTGAAC**	1,068,678–1,068,718
45	hs394e11'	**GGCCGCGGGAATTCGATT**GTCGGGATCCTC AATATTCCC	TGAAGATGGGCTCTGCTGGAC**CACTAG TGAATTCGCGGC**	1,085,683–1,085,724
46	hs344f5	**GGCGATTAAGTTGGGTAAC**GAAGGGGCTGGTG GCTCATCTTCTCCTTGGGAGAG	CACGGTGAGGGATGGTGTCTCTGAGG**CTGTTT CCTGTGTGAAC**	1,124,629–1,124,663
47	hs.84285	**GGCGATTAAGTTGGGTAAC**GGTTGTTTTCTTTT CTCTCAGGGTTTCGTGGCTGTCCCAA	CAAAAAATCCCGATGGCACGATGAACCTCA**GCTGT TTCCTGTGTGAAC**	1,304,274–1,304,313
48	hs349e11	**GGCGATTAAGTTGGGTAAC**CGAATAAGGCAAG TTCCCACTCCTC	GATGCGGGCTGTGCCTTATTTAT**GCTGTTTC CTGTGTGAAC**	1,318,158–1,318,182
49	hs.134846	**GGCGATTAAGTTGGGTAAC**GCTGCGGCTGCACA AATGTTCCACCCGCGAGTGCATCGAGCAGTT	CTACCTGGACAAGCTCAAACAGGTAGGGA**GCTGT TTCCTGTGTGAAC**	1,337,561–1,337,605

No.	Name	Upstream hybridising sequence	Downstream hybridising sequence	Positions[a]
50	hs.118261	**GGCGATTAAGTTGGGTAAC**GGTGCACAACAGAGC CCACATAGGAACACCAGGGCTGCGAGGGGGAGAT	GCACACGGCCACAAAAGTTCCCA**GCTGTT TCCTGTGTGAAC**	1,372,574–1,372,622
51	hs305c8'	**GGCGATTAAGTTGGGTAAC**GGTGCCTTTGCCGGG ATTCCTGAAAGGCAGGGTCCATGGT	TTGCACCGAGGAACTGGATTTTGG**GCTGTT TCCTGTGTGAAC**	1,495,719–1,495,758
52	WDTC2 (3)	**GGCCGCGGGAATTCGATT**GCTGCTCCGACTT CTTCATCGAGC	ACAGTCAGTACGAGAGGGCGGTAG**CACTAG TGAATTCGCGGC**	1,513,567–1,513,566
53	CRAMP1L	**GGCCGCGGGAATTCGATT**CCGCCAGCTTTCTGG AACAGATCAC	TGTACGGTGCTTTCTCTCCAAGGAGCT**CACTA GTGAATTCGCGGC**	1,615,928–1,615,979
54	hs.88500	**GGCGATTAAGTTGGGTAAC**GCCCACACCCTCTA ACACCTGTCTC	CTTATCGTCACATGCTGACCGTCCG**GCTGTTT CCTGTGTGAAC**	1,730,426–1,730,450
55	hs.155482	**GGCGATTAAGTTGGGTAAC**CTGCTCAGAATGTG GGAACAGGC	TGGGCGCTTTTTACACACGCAGC**GCTGTTT CCTGTGTGAAC**	1,806,637–1,806,659
56	FAHD1	**GGCCGCGGGAATTCGATT**GGAGATATTATCTTG ACTGGGACGCCAAAGG	GAGTTGGACCGGTTAAAGAAAACGATGAGAT**C ACTAGTGAATTCGCGGC**	1,817,784–1,817,845
57	HSAC76P10	**GGCCGCGGGAATTCGATT**CATTGTTGCAAATGG ACACAGTCTTAATGG	GAGGATTATTAACGTGCTTGCAGCTGTGAAATCG**CA CTAGTGAATTCGCGGC**	1,847,771–1,847,770
58	HSAC76P10 (2)	**GGCCGCGGGAATTCGATT**CAAGGAGAACGCTTGGC GTGTCCTCCTTTGGTCCAGC	CTGGGTTGCAGCTTCGATGCTCCCTCTGT**CACTAGT GAATTCGCGGC**	1,880,212–1,880,277
59	HS3ST6	**GGCCGCGGGAATTCGATT**CCACTTCAGATGCCTGA TTTGCCCGAGCT	CATAAAGGATGTGGTTCCCTCCCAGGGAG**CACTAG TGAATTCGCGGC**	1,913,866–1,913,923
60	hs.48384	**GGCGATTAAGTTGGGTAAC**GGGAGTGTCCTGATGTTT CAACCACTGCGCCCTGCCTTCCTGTCTCGACG	CCCAGGGTTCCAACTCCAAGGTGGAAT**GGCTGTTT CCTGTGTGAAC**	1,928,982–1,929,031
61	hs.198274	**GGCGATTAAGTTGGGTAAC**CCAATCCCATCGTCTACA TGATGAAAGCGTTCGACCTCATCGTGGACCG	ACCCGTGACCCTCGTGAGAGGTACGAA**GCTGTTT CCTGTGTGAAC**	1,949,688–1,949,736
62	PKD 1	**GGCGATTAAGTTGGGTAAC**GGTCATATAGAGGTTAC CTTGTATGTAGTCACG	CACATAGTCACGCACATGGCAGCCG**GCTGTTTC CTGTGTGAAC**	2,086,449–2,086,481

Letters in bold signify the 5' and 3' universal tags for amplification in the MLPA reaction
a UCSC Genome Browser (May 2004) chromosome 16p13.3

Figure 3. Scatterplots showing the MLPA results of the patients G.Z.



(a), P.V. (b), H.N. (c) and F.T. (d). The *Y*-axis represents the ratio peak height of patient divided by that of the normal control, the *X*-axis shows the chronological position of MLPA probes along the p-arm of chromosome 16.

Table 2. Hematological parameters of patients G.Z., P.V., N.M. and F.T. and family members.

Case	Sex–age (years)	Hb (g/dl)	RBC ($\times 10^{12}$/l)	MCV (fl)	MCH (pg)	MCHC (g/dl)	ZPP (umol zp/mol heme)	A_2 (%)	I.B.	α Globin genotype
Propositus G.Z.	M–31	12.7	5.84	70	22.1	31.6	53	2.8	+	aa/-GZ
Mother of G.Z.	F–55	11.6	5.50	67	21.1	31.7	ND	ND	ND	aa/-GZ
Propositus P.V.	M–11	10.5	4.66	69	22.4	32.2	29	2.9	+	aa/-PV
Father of P.V.	M–49	15.1	5.05	89	29.8	33.3	20	3.1	–	aa/aa
Mother of P.V.	F–48	13.7	4.34	91	31.6	34.6	13	2.7	–	aa/aa
Propositus H.N.	M–5	10.3	4.75	71	21.7	30.6	74	2.7	+	aa/-HN
Father of H.N.	M–41	14.0	4.90	89	28.5	32.0	35	2.7	–	aa/aa
Mother of H.N.	F–40	12.9	4.62	87	28.0	32.0	50	2.9	–	aa/aa
Propositus F.T.	F–30	10.6	5.14	74	20.8	28.2	56	2.6	ND	aa/-FT
Father of F.T.	M–60	14.8	4.84	93	30.8	33.0	32	2.6	ND	aa/aa
Mother of F.T.	F–60	15.0	4.72	95	31.9	33.6	39	2.6	ND	aa/aa
Sister of F.T.	F–27	13.8	4.67	91	29.8	32.7	30	3.0	ND	aa/aa

ZPP Zinc Proto Porphyrin (lmol ZP/mol Heme), *ND* not determined
+/– = Positive/negative Inclusion Bodies test

(α/β = 0.6). The father showed a completely normal hematology and balanced chain synthesis (α/β = 1.1).

The three patients with MR and α-thalassemia all have healthy parents presenting with normal hematological parameters (Table 2), supporting a *de novo* event.

Molecular analysis

Case G.Z.

The MLPA probes 1–35 appeared to be deleted in this patient as shown in the previous study (Harteveld *et al*. 2005), indicating that the 3'breakpoint was located somewhere between the *MSLN* gene and the *TSC* gene, approximately 2 Mb from the telomere of 16p. Additional MLPA primer pairs designed in this study, revealed the deletion breakpoint to be localized between probe no. 37 and 40, which spans a region of approximately 100 kb (Fig. 3a). Two probes, no. 38 and 39 were designed to narrow the breakpoint region to 31 kb between positions 869,698 and 900,907. The deletion involves the telomeric region including the complete α-globin gene cluster but leaving the *SOX8*-gene intact. According to the hematological analysis of the parents, the deletion is inherited from the mother, who presented with microcytic hypochromic

anemia, similar to the propositus. Unfortunately, no DNA of the parents was available for MLPA analysis. No clinical phenotype other than α^0-thalassemia seems to be associated with this deletion.

Case P.V.
During an initial screening, a telomeric deletion was observed extending to *KIAA0683* at position 1,495,758 (probe no. 51). Two probe pairs (nos. 52 and 53) were designed between the last MLPA probe pair deleted in this patient and the first still present (probe no. 54) to reduce the breakpoint region. Both probes 52 and 53 appeared to be involved in the deletion reducing the breakpoint region to 114 kb between positions 1,615,979 and 1,730,426 (Figs. 2, 3b). This deletion of approximately 1.62–1.73 Mb causes monosomy for several genes including the *SOX8*-gene. No other chromosomal aberrations were found in this patient, neither at the cytogenetic level, by FISH analysis, nor by MAPH analysis (Kriek *et al*. 2004).

Case H.N.
After an initial screening the deletion was found to involve a 1.8–1.93 Mb region from the telomere. An additional MLPA analysis using newly designed probe pairs 56–59 identified the deletion breakpoint between probes 59 and 60, which limits the breakpoint region to 15 kb between positions 1,913,923–1,928,982 (Figs. 2, 3c).

Case F.T.
Initially, the deletion length appeared similar to that of patient H.N.. Fine mapping using MLPA probe pairs 56–59 revealed the 3'breakpoint to be different. The breakpoint is localized between probe pairs 58 and 59, which equal a deletion length of 1.9 Mb from the telomere to position 1,880,277–1,913,866 (Figs. 2, 3d). This region is extremely rich in Alu repeats, which may have played a role in the mechanism leading to these large deletions. This also might explain the observed clustering of 3¢breakpoints in patients F.T. and H.N.

Discussion

In contrast to the ATR-X syndrome, ATR-16 does not present with a clearly defined phenotype. Sixteen cases have been described in literature and in most cases it was not clear whether the dysmorphic features were attributable to the monosomy for 16p

or to the additional chromosomal aberrations found in these patients (Brown *et al*. 2000; Eussen *et al*. 2000; Gallego *et al*. 2005; Warburton *et al*. 2000; Wilkie 1990a). Only five patients (indicated as BA, TN, BO, IM and LIN in Fig. 2) were described with a clear monosomy for a telomeric deletion of 16p (Daniels *et al*. 2001; Fei *et al*. 1992; Lamb *et al*. 1993; Lindor *et al*. 1997; Wilkie *et al*. 1990a). We present the high resolution mapping by MLPA of ATR-16 deletions in four Caucasian patients affected with microcytic hypochromic anemia without iron deficiency, due to a large deletion including the complete a-gene cluster. One patient (G.Z.) with a deletion of 870–900 kb presented no dysmorphic features or MR, while three patients showing deletions ranging from 1.7 to 1.9 Mb presented with mild MR (IQ 50–60) and a variety of dysmorphic features. The samples have been checked for unbalanced translocations and partly for deletions/duplications in other parts of the genome, either using telomere MLPA (Schouten *et al.* 2002; Kriek *et al*. 2004) for samples H.N. and P.V.) or a 1420-plex bead-based MLPA (for samples H.N. and F.T., Fan *et al*. 2006; Aten *et al*., in preparation). No other deletions/ duplications or unbalanced translocations were found in H.N., F.T. and P.V. other than the deletions at the tip of chromosome 16p13.3 confirmed by MLPA analysis using the probes described in this manuscript.

The two patients, G.Z. and P.V., showing respectively the largest deletion with only α^0-thalassemia and the smallest deletion clearly associated with the classical ATR16 features, might give a better insight into the genes for which haploinsufficiency contributes to the syndrome. Three cases - $-^{BO}$,- $-^{IM}$ and - $-^{LIN}$ are known from the literature without additional chromosomal rearrangements besides the deletion causative of ATR-16 (Daniels *et al*. 2001; Fei *et al.* 1992; Lindor *et al*. 1997). These deletions are larger than the presently described case, which restricts the region for which monosomy seems to contribute to the ATR-16 associated phenotype to an approximately 800 kb region between 0.9 and 1.7 Mb from the telomere of 16p. Approximately 14 genes and gene families of known function are located in this area.

The *SOX8* gene is a member of the *SOX* (*SRY*-related HMG-box) family and encodes for a transcription factor involved in regulation of embryonic development and in determination of cell fate. The *SOX8*-protein is suggested to be involved in brain development and function and is strongly expressed in brain and less abundant in other tissues. Therefore *SOX8* is considered to be a good candidate gene for which haploinsufficiency may contribute to the MR phenotype seen in ATR-16 patients (Holinski-Feder *et al*. 2000; Pfeifer *et al*. 2000). However, MLPA analysis in several members of a Brazilian family without MR or dysmorphic features using the 62 probes showed a deletion of the tip of the short arm of chromosome 16 including *SOX8* (manuscript in preparation).

Other disease genes located in the 800 kb region deleted between G.Z. and P.V. are *CACNA1H*, *GNPTG* and *CLCN7*. These genes are associated respectively with childhood absence epilepsy, autosomal recessive pseudo Hurler polydystrophy and autosomal dominant Albers-Schonberg osteopetrosis type II (Cleiren *et al.* 2001; Liang *et al.* 2007; Perez-Reyes 2006; Tiede *et al.* 2006). It is not clear, however, how haploinsufficiency would lead to the phenotypic features seen in ATR-16. Of the three patients described in this report, only one (P.V.) suffers from epilepsy, while the others do not. No other features typical for pseudo-Hurler or Albers-Schonberg disease are seen in the ATR-16 patients described. Members of a Tryptase precursor gene family (*TPSG1, AB1, B2* and *D1*) are located between 1.21 and 1.26 Mb, and are believed to play a role in the pathophysiology of the polygenic disorder of asthma (Pallaoro *et al.* 1999). Also in P.V. and H.N. asthma was reported, supporting the assumed involvement of the Tryptase family genes. Of the other genes in this region, such as *C1QTNF8, UBE2I, BAIAP3, IFT140, C16orf30* and *CRAMP1L*, only *BAIAP3* is highly expressed in brain. This gene encodes a transmembrane protein, a member of the secretin receptor family, which interacts with the cytoplasma specific angiogenesis inhibitor1 and may be involved in synaptic functions (Shiratsuchi *et al.* 1998). To determine how haploinsufficiency for these genes may be of influence on the intellectual development and variability of dysmorphic features seen in these patients, more deletions should be studied.

Some common features associated with ATR-16 include a severe delay in active language ability, downslant of the palpebral fissures, mild hypertelorism, a broad nasal bridge and small ears and a short neck with webbing. Most of these features are seen in two of our patients (P.V. and H.N.), who are considered monosomic for the telomeric deletion on chromosome 16p. On the other hand, patient F.T. who has a deletion length similar to H.N. shows much less pronounced dysmorphic features presumed to be characteristic for ATR-16, which subscribes the variability in expression of this syndrome. All patients have at least one common feature measurable at the hematological level, i.e. α^0-thalassemia. Nevertheless only F.T. was recognized as a possible ATR-16 syndrome at the hematological level, because of a persistent microcytic hypochromic anemia at normal ferritin levels. The other two patients (P.V. and H.N.) were identified by MAPH screening using subtelomeric probes and mapped in detail by MLPA in the present study. Because dysmorphic features associated with ATR16 are not always very specific, we would like to plead for incorporating a simple hematological test if ATR-16 is suspected and, when positive, recommend a molecular screening using the 62 MLPA probes as described in this study.

Ultimately cloning of the breakpoint and subsequent sequence analysis is the only way to determine the nature of the deletion found, e.g. healed telomere break, interstitial deletion or translocation to another non-coding subtelomeric region (with the loss of the area involved in ATR16) and this will be subject for future study. On the other hand MLPA is a strong diagnostic tool in determining whether a genomic region is deleted, the extent of the deletion and the location of breakpoints. Because MLPA employs standard technology operational in most diagnostic laboratories, it is highly suitable for rapid testing for these disorders, which are believed to be under diagnosed (Daniels *et al*. 2001; Wilkie *et al*. 1990a).

Acknowledgments

Chiara Refaldi and Emmanora di Capua for technical assistance in performing the MLPA analysis and primer design, Johan den Dunnen and Martijn Breuning for fruitful discussion and advice. Part of this work has been made possible thanks to European Project ITHANET.

References

Bernini LF, Harteveld CL (1998) Alpha-thalassaemia. *Baillieres Clin Haematol* **11**:53–90

Brown J, Horsley SW, Jung C, Saracoglu K, Janssen B, Brough M, Daschner M, Beedgen B, Kerkhoffs G, Eils R, Harris PC, Jauch A, Kearney L (2000) Identification of a subtle t(16;19)(p13.3;p13.3) in an infant with multiple congenital abnormalities using a 12-colour multiplex FISH telomere assay, M-TEL. *Eur J Hum Genet* **8**:903–910

Cardoso C, Timsit S, Villard L, Khrestchatisky M, Fontes M, Colleaux L (1998) Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. *Hum Mol Genet* 7:679–684

Cardoso C, Lutz Y, Mignon C, Compe E, Depetris D, Mattei MG, Fontes M, Colleaux L (2000) ATR-X mutations cause impaired nuclear location and altered DNA binding properties of the XNP/ATR-X protein. *J Med Genet* **37**:746–751

Chong SS, Boehm CD, Higgs DR, Cutting GR (2000) Single-tube multiplex-PCR screen for common deletional determinants of alpha-thalassemia. *Blood* **95**:360–362

Cleiren E, Benichou O, Van Hul E, Gram J, Bollerslev J, Singer FR, Beaverson K, Aledo A, Whyte MP, Yoneyama T, deVernejoul MC, Van Hul W (2001) Albers-Schonberg disease (autosomal dominant osteopetrosis, type II) results from mutations in the ClCN7 chloride channel gene. *Hum Mol Genet* **10**:2861–2867

Dacie JV, Lewis SM (1991) Practical haematology, 7th edn. Churchill Livingstone, Edinburgh

Daniels RJ, Peden JF, Lloyd C, Horsley SW, Clark K, Tufarelli C, Kearney L, Buckle VJ, Doggett NA, Flint J, Higgs DR (2001) Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16. *Hum Mol Genet* **10**:339–352

Eussen BH, Bartalini G, Bakker L, Balestri P, Di Lucca C, Van Hemel JO, Dauwerse H, van Den Ouweland AM, Ris-Stalpers C, Verhoef S, Halley DJ, Fois A (2000) An unbalanced submicroscopic translocation t(8;16)(q24.3;p13.3)pat associated with tuberous sclerosis complex, adult polycystic kidney disease, and hypomelanosis of Ito. *J Med Genet* **37**:287–291

Fan J, White SJ, Bibikova M, Zhou L, Chen J, Wickham-Garcia E, Kalf M, van Ommen GJ, Breuning MH, Guo L, Lu S-H, Zhan Q, Jiang W, Chan O, Wang-Rodriguez J, Barker DL, den Dunnen JT (2006) Screening of deletions and duplications in 1500 genomic loci in a single assay. *Am J Hum Genet* (56th Annual meeting, New Orleans, Louisiana), **25**, 10 September 2006

Fei YJ, Liu JC, McKie VC, Huisman TH (1992) Hb H disease and mild mental retardation in a black girl with a Hb S heterozygosity. *Hemoglobin* **16**:431–434

Flint J, Rochette J, Craddock CF, Dode C, Vignes B, Horsley SW, Kearney L, Buckle VJ, Ayyub H, Higgs DR (1996) Chromosomal stabilisation by a subtelomeric rearrangement involving two closely related Alu elements. *Hum Mol Genet* **5**:1163–1169

Gallego MS, Zelaya G, Feliu AS, Rossetti L, Shaffer LG, Bailey KA, Bacino CA, Barreiro CZ (2005) ATR-16 due to a *de novo* complex rearrangement of chromosome 16. *Hemoglobin* **29**:141–150

Gibbons RJ, Higgs DR (2000) Molecular-clinical spectrum of the ATR-X syndrome. *Am J Med Genet* **97**:204–212

Gibbons RJ, Suthers GK, Wilkie AO, Buckle VJ, Higgs DR (1992) X-linked alpha-thalassemia/mental retardation (ATR-X) syndrome: localization to Xq12-q21.31 by X inactivation and linkage analysis. *Am J Hum Genet* **51**:1136–1149

Giordano PC, van Delft P, Batelaan D, Harteveld CL, Bernini LF (1999) Haemoglobinopathy analyses in

the Netherlands: a report of an in vitro globin chain biosynthesis survey using a rapid, modified method. *Clin Lab Haematol* **21**:247–256

Harteveld CL, Voskamp A, Phylipsen M, Akkermans N, den Dunnen JT, White SJ, Giordano PC (2005) Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha-and beta-thalassaemia characterised by high resolution multiplex ligation-dependent probe amplification. *J Med Genet* **42**:922–931

Higgs DR (1993) alpha-Thalassaemia. *Baillieres Clin Haematol* **6**:117–150

Holinski-Feder E, Reyniers E, Uhrig S, Golla A, Wauters J, Kroisel P, Bossuyt P, Rost I, Jedele K, Zierler H, Schwab S, Wildenauer D, Speicher MR, Willems PJ, Meitinger T, Kooy RF (2000) Familial mental retardation syndrome ATR-16 due to an inherited cryptic subtelomeric translocation, t (3;16) (q29;p13.3). *Am J Hum Genet* **66**:16–25

Horsley SW, Daniels RJ, Anguita E, Raynham HA, Peden JF, Villegas A, Vickers MA, Green S, Waye JS, Chui DH, Ayyub H, MacCarthy AB, Buckle VJ, Gibbons RJ, Kearney L, Higgs DR (2001) Monosomy for the most telomeric, gene-rich region of the short arm of human chromosome 16 causes minimal phenotypic effects. *Eur J Hum Genet* **9**:217–225

Kriek M, White SJ, Bouma MC, Dauwerse HG, Hansson KB, Nijhuis JV, Bakker B, van Ommen GJ, den Dunnen JT, Breuning MH (2004) Genomic imbalances in mental retardation. *J Med Genet* **41**:249–255

Lamb J, Harris PC, Wilkie AO, Wood WG, Dauwerse JG, Higgs DR (1993) *De novo* truncation of chromosome 16p and healing with (TTAGGG)n in the alpha-thalassemia/mental retardation syndrome (ATR-16). *Am J Hum Genet* **52**:668–676

Liang J, Zhang Y, Chen Y, Wang J, Pan H, Wu H, Xu K, Liu X, Jiang Y, Shen Y, Wu X (2007) Common polymorphisms in the CACNA1H gene associated with childhood absence epilepsy in Chinese Han population. *Ann Hum Genet* **71**:325–335

Lindor NM, Valdes MG, Wick M, Thibodeau SN, Jalal S (1997) *De novo* 16p deletion: ATR-16 syndrome. *Am J Med Genet* **72**:451– 454

Liu YT, Old JM, Miles K, Fisher CA, Weatherall DJ, Clegg JB (2000) Rapid detection of alpha-thalassaemia deletions and alpha-globin gene triplication by multiplex polymerase chain reactions. *Br J Haematol* **108**:295–299

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**:1215

Pallaoro M, Fejzo MS, Shayesteh L, Blount JL, Caughey GH (1999) Characterization of genes encoding known and novel human mast cell tryptases on chromosome 16p13.3. *J Biol Chem* **274**:3355–3362

Perez-Reyes E (2006) Molecular characterization of T-type calcium channels. *Cell Calcium* **40**:89–96

Pfeifer D, Poulat F, Holinski-Feder E, Kooy F, Scherer G (2000) The SOX8 gene is located within 700 kb of the tip of chromosome 16p and is deleted in a patient with ATR-16 syndrome. *Genomics* **63**:108–116

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

Shiratsuchi T, Oda K, Nishimori H, Suzuki M, Takahashi E, Tokino T, Nakamura Y (1998) Cloning

and characterization of BAP3 (BAI-associated protein 3), a C2 domain-containing protein that interacts with BAI1. *Biochem Biophys Res Commun* **251**:158– 165

Tiede S, Cantz M, Spranger J, Braulke T (2006) Missense mutation in the *N*-acetylglucosamine-1-phosphotransferase gene (GNPTA) in a patient with mucolipidosis II induces changes in the size and cellular distribution of GNPTG. *Hum Mutat* **27**:830–831

Villard L, Lossi AM, Cardoso C, Proud V, Chiaroni P, Colleaux L, Schwartz C, Fontes M (1997) Determination of the genomic structure of the XNP/ATRX gene encoding a potential zinc finger helicase. *Genomics* **43**:149–155

Warburton P, Mohammed S, Ogilvie CM (2000) Detection of submicroscopic subtelomeric chromosome translocations: a new case study. *Am J Med Genet* **91**:51–55

Weatherall DJ, Higgs DR, Bunch C, Old JM, Hunt DM, Pressley L, Clegg JB, Bethlenfalvay NC, Sjolin S, Koler RD, Magenis E, Francis JL, Bebbington D (1981) Hemoglobin H disease and mental retardation: a new syndrome or a remarkable coincidence? *N Engl J Med* **305**:607–612

White SJ, Vink GR, Kriek M, Wuyts W, Schouten J, Bakker B, Breuning MH, den Dunnen JT (2004) Two-color multiplex ligation-dependent probe amplification: detecting genomic rearrangements in hereditary multiple exostoses. *Hum Mutat* **24**:86–92

Wilkie AO, Buckle VJ, Harris PC, Lamb J, Barton NJ, Reeders ST, Lindenbaum RH, Nicholls RD, Barrow M, Bethlenfalvay NC (1990a) Clinical features and molecular analysis of the alpha thalassemia/mental retardation syndromes. I. Cases due to deletions involving chromosome band 16p13.3. *Am J Hum Genet* **46**:1112–1126

Wilkie AO, Zeitlin HC, Lindenbaum RH, Buckle VJ, Fischel-Ghodsian N, Chui DH, Gardner-Medwin D, MacGillivray MH, Weatherall DJ, Higgs DR (1990b) Clinical features and molecular analysis of the alpha thalassemia/mental retardation syndromes. II. Cases without detectable abnormality of the alpha globin complex. *Am J Hum Genet* **46**:1127–1140

Wilkie AO, Gibbons RJ, Higgs DR, Pembrey ME (1991) X linked alpha thalassaemia/mental retardation: spectrum of clinical features in three related males. *J Med Genet* **28**:738–741

Yntema HG, Poppelaars FA, Derksen E, Oudakker AR, van Roosmalen T, Jacobs A, Obbema H, Brunner HG, Hamel BC, van Bokhoven H (2002) Expanding phenotype of XNP mutations: mild to moderate mental retardation. *Am J Med Genet* **110**:243–247

Chapter III-4

Comparison of four genome wide platforms using four overlapping interstitial 2p alterations

Marjolein Kriek[1], Claudia A. L. Ruivenkamp[1], Yavuz Ariyurek[1], Margot E. Kalf[1], Jeroen Knijnenburg[2], Arie van Haeringen[1], Emanuela Lucci-Cordisco[3], Maurizio Genuardi[4], Carla Rosenberg[5], Ana C. Krepischi-Santos[5], S.R. Sanders[6], Stefan J. White[1], K. Szuhai[2], Martijn H. Breuning[1], Johan T. den Dunnen[1]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands
[2]Dept. Molecular Cell Biology, Leiden University Medical Center, Leiden, The Netherlands
[3]EL-C: Institute of Medical Genetics, Catholic University "A. Gemelli" School of Medicine, Rome, Italy
[4]MG: Medical Genetics Unit, Department of Clinical Pathophysiology, University of Florence Medical School, Florence, Italy
[5]Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil.
[6]Section of Genetics & Metabolism, Health Science Centre, Winnipeg, Canada

Submitted

Abstract

Molecular karyotyping by array-based techniques represents a giant leap forward compared to microscopic metaphase banding. We compared the performance of four different array based platforms to identify and map the breakpoints in four patients with different interstitial 2p deletions, all localised within 2p16.1-p21. Currently, there are two main array formats, array-CGH and SNP-based. For array-CGH the probes used are (3K–32K) genomic clones or up to 244K oligonucleotides, with the size and number determining the resolution of analysis. SNP arrays, containing 10K–1000K loci have proven to facilitate, in addition to genome-wide association studies, the detection of deletions and duplications. The resolution of these arrays depends on the number of SNP loci present and on their distribution across the genome.

In this study, the platforms used include a 3K large genomic insert clone array, a 44K (long oligo) microarray and two SNP- based arrays (250-500K, and 317K). Our analysis showed that the size of the 2p deletions varied, from ~10.6Mb in patient 1, to ~2.4Mb in patient 4. The minimum region of overlap of the deletions was ~1.3Mb encompassing 8 genes. The *MSH6* gene was deleted in minimally three out of four patients, indicating that they have a 60-90% chance of developing colon carcinoma. No clear genotype/phenotype correlation emerged from the comparison of the four patients.

Comparing cross-platform the breakpoint mapping gave similar results in the majority of cases.

Introduction

For decades trypsin Giemsa banding of metaphase spreads has been the standard diagnostic method to detect chromosomal rearrangements. The method has several advantages; all chromosomes are seen under the microscope, and individual cells can be karyotyped, permitting clonal analysis[1,2] and the study of mosaicism. A major limitation is the fact that due to the contraction of chromosomes during metaphase and the resolution of the light microscope, G banding is not capable of identifying rearrangements smaller than 3-5 Mb.

Fluorescence in Situ Hybridization (FISH)[3,4] partly overcomes this problem, allowing direct testing for the presence, absence or amplification of specific genomic regions. This method is especially used for the confirmation of microdeletion syndromes and the analysis of potential subtelomeric rearrangements. FISH analysis can also be used for the detection of mosaicism to a very low level, depending on the number of cells analysed. However, it has several drawbacks, as detecting rearrangements using FISH analysis is only possible when cells are available, an obvious, specific phenotype is present that is recognized by a specialist, and when a specific FISH probe exists. Finally, although multi-colour methodologies have been developed[5,6,7] the number of loci that can be analysed simultaneously is limited.

Recently, array-based technologies have been developed that provide both genome-wide and high resolution analysis. In contrast to FISH, where fragments of DNA are labeled and hybridized to chromosome spreads, array-based approaches label genomic DNA, which is then hybridized to DNA spotted on a solid support, typically a glass slide. The size of the DNA probe and the number of probes on the array determine the resolution of analysis.

The first arrays used relatively large DNA fragments (~150 kb) isolated from Bacterial Artificial Chromosome (BAC) or P1 derived Artificial Chromosome (PAC) clones.[8-10] A newer format uses oligonucleotide probes of 25 to 60 nt in length.[11,12] Due to the smaller size of these probes and the much larger number of loci analysed, it is possible to detect much smaller copy number variations (CNVs) with greater precision compared to those that can be revealed using BAC-PAC clone arrays. The 25-mer probe arrays were originally designed for SNP analysis. However, they were quickly used to estimate copy number changes by using both signal strength and allele scoring. Initial studies used the Affymetrix 10K array, which demonstrated the principle that the arrays could provide quantitative data.[13] Subsequent work has taken advantage of

higher resolution chips, currently up to 500-1000K.[14] In practice, these arrays have an effective resolution below 10 kilobases. However, despite their extremely high resolution, it should be noted that these tools can not be used to detect copy neutral rearrangements like translocations and inversions.

In this study, we have analysed four patients with different sizes of interstitial 2p deletions, all localised within the chromosome region 2p16.1-p21. We have compared different platforms for identifying the deletions as well as their ability to define breakpoints. In addition, we have collated and compared the clinical data of these patients. It appears that psycho-motor delay is the only common clinical feature that corresponds to a deletion within this area.

Methods

Array- Comparative Genomic Hybridisation (Array-CGH)

The array-CGH procedures were performed as previously described.[15] The clones were provided by the Wellcome Trust Sanger Institute (UK), and information regarding the full set is available at the Ensembl web site.

The array contained ~3500 large genomic insert clones spaced at ~1 Mb intervals over the genome, meaning that the resolution of the array varies between 0.2-3 Mb. Profiles were displayed by using the Log(2) ratio of test and reference sample. The thresholds were set at -0.3 and 0.3. BACs with a Log(2) ratio outside this interval were considered to be altered.[15,16]

Agilent microarray

Agilent Human Genome CGH Microarrays consist of ~44,000 60-mer oligonucleotide gene focused probes that span coding and non-coding sequences with an average spatial resolution of ~35 kb. Both genes with known function and hypothetical genes were included in the array. We used a loop-hybridisation design to analyse six DNA samples, including three patients with an interstitial 2p deletion. In a loop hybridisation design DNA sample 1 and 2 are differently labelled (Cy5 and Cy3, respectively) and subsequently hybridised on the same array. The second array includes DNA sample 2 and 3 that are labelled using Cy5 and Cy3, respectively. In this way, one sample is hybridised twice on an independent array in two different colors (= dye swap), enabling its own internal quality control. Arrays were hybridised according to the recom-

mendations of the supplier (www.agilent.com). Data were analysed using the Agilent CGH Analytics 3.4 software with a moving averaging window of 1Mb. The size of the three different deletions was calculated using the Log(2) ratio. The thresholds were set at -0.3 and 0.3.

Affymetrix 500K Genechip

The Genechip Human Mapping 500K array set was used. The procedure was performed as described in the Affymetrix GeneChip Human Mapping 500K Manual (http://www.affymetrix.com). The set comprises two SNP arrays contain ~250.000 25-mer oligonucleotides each. Using this protocol, the human genome is cut by restriction enzymes (NspI and StyI); one restriction enzyme is used per array, enabling the analysis of 250,000 loci. The use of a second restriction enzyme is necessary for the analysis of another 250,000 loci. For data analysis, DNA-Chip Analyzer (dChip) software (version release 02-16-06) was used.[17,18] Regions of copy number gain and loss were detected using the hidden Markov model output of dChip. The thresholds for this platform were set between 1.6 and 2.4 using a linear scale, in where 2.0 represents two copies of a given locus.[11,19]

Illumina 317K beadchip

The Illumina humanhap 317K genotyping beadchip work up has been performed as suggested by the manufacturer (www.illumina.com). The SNP array consists of 317,000 25-mer oligonucleotide probes. For data analysis, the beadstudio data analysis software provided by Illumina was used.

In this platform, the regions for CNVs are detected based on the LogR ratio. This tool combines data of both heterozygosity (SNP call) and signal strength. The thresholds were set at -0.3 and +0.3. In addition to the Log R ratio, the data analysis software also provides B allele frequency, Loss of Heterozygosity (LOH) and Copy number (CN) score.

Patient samples

The four patient samples were gathered from the Netherlands (patient 1), Canada (patient 2), Brazil (patient 3) and Italy (patient 4), respectively. Two of them have been described previously.[20,21]

The DNA of the patients was applied to each platform once, except for the Agilent array (due to the dye swap procedure). This study was approved by the Institutional Review Board of the Leiden University Medical Center, conforming with Dutch law.

Table 1. Overview of the BAC array results.

Intnl Clone name	Chrom.	Patient 1	Patient 2	Patient 3	Patient 4
RP11-204D19	2				
RP11-24I5	2	■	■		
RP11-421J10	2	■	■		
RP11-27C22	2	■	■		
RP11-110G2	2	■	■		
RP11-1084a21	2	■	■		■
RP11-436K12	2	■	■		■
RP5-960D23	2	■	■		■
RP11-19A8	2	■	■	■	■
RP11-436L21b	2	■	■	■	■
RP11-436L21	2	■	■	■	■
RP11-460M2	2	■	■	■	
RP11-319N5	2	■	■	■	
RP11-5M9	2	■	■	■	
RP11-391D19	2	■	■	■	
RP11-389K20	2	■	■	■	
RP11-335O22	2	■	■	■	
RP11-7H13	2	■	■		
RP11-508L23	2	■			
RP11-30C22	2				

Minimum region of overlap

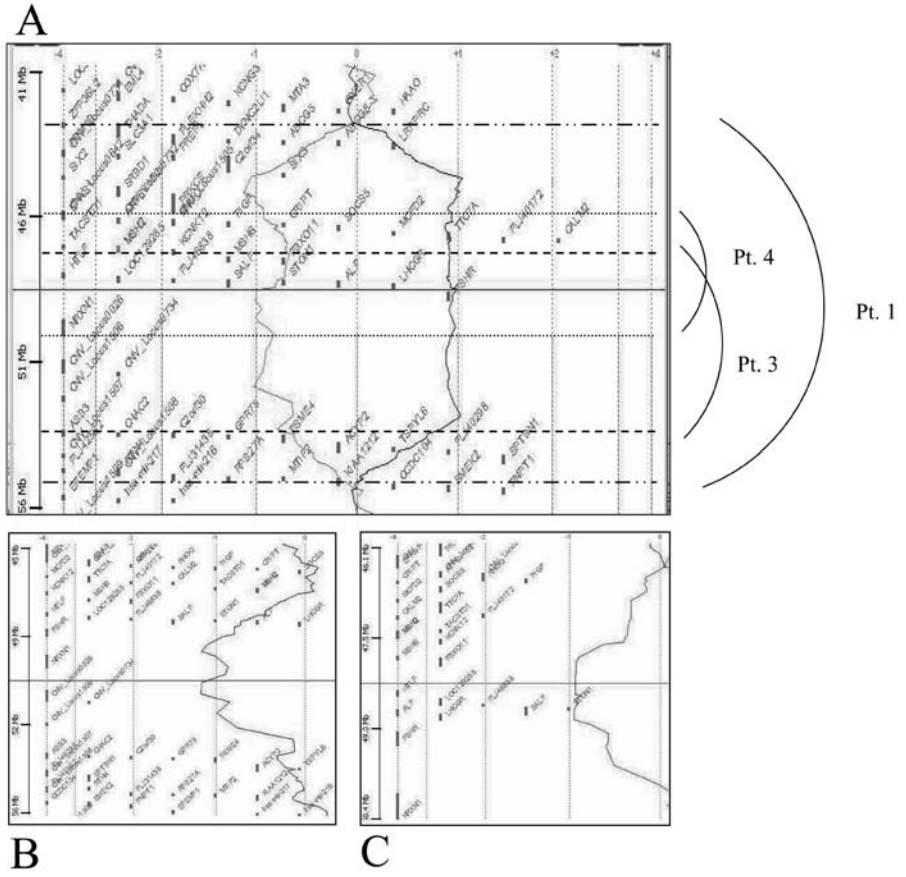
Deleted BACs per patient, depicted as grey bars.

Results

Initial chromosome analysis of patient 1 did not reveal any abnormalities. However, by using both Multiplex Amplifiable Probe Hybridisation (MAPH)[22] and 3K array-CGH, it was possible to identify a deletion of chromosome region 2p16.2-p21. Retrospective analysis of the karyogram (G-banding) did detect the interstitial 2p deletion. The banding pattern of the short arm of chromosome 2 of this patient was compared to that of the previously described patient with a deletion within this region,[20] and was found to be similar (data not shown).

To study deletions in this region and their phenotypic consequences in more detail, we collected DNA from three additional patients with overlapping interstitial 2p deletions. These DNA samples were hybridised on four different array platforms to test the performance of these platforms and to map the deletion breakpoints as precisely as possible.

Figure 1. The result of patients using Agilent microarray platform.



(A) Due to the dye swap, the deletion is depicted in two colors resulting in a symmetrical profile pattern. All genes localized within the deleted region are visualized using the Agilent software tool. The deleted region of patient 3 (B) and patient 4 (C) are outlined by the dashed line and the dotted line, respectively. For the size of the deletion see table 2. Pt= patient. [See appendix: colour figures.]

Platform 1: 3K BAC array

A summary of the results obtained by array-CGH analysis is shown in table 1. The deletion of patient 1 closely resembles that of patient 2 although it extends one centromeric BAC further. The deletion of patient 4 is the smallest. The minimal region of overlap is defined by the telomeric breakpoint of patient 3 and the centromeric breakpoint of patient 4 and it is estimated to be 1.4 - 1.5 Mb.

Table 2. Localisation of the breakpoints of the four deletions identified by the 3K BAC array, the 44K micro-array of Agilent, the 500K Affymetrix Genechip and the 317K beadchip of Illumina.

Patient 1	**Last probe +**	**Genome position**	**First probe -**	**Genome position**	**Last probe -**	**Genome position**	**First probe +**	**Genome position**	**Maximum size**	**Minimum size**
aCGH	RP11-204D19	43,517Kb	RP11-24I5	44,273Kb	RP11-508L23	53,793Kb	RP11-30C22	54,808Kb	11,291Kb	9,520Kb
Agilent	A_14_P115860	44,049Kb	A_14_P119582	44,076Kb	A_14_P135447	54,658Kb	A_14_P126243	54,703Kb	10,654Kb	10,582Kb
Affymetrix	rs6736282	44,058Kb	rs17031803	44,066Kb	rs4387841	54,608Kb	rs4671950	54,613Kb	10,555Kb	10,542Kb
Illumina	rs6723119	44,041Kb	rs11124960	44,069Kb	rs10496032	54,577Kb	rs11896012	54,631Kb	10,590Kb	10,508Kb
Patient 2	**Last probe +**		**First probe -**		**Last probe -**		**First probe +**			
aCGH	RP11-204D19	43,517Kb	RP11-24I5	44,273Kb	RP11-7H13	52,903Kb	RP11-508L23	53,793Kb	10,276Kb	8,630Kb
Agilent	N.D.		N.D.		N.D.		N.D.			
Affymetrix	rs7562014	43,991Kb	rs4953037	44,003Kb	rs1318578	53,557Kb	rs6727792	53,558Kb	9,567Kb	9,554Kb
Illumina	rs4953037	44,003Kb	rs12712900	44,010Kb	rs10164725	53,560Kb	rs4672456	53,588Kb	9,585Kb	9,550Kb
Patient 3	**Last probe +**		**First probe -**		**Last probe -**		**First probe +**			
aCGH	RP5-960D23	47,773Kb	RP11-460n15	47,795Kb	RP11-335O22	52,192Kb	RP11-7H13	52,903Kb	5,130Kb	4,397Kb
Agilent	A_14_P122998	47,627Kb	A_14_P102713	47,658Kb	A_14_P131858	52,224Kb	A_14_P115721	52,790Kb	5,163Kb	4,566Kb
Affymetrix	rs13401500	48,028Kb	rs6729916	48,028K	rs1861980	53,624Kb	rs4672481	53,627Kb	5,599Kb	5,596Kb
Illumina	rs2651767	47,926Kb	rs2134056	47,927Kb	rs2287511	53,626Kb	rs6746107	53,639Kb	5,713Kb	5,699Kb
Patient 4	**Last probe +**		**First probe -**		**Last probe -**		**First probe +**			
aCGH	RP11-110G2	46,104Kb	RP11-1084a21	?	RP11-436L21	49,158Kb	RP11-460M2	49,302Kb	3,198Kb	?
Agilent	A_14_P103946	46,819Kb	A_14_P105713	46,847Kb	A_14_P111706	49,272Kb	A_14_P101515	49,560Kb	2,741Kb	2,425Kb
Affymetrix	rs17035674	46,884Kb	rs2289929	46,898Kb	rs4971697	49,269Kb	rs12713041	49,270Kb	2,386Kb	2,371Kb
Illumina	rs1053952	46,903Kb	rs1374274	46,909Kb	rs6743414	49,278Kb	rs12622540	49,290Kb	2,387Kb	2,369Kb

Agilent data was not obtained from patient 2. The size of the deletion of patient 3 using Affymetrix genechip was calculated based on 250K resolution. Three out of six of the breakpoint locations defined by Agilent were not in agreement with the results provided by the SNP arrays (proximal and distal breakpoint of patient 3 and the distal breakpoint of patient 4). In three of the breakpoints there is a small difference in localisation of the breakpoints obtained by Affymetrix and Illumina (distal breakpoint of patient 2 and 4; proximal breakpoint of patient 4). One exception includes the distal breakpoint of patient 3, in which there is a difference of five SNPs encompassing 100 Kb of genomic sequence. Data were based on the Ensemble website, assembly December 2006.
?: the localisation of this BAC is unknown in Ensemble. N.D.: not determined

Platform 2: 44K Agilent Technologies microarray

At the time this study was performed, the Agilent 44K oligo array was the only Agilent array available, covering only gene-based sequences of the human genome. Three of the four interstitial 2p deletions were tested using this oligo array (figure 1). Patient 2 could not be tested, as there was not enough material available.

Due to the loop-hybridisation set up (see Material and Methods) in combination with a dye swap, the samples were effectively analysed in two independent hybridisations. The analysis maps the proximal breakpoint of patient 3 to the region between the *NRXN1* and the *ASB3* genes, a large region (2.8 Mb) devoid of known genes. Consequently, the estimation of this breakpoint might be less accurate. The results of this platform agree with the outcome of the array-CGH. Deletion size varied, from a maximum in patient 1 (10.6 – 10.7 Mb) to a minimum in patient 4 (2.4 – 2.7 Mb) (table 2). The size of the minimum region of overlap calculated based on the Agilent data is 1.6 Mb.

Platform 3: Affymetrix 250K / 500K Genechip

Three out of four patients were analysed using a 500K Genechip (patients 1, 2, 4). Patient 3 was analysed only by the 250K Genechip using the NspI restriction enzyme (figure 2). Interestingly, the sizes of the deletions calculated based on 250K analysis were comparable with those obtained from the combined data of both arrays (500K) (data not shown), indicating that for the calculation of large CNV the use of only one restriction enzyme can be sufficient.

The minimal region of overlap between the different interstitial deletions on chromosome band 2p is 1.2 Mb (table 2).

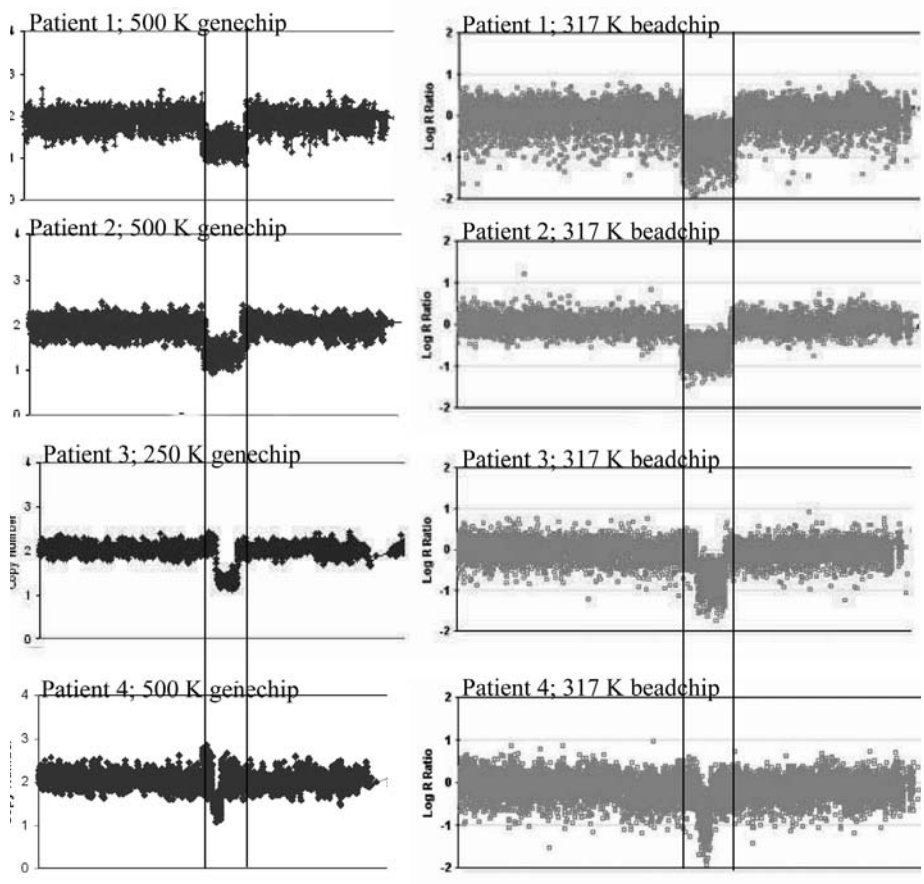
Platform 4. Illumina 317K beadchip

The results obtained for all four patients are depicted in figure 2. The results regarding the sizes are in agreement with the results obtained using the other 3 platforms (table 2). The minimal region of overlap, based on the results of the beadchip, is 1.4 Mb.

Discussion

In this study, different high resolution genome wide screening platforms were compared, including array-CGH using large insert clones, the long-oligo array of Agilent, the Affymetrix Genechip and the beadchip of Illumina. The genechip and the beadchip are SNP based arrays and they both use short-oligos.

Figure 2. The interstitial 2p deletion of the four patients analysed by Affymetrix genechip (left) and the beadchip of Illumina (right).



The deletions of the different patients are shown separately. Patient 3 was only analysed using 250K NspI genechip. A normal copy number of two is represented by a copy number between 1.6 and 2.4 for the Affymetrix genechip or by a LogR ratio between –0.3 and +0.3 for the beadchip of Illumina. The vertical lines represent the size of the largest deletion. In general, the variation of the data points obtained by the beadchip is larger than that of the genechip. Especially in patient 3, the difference in variation is remarkable. [See appendix: colour figures.]

Comparing across platforms, we found that the localisation of both proximal and distal breakpoints was largely in agreement (table 2). Nearly all BACs that showed 2 copies did not have overlap with regions that were deleted according to the results obtained by the SNP arrays and vice versa. One exception was the proximal breakpoint in patient 3 in which BAC RP11-7H13 should have been deleted according to the data

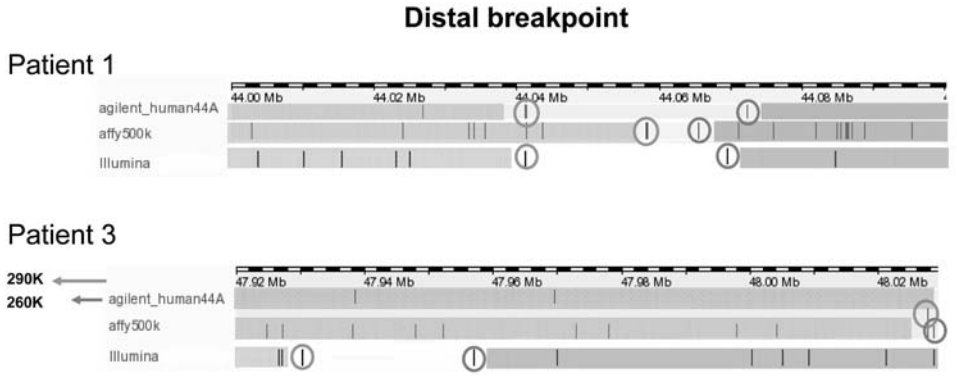
obtained by two SNP platforms. Notably, an 'aberrant' Agilent result was present at the same breakpoint as was the 'aberrant' BAC (proximal breakpoint in patient 3). In fact, the breakpoint mapping of the two array-CGH platforms was similar (breakpoint at ~52.8Mb), as was the outcome of both SNP platforms for the proximal breakpoint of patient 3 (localised at ~53.6Mb)(table 2). This might be explained by the difference in probe density near the breakpoints localised by the different platforms (see also Results). This idea is strengthened by the fact that, based on in silico data of the 244K array, (an improved version of the Agilent array), the number of oligonucleotides localised near the proximal breakpoint of patient 3 was significantly increased; 25 probes were localised within the breakpoint interval determined by the 44K array. The number of 'extra' probes present at the rest of the breakpoints defined by the 44K array is 5-10.

There is also some discrepancy between the outcomes of the two SNP platforms. In general, the data obtained using the beadchip showed more variation in all patients compared to that of genechip (figure 2). The maximum number of SNPs that were in discordance between the two SNP arrays was five (the distal breakpoint of patient 3) (figure 3). These five SNPs are in a genomic region covering more than 100 Kb. The number of data points for both SNP based arrays is similar at this breakpoint (based on the Ensembl database) (figure 3), indicating that both SNP platforms should be equally informative. We do not have a satisfactory explanation for this difference. The other differences in breakpoint mapping between the two SNP based tools included either only one or two SNPs (distal breakpoint of patient 2 and 4) or the differences in localisation of the breakpoints were very small (proximal breakpoint of patient 4; breakpoint mapping difference 7 kb). The observed difference in breakpoint mapping can be related to the use of different scoring algorithms that differ between platforms. This indicates that sequencing of the breakpoints is still needed to obtain information about the exact localisation of the breakpoint.

Patient 1, 3 and 4 did not show any copy number alterations outside chromosome region 2p16.2-p21. Patient 2, however, had a deletion on chromosome band 6q22.31 of ~1.2 Mb in size. This was identified using the BAC-array (BAC clone RP11-475J3) and the two SNP-based platforms. It has previously been found within the healthy population[23] and there are no known genes within the deleted region. It is therefore thought to be a neutral variant.

Currently, there is no golden standard available to determine which platform is the most accurate. It might be argued that high density SNP genotyping would be the

Figure 3. Overview of the distal breakpoints of patient 1 and patient 3 defined by Agilent, the Affymetrix genechip and the beadchip of Illumina.



The deleted region is depicted in red, whereas regions showing two copies are depicted in green. A green circle represents the last data point that showed a normal copy of two. A red circle represents the first data point that showed a deletion.
The number of data points per platform is comparable at the location of the distal breakpoint of patient 1 and 3. In patient 1, the breakpoint mapping of all platforms is concordant. In contrast, there is a huge difference in breakpoint mapping in patient 3. According to the results obtained by Agilent platform, the distal breakpoint of the deletion is located 290-260K outside the most distal point of the picture (47,92 Mb) (green and red arrow). The results of the Affymetrix platform show that the deleted region starts more proximally at ~48.03Mb (black arrow). The beadchip of Illumina defines the distal breakpoint of the deletion between these two points. [See appendix: colour figures.]

most appropriate to implement for the screening copy number alteration, as this tool offers the simultaneous measurement of copy number changes and copy-neutral loss of heterozygosity (i.e uniparental disomy). On the other hand, the SNPs have been selected based on criteria such as heterozygosity, confirmation with Hardy-Weinberg equilibrium. Although these features are important for association studies, where SNPs need to be informative, they are less critical for copy number analysis where even spacing is more important. Indeed, many regions prone to rearrangements (e.g. duplicons) are lacking or are underrepresented on these arrays, as the associated SNPs did not meet the required quality criteria. This is in contrast to array-CGH in which the location of the oligonucleotides is not limited to known SNPs, and, therefore, it is possible to analyse regions of the genome where no validated SNPs are available.[24] Calculating the spacing between the consecutive data points per platform within chromosome region 2p16.2-p21, shows that the median spacing of genechip was 2.40 kb, with a maximum of 65.10 kb, the median spacing of the beadchip was 4.57 kb (with a maximum of 71.85 kb) and finally, that of Agilent using the 244K was 9.85 kb (with

a maximum of 47.40 kb). Thus, although the median spacing of Agilent is the largest (as it is gene-based), this platform might be the most valuable tool for investigating CNVs, depending on the genome region of interest (based on the maximum spacing of the three platforms).

Indeed, a previous study[24] has shown that in addition to the SNP-arrays, array-CGH analysis is required to cover all CNV regions in the human genome, with at least one third of CNVs >50 kb otherwise being missed. New arrays of both Affymetrix and Illumina are closing this gap by combining both SNP- and non-SNP probes on one array.

The beadchip has several clear advantages over the genechip, such as a higher SNP call rate, which is important when the expected size of the CNV is small. In our study, about 5-10%, sometimes even more, of all SNPs on the Affymetrix platform could not be scored (data not shown), resulting in a significant reduction of its resolution. Of course, the cause of such reduction might lie in a suboptimal quality of the DNA, however the identical DNA was used on the Illumina arrays. In addition, the genechip needed two arrays (this experiment) for a resolution comparable to that of the beadchip, which is especially of interest for the detection of small CNVs, and nearly all steps of the Illumina protocol can be automated. At the time these experiments were performed, only Illumina provided customer friendly software. Recently, however, software enabling easy calculation of the data generated by Affymetrix has become available, demonstrating the fast adaptation of products and application within this field. An important argument in favor of the genechip is the fact that they have started to validate these arrays to allow implementation in a diagnostic setting.

Looking at the breakpoints of the four patients, it can be concluded that the deletion of patients 1, 2 and 4 includes both the *MSH2* and the *MSH6* genes. The distal breakpoint of patient 3 is localised within or nearby (depending on the platform applied) the *MSH6* gene; the *MSH2* gene is not deleted in patient 3. This means that at least three out of four patients have a twenty fold increased chance of developing colon cancer or other Lynch syndrome-related tumors[25] compared to the healthy population. For this reason, it is of high clinical interest to diagnose the breakpoints of interstitial 2p deletions. However, when comparing the phenotype of the four patients (table 3), the only feature in common is mental retardation, which is a non-specific feature of nearly all chromosome anomalies. The lack of a common phenotype could be due to the different ages at observation (table 3) of the patients and the difference in size of the deletion.

Table 3. Overview of clinical features of the four patients with different sizes of interstitial 2p deletions.

	Patient 1	**Patient 2**	**Patient 3**	**Patient 4**
Localisation	2p16.2-p21	2p16.2-p21	2p16.3	2p16.3-p21
Cytogenetically visible	Yes	Yes	No	No
De novo	Yes	Yes	N.D.	Father not tested
Age of examination	6, 13, 36 months	5, 13.6, 17 month	7 years	37 years
Psychomotor delay	**Present**	**Present**	**Present**	**Present**
Length	Short stature (-2 SD)	Tall stature	75th percentile	Short stature
Weight	10th percentile	95th percentile	50-75th percentile	> 97th percentile
Size of skull	< 25th percentile	Microcephaly	50th percentile.	< 25th percentile
Shape of skull	Flattening of the occipital region	Flattening of posterior parietal region	Turricephaly	Brachycephaly with narrow forehead
Others	Aorta descendens P97. Palatoschisis Cataract, Nystagmus, Strabismus convergens	ASD Mild astigmatism Hypothelorism	Joint hyperextensibility with tendency to dislocation, High arched, narrow palate oblong face, large mouth, thin upper lip	
Colon cancer	Too young	Too young	Too young	Yes

N.D.: not determined

The minimum region of deletion overlap is localised between the distal breakpoint of patient 3 and the proximal breakpoint of patient 4. This region is ~1.3 Mb in size and encompasses 8 genes, from *MSH6* to *FSHR*. So far, this region was not found altered among healthy individuals.[24] The *FOXN2* gene, located between *MSH6* and *FSHR* might be of interest in relation to the phenotype of the patients. It is known that deregulation of FOX family genes can lead to congenital disorders in addition to its involvement in several types of cancer. Furthermore, the *FBXO11* gene coding for F-box protein family, might be involved in some of the developmental anomalies, as it related to phosphorylation-dependent ubiquitination. Mutations within the *LHCGR* and the *FSHR* genes are related to aberrant external and/or internal genital organs. No mutations with specific pathogenetic consequences have been reported for the remaining two genes (*CCDC128, STON1*).

Recently, the whole genome of Nobel laureate Jim Watson was sequenced (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi), revealing as much as 600,000 single nucleo-

tide variants that had not been reported before. The cost involved of this project was substantial and therefore this way of screening the human genome is not applicable on large scale yet. It can be expected, however, that affordable sequence-based whole genome genotyping will become possible within the coming two years. As a result, SNP typing and array-CGH will be superseded fairly soon by next generation sequencing. The first step towards the implementation of genome wide sequencing in a diagnostic setting would be to type "harmless" variations in a large group of normal individuals, since on average 1 in 1000 nucleotide on the human genome of a healthy individual varies. In addition, screening large cohorts of affected individuals with well-defined clinical features is essential to be able to interpret this new data.[26]

References

1. Caspersson T, Lomakka G, Zech L. The 24 fluorescence patterns of the human metaphase chromosomes - distinguishing characters and variability. *Hereditas* 1972: **67**(1):89-102.
2. Yunis JJ. High resolution of human chromosomes. *Science* 1976: **191**(4233):1268-1270.
3. Landegent JE, Jansen in dW, van Ommen GJ, Baas F, de Vijlder JJ, Van Duijn P, Van der PM. Chromosomal localization of a unique gene by non-autoradiographic in situ hybridization. *Nature* 1985: **317**(6033):175-177.
4. Ried T, Mahler V, Vogt P, Blonden LAJ, Van Ommen GJB, Cremer T, Cremer M. Direct carrier detection by in situ suppression hybridization with cosmid clones of the Duchenne/Becker muscular dystrophy locus. *Hum Genet* 1990: **85**:581-586.
5. Raap AK, Van De Corput MPC, Vervenne RAW, Van Gijlswijk RPM, Tanke HJ, Wiegant J. Ultra-sensitive FISH using peroxydase mediated deposition of biotin- or fluorochrome tyramides. *Hum Mol Genet* 1995: **4**:529-534.
6. Liehr T, Starke H, Heller A, Kosyakova N, Mrasek K, Gross M, Karst C, Steinhaeuser U, Hunstig F, Fickelscher I, Kuechler A, Trifonov V, Romanenko SA, Weise A. Multicolor fluorescence in situ hybridization (FISH) applied to FISH-banding. *Cytogenet Genome Res* 2006: **114**(3-4):240-244.
7. Liehr T, Starke H, Weise A, Lehrer H, Claussen U. Multicolor FISH probe sets and their applications. *Histol Histopathol* 2004: **19**(1):229-237.
8. 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**(4):399-407.
9. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998: **20**(2):207-211.
10. Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, Law S, Myambo K, Palmer J, Ylstra B, Yue JP, Gray JW, Jain AN, Pinkel D, Albertson DG. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001: **29**(3):263-264.
11. Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen TH, Girard L, Minna J, Christiani D, Leo C, Gray JW, Sellers WR, Meyerson M. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 2004: **64**(9):3060-3071.
12. Scott DA, Klaassens M, Holder AM, Lally KP, Fernandes CJ, Galjaard RJ, Tibboel D, de Klein A, Lee B. Genome-wide oligonucleotide-based array comparative genome hybridization analysis of non-isolated congenital diaphragmatic hernia. *Hum Mol Genet* 2007: **16**(4):424-430.
13. Herr A, Grutzmann R, Matthaei A, Artelt J, Schrock E, Rump A, Pilarsky C. High-resolution analysis of chromosomal imbalances using the Affymetrix 10K SNP genotyping chip. *Genomics* 2005: **85**(3):392-400.
14. Komura D, Shen F, Ishikawa S, Fitch KR, Chen W, Zhang J, Liu G, Ihara S, Nakamura H, Hurles ME, Lee C, Scherer SW, Jones KW, Shapero MH, Huang J, Aburatani H. Genome-wide detection of human copy number variations using high-density DNA oligonucleotide arrays. *Genome Res* 2006: **16**(12):1575-1584.
15. Knijnenburg J, Szuhai K, Giltay J, Molenaar L, Sloos W, Poot M, Tanke HJ, Rosenberg C. In-

sights from genomic microarrays into structural chromosome rearrangements. *Am J Med Genet* A 2005: **132**(1):36-40.

16. Vissers LE, De Vries BB, Osoegawa K, Janssen IM, Feuth T, Choy CO, Straatman H, Van D, V, Huys EH, Van Rijk A, Smeets D, Ravenswaaij-Arts CM, Knoers NV, Van DB, I, De Jong PJ, Brunner HG, van Kessel AG, Schoenmakers EF, Veltman JA. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. Am *J Hum Genet* 2003: **73**(6):1261-1270.
17. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 2001: **98**(1):31-36.
18. Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* 2004: **20**(8):1233-1240.
19. Friedman JM, Baross A, Delaney AD, Ally A, Arbour L, Asano J, Bailey DK, Barber S, Birch P, Brown-John M, Cao M, Chan S, Charest DL, Farnoud N, Fernandes N, Flibotte S, Go A, Gibson WT, Holt RA, Jones SJ, Kennedy GC, Krzywinski M, Langlois S, Li HI, McGillivray BC, Nayar T, Pugh TJ, Rajcan-Separovic E, Schein JE, Schnerch A, Siddiqui A, Van Allen MI, Wilson G, Yong SL, Zahir F, Eydoux P, Marra MA. Oligonucleotide microarray analysis of genomic imbalance in children with mental retardation. *Am J Hum Genet* 2006: **79**(3):500-513.
20. Sanders SR, Dawson AJ, Vust A, Hryshko M, Tomiuk M, Riordan D, Prasad C. Interstitial deletion of chromosome 2p16.2p21. *Clin Dysmorphol* 2003: **12**(3):183-185.
21. Lucci-Cordisco E, Zollino M, Baglioni S, Mancuso I, Lecce R, Gurrieri F, Crucitti A, Papi L, Neri G, Genuardi M. A novel microdeletion syndrome with loss of the MSH2 locus and hereditary non-polyposis colorectal cancer. *Clin Genet* 2005: **67**(2):178-182.
22. White S, Kalf M, Liu Q, Villerius M, Engelsma D, Kriek M, Vollebregt E, Bakker B, van Ommen GJ, Breuning MH, Den Dunnen JT. Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet* 2002: **71**(2):365-374.
23. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. Large-scale copy number polymorphism in the human genome. *Science* 2004: **305**(5683):525-528.
24. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature* 2006: **444**(7118):444-454.
25. Hendriks YM, de Jong AE, Morreau H, Tops CM, Vasen HF, Wijnen JT, Breuning MH, Brocker-Vriends AH. Diagnostic approach and management of Lynch syndrome (hereditary nonpolyposis colorectal carcinoma): a guide for clinicians. *CA Cancer J Clin* 2006: **56**(4):213-225.
26. Ropers HH. New perspectives for the elucidation of genetic disorders. *Am J Hum Genet* 2007: **81**(2):199-207.

Chapter IV

Discussion
Summary
Nederlandse Samenvatting

Chapter IV-1

Discussion

We are currently able to apply genome-wide screening tools with an unprecedented resolution to detect progressively smaller variants. It can be argued that this will improve the basis of genetic counselling significantly, as the probability of finding variants that may be related to the impairment of development and health in a patient will increase, and more information can be given about recurrence risks. It also enables us to verify assumptions that were made long before karyotyping and molecular diagnostic tools were invented. In these early days, geneticists have categorised large patient groups as having a multifactorial cause for their developmental delay. With the identification of variants that show a clearly detectable, but incomplete association with MR, one can now 'prove' on a molecular basis that the assumptions made were correct.

Although the identification of new variants is gratifying, it is accompanied by a progressively more difficult task for the people working in human and clinical genetics. After the introduction of karyotyping, a relatively small number of heteromorphisms (e.g. variants not related to human disease) were recognised and documented. This is in contrast to the current situation, where the number of variants with an unknown contribution to genomic disorders is huge. It has recently become clear that, by zooming in on the human genome using array-based platforms, variations exist at an unexpectedly high frequency among healthy individuals; as much as 12% of the human genome show CNVs that are probably not related to any clinical feature (Redon *et al.* 2006).

So, the more we learn about the human genome, the more we are confronted with questions about the implications of new findings. Does it involve a disease-causing alteration or is it a neutral variant?

In many reports the authors have only considered *de novo* variants to be causative. As soon as it became clear that one of the parents carried the same CNV, as the one detected in the affected child, it was thought to be a neutral variant. This is not always a correct assumption, as familial variants might be related to genomic disorders due to phenotypic variability (Ullmann *et al.* 2007), the presence of an autosomal recessive disorder (chapter III-2), or related to a deletion of an imprinted region that may be silent in one parent and disease-causing in the next generation. To complicate the picture even further, genetic disorders can also originate by a combination of two or more variations inherited from two parents, where each of which alone will not result in disease (Klopocki *et al.* 2007) (Lupski 2007). In addition, we can expect that some of the regions showing CNVs among healthy individuals contribute to genetic disease. This would indicate that CNVs present in regions described in the Human Variation database are not always neutral variants. In contrast, as pointed out in chapter I-5, the

finding of a *de novo* variation in an affected individual does not automatically mean that the alteration is disease causing. All these observations impact greatly on genetic counselling and this also underlines the main drawback of using the new platforms, as we sometimes lack the knowledge to adequately inform the patient and the family of the consequence of any finding. To resolve this, it is of great importance to collate CNV data in databases that are accessible to everyone working in this field. Two of such databases are available at this moment (ECARUCA and DECIPHER).

New tools for genome analysis reach the diagnostic laboratory at a quick pace. As a result, one can consider several technical approaches to help diagnose the patient with mental retardation and / or congenital malformation. In chapter II-3, we proposed a strategy in which MLPA covering the chromosome ends and regions related to micro deletions/ duplications should be used first, and if uninformative be followed by whole genome analysis. As pointed out by Rauch *et al.* (2006), this first step will detect an aberration in 5-20% of the MR patients, depending on the criteria used for selection. Since these rearrangements are also readily detected by currently available genome-wide screening tools (arrays), the use of these arrays as a first step now seems a more logical way to go, if it is possible to implement this in the diagnostic setting. Due to the necessity of guaranteeing Standard Operating Procedures in the diagnostic laboratories, it is often difficult to implement the most recent technologies that have proven to be efficient in a research setting. The rapid evolution of technology demands constant adaptation from both the clinician (who has to explain the outcome of the screening towards the patient) and the laboratory (validation and implementation of a new technique) in order to continue applying state of the art diagnostic methods.

At this moment, there is no golden standard available for determining which genome-wide screening platform provides the most relevant data for diagnostic purposes. The advantages of both CGH array based screening and high-density SNP genotyping have been discussed in section I.6.3.5. A recent study (Redon *et al.* 2006) has shown that in addition to the SNP-arrays, array-CGH analysis is required to cover all CNV regions in the human genome, with at least one third of CNVs >50 kb otherwise being missed. New arrays of both Affymetrix and Illumina are closing this gap by combining both SNP- and non-SNP probes on one array. In addition, Nimblegen now has a 42 M non-SNP array available enabling the detection of variants as small as 500 bps.

Although we already struggle to arrive at correct and comprehensive interpretation of high resolution array analysis in a diagnostic setting, a next generation of technical advance is approaching.

Recently, the whole genome of Nobel laureate Jim Watson was sequenced, revealing as many as 600,000 variants that had not been reported before. The cost involved of this project was substantial and therefore this way of screening the human genome is not yet applicable on large scale. It can be expected, however, that affordable sequence-based whole genome genotyping will become possible within the coming 2-5 years. As a result, SNP typing and array-CGH will be superseded fairly soon by next generation sequencing. The first step towards the implementation of genome wide sequencing would be increasing the knowledge about "harmless" variations in a large group of normal individuals, since on average 1 in 1000 nucleotides in the genome of a two healthy individuals varies. In addition, screening large cohorts of affected individuals with well-defined clinical features is essential to be able to interpret this new data (Ropers *et al*. 2007).

The possibility of 'reading' the whole human genome at the nucleotide level will also provide information about susceptibility for diseases that are not related to the patients' reason for consulting a specialist. This issue should be discussed with the patient or the parents during the counseling prior to genome-wide testing. One might choose to communicate only the variants that are thought to be causally related to the patients' phenotype or those that are well known to have a great potential influence on the patients' health (for example inactivation of tumor-suppressor genes). Two examples of alterations in tumor-suppressor genes detected after screening the human genome for MR-related CNVs are described in chapter III-2 and III-4. The patients in chapter II-2, carrying an interstitial 2p deletion, have a high chance of developing a HNPCC related tumor, as the deletion includes the *MSH6* gene. In chapter III-4, the mother and the maternal grandmother of the two index patients with Peters Plus syndrome were found to have a 1.5 Mb deletion encompassing a part of the *BRCA2* gene. Despite the fact that both women already developed breast cancer, they are now confronted with a high recurrence risk and a moderate increased risk of developing ovarian cancer. These 'side effects' of screening can't be avoided. However, a positive consequence of this knowledge is the fact that these patients can now be included in a screening program.

In summary, we can conclude that the plasticity of the genome creates a conundrum of Babylonic proportions. Nevertheless, it is expected that the implementation of new screenings technologies will give greater insight into a range of genetic diseases, and will hopefully lead to a better understanding of the many different causes of intellectual disability and congenital malformations.

Chapter IV-2

Summary

For decades, trypsin Giemsa banding of metaphase spreads has been the standard diagnostic method to detect chromosomal rearrangements. The method has several advantages; all chromosomes are seen under the microscope, and individual cells can be karyotyped, permitting clonal analysis (Caspersson, Lomakka, and Zech 1972; Yunis 1976). A major limitation is the fact that due to the contraction of chromosomes during metaphase, and the resolution of the light microscope, G banding is not capable of identifying rearrangements that are smaller than 3-5 Mb.

Fluorescence in Situ Hybridisation (FISH) (Landegent *et al*. 1985) (Ried *et al*. 1990) partly overcame this problem, allowing direct testing for the presence or absence of a specific genomic region. This method is especially used for the identification of micro deletion syndromes and subtelomeric rearrangements. It has several drawbacks though, as detecting rearrangements using FISH analysis is only possible when an obvious, specific phenotype is present that is recognized by a specialist, and when a specific FISH probe is available. Multi-colour methodologies have been developed (Knight *et al*. 1997) (Engels *et al*. 2003), however, the number of loci that can be analysed simultaneously remains limited.

Assays for gene copy number or gene dosage have long been utilized in the clinical molecular genetic laboratories. For many years, Southern blotting analysis (Southern 1975) followed by densitometry was the main assay available to assess for a small number of copy number variations. The development of real-time Polymerase Chain Reaction (PCR), Multiplex Amplifiable Probe Hybridisation (MAPH) (Armour *et al*. 2000) and Multiplex Ligation-dependent Probe Amplification (MLPA) (Schouten *et al*. 2002) allowed more widespread analysis of gene dosage. MAPH and MLPA are PCR-based methods to simultaneously determine the copy number of a large set, currently up to 60, of different chromosomal loci. The advantage of MAPH/MLPA compared to other techniques, such as FISH, is that the resolution of detection is limited only by the size of the probes used (100-500 bp) and it facilitates the parallel screening of several tens of patients at many different loci in one experiment.

Chapter II-1 describes the use of MAPH probe sets covering different genomic loci, including subtelomeric regions, regions involved in microdeletion syndromes and a set of genes evenly spread through out the rest of the genome. Using these probe sets, 184 mentally retarded patients were screened. Results included the detection of rearrangements in subtelomeric and pericentromeric regions, as well as several interstitial alterations, indicating that submicroscopic alterations with a higher frequency were not limited to the ends of the chromosomes.

In **Chapter II-2**, a MAPH assay was designed containing exon-specific single copy

sequences from within a selection of the 169 regions flanked by duplicons that were identified, at a first pass, in 2001. Subsequently, the frequency of chromosomal rearrangement among patients with mental retardation (MR) and/or congenital malformations (CM) was determined. The same study population was tested for rearrangements in regions with no known duplicons nearby, using a set of probes derived from function-selected genes. As expected, the alteration frequency per unit of DNA is much higher in regions flanked by duplicons (fraction of the genome tested: 5.2%) compared to regions without known duplicons nearby (fraction of the genome tested: 24.5%). Thus, the data supported the emerging hypothesis that regions flanked by duplicons are enriched for copy number variations.

Recently, technologies have been developed that provide both genome-wide and high resolution analysis. The first arrays used relatively large fragments of DNA (~150 kb) isolated from mainly Bacterial Artificial Chromosome (BAC) (Solinas-Toldo *et al.* 1997; Pinkel *et al.* 1998; Snijders *et al.* 2001). A newer format uses oligonucleotide probes of 25 to 60 nt in length. Due to the smaller size of these probes and the much larger number of loci analysed, it is possible to detect much smaller copy number variations (CNVs) compared to the variations that can be revealed using BAC-PAC clone array. The 25-mer probe arrays were originally designed for use in genome-wide SNP analysis, for linkage and association studies. As successive arrays have provided ever increasing coverage, currently up to 500-1000K, they were quickly used to estimate copy number changes by using both the signal strength and SNP score. The 60 nt oligo based arrays give stronger signal intensity and CNVs can therefore be detected using solely the signal intensity.

To assess the value of MLPA and array based techniques in clinical diagnosis, 58 developmentally delayed (DD) patients with a normal karyotype were independently tested with both array-CGH and MLPA. The results are described in **chapter II-3**. It shows that both methods are effective and represent an improvement to classical and molecular cytogenetics, as currently applied. Considerations balancing cost-efficiency and complexity promote a format where MLPA screening precedes array-CGH analysis. In addition, an alternative screening strategy, encompassing MLPA testing prior to karyotyping for unselected samples is described. However, at this moment, whole genome array analysis has become affordable, making MLPA analysis as a first step not necessary anymore.

Chapter III-1 highlights the value of using different genomic approaches to unravel chromosomal alterations and their phenotypic impact. Albeit was possible to identify a previously unreported rearrangement within the 22q11 region, e.g. a dele-

tion of the proximal part of chromosome band 22q11. It is argued that a deletion in this chromosome band is unrelated to the phenotypic trait seen in the index patient. The other chromosome 22 homologue carries a duplication of the Velocardiofacial/DiGeorge syndrome (VCFS/DGS) region. In addition, a previously undescribed deletion of 22q12.1, located in a relatively gene-poor region, was identified.

Chapter III-2 describes the identification of the gene involved in Peters Plus syndrome. This shows how the application of new techniques leads to the identification of the pathogenetic mutations of this autosomal recessive syndrome.

Chapter III-3 shows that MLPA testing is capable of fine mapping the breakpoints of different interstitial 16p deletions. As the clinical features are not very distinctive, the screening of the ATR-16 region in patients presenting with mild to moderate MR and microcytic hypochromic anemia with normal ferritin levels is proposed.

Finally, four platforms (a 3 K BAC clone array, a 44 K microarray and two SNP-based arrays (250-500K, and 317K) have been compared for their ability to identify the breakpoints in four patients with different sizes of interstitial 2p deletions, all localised within 2p16.1-p21(**chapter II-4**). All platforms identified the deletion and the results were comparable cross-platform.

It is evident that the genome-wide high resolution arrays provide an enormous improvement of the resolution of chromosome analysis. However, preliminary studies indicate that the extent of 'normal' copy number variation in the human genome may amount to at least 12 % (Redon *et al*. 2006). This percentage is still rather small compared to the number of variants that will be revealed using next generation sequencing. This sequence-based whole genome genotyping of a patient will soon be possible on a large scale. Consequently, the problem we are now facing with CNVs based on the outcome of array screening will be amplified substantially. Collecting data of sequence variation in very large groups of healthy individuals as well as of well-characterised patients will be needed to understand the results in the near future.

Thus, for each apparent rearrangement detected, it is necessary to determine its phenotypic consequences.

Chapter IV-3

Nederlandse samenvatting

In 1956 werd door Tjio en Levan het correcte aantal chromosomen in een menselijke cel gepubliceerd (n = 46). Op basis van deze bevinding werd een techniek ontwikkeld om chromosomen nader te onderzoeken; karyotypering met behulp van de de lichtmicroscoop (Caspersson, Lomakka, and Zech 1972; Yunis 1976). Een andere belangrijke doorbraak was de ontdekking van de Fluorescent *In Situ* Hybridisation (FISH) techniek (Ried *et al*. 1990; Landegent *et al*. 1985). Dit maakte het mogelijk om gericht relatief kleine veranderingen in het erfelijk materiaal van de mens te identificeren. Het werd echter duidelijk dat deze microscopische technieken beperkingen kennen, arbeidsintensief en kostbaar zijn. De belangrijkste beperking is dat veranderingen in het erfelijk materiaal kleiner dan 5-10 miljoen bouwstenen (= megabasen=Mb) zonder duidelijke specifieke klinische kenmerken bij een patiënt niet kunnen worden gediagnosticeerd. In de afgelopen jaren is een scala aan moleculaire technieken ontwikkeld met een hogere resolutie in vergelijking met karyotypering. Aanvankelijk gaven multicolour en multiprobe FISH uitkomst, echter deze technieken zijn niet in staat om veranderingen in het erfelijk materiaal kleiner dan ~2 Mb op te sporen (figuur 3). Southern blotting (Southern 1975) en Pulse field gel electrophoresis (PFGE) (van Ommen *et al*. 1986; Den Dunnen *et al*. 1987) zijn wel in staat deze submicroscopische veranderingen te detecteren, echter zij zijn arbeidsintensief en hebben een lage doorvoersnelheid. In 2000 en 2002 werden, respectievelijk, Multiplex Ampliable Probe Hybridisation (MAPH) (Armour *et al*. 2000) en Multiple Ligation-dependent Probe Amplification (MLPA) (Schouten *et al*. 2002) technieken geïntroduceerd. Deze, op kwantitatieve PCR-gebaseerde, technieken zijn in staat om met een zeer hoge resolutie (overeenkomend met de probe lengte ~100-500 baseparen) meerdere plaatsen op het genoom te testen op de aanwezigheid van kopie-verschillen bij 96 patiënten in één reactie.

Hoofdstuk II-1 en hoofstuk II-2 beschrijven twee studies waar gebruik is gemaakt van de MAPH techniek. De MLPA techniek is in deze studies gebruikt voor de verificatie van de gevonden veranderingen. In **hoofstuk II-1** worden de 'chromosoom-eind' en interstitiële veranderingen (verandering binnen het chromosoom) samengevat, die gevonden zijn na het testen van 184 patiënten met een verstandelijke beperking. Ruim 4% van deze studiepopulatie had een verandering aan het einde van de chromosomen. Een onderverdeling in patiënten met een verstandelijke beperking met en zonder aanvullende dysmorfe / aangeboren afwijkingen resulteerde in de conclusie dat de kans op het vinden van veranderingen aan het einde van de chromosomen het grootst is bij patiënten met aanvullende afwijkingen. Deze bevinding is in overeenkomst met data uit de literatuur. Daarnaast werd met een relatief klein aantal geteste interstitiële gebieden (n = 112, inclusief gebieden die gerelateerd zijn aan microdeletie syndromen), zeven

veranderingen gedetecteerd. Dit bevestigde het idee dat het voorkomen van submicroscopische veranderingen in het erfelijk materiaal niet beperkt is tot de uiteinden van de chromosomen, maar dat overal langs de armen van de chromosomen afwijkingen kunnen ontstaan.

Hoofdstuk II-2 beschrijft de bevindingen van het testen van stukken genoom die geflankeerd worden door segmentale duplicaties. Dit zijn stukken erfelijk materiaal met een zeer hoge homologie (volgorde van de bouwstenen zijn vrijwel gelijk), waardoor ongelijke paring gevolgd door ongelijk 'overstappen' tot kopie-verschillen kunnen leiden (figuur 2A). Op basis van onze bevinding kon geconcludeerd worden, overeenkomend met de data uit de literatuur, dat kopie-verschillen vaker voorkomen tussen deze zogenaamde homologe gebieden dan elders in het genoom.

Ondanks dat MAPH, maar vooral MLPA momenteel wordt toegepast in meerdere, vooral Europese, diagnostische laboratoria voor het opsporen van veranderingen in vele verschillende genen, zijn zij niet in staat om genoom-breed te screenen op de aanwezigheid van mogelijke kopie-verschillen. Array-gebaseerde technieken (BAC-, oligo- en SNP arrays) zijn wel in staat om in één proef het gehele erfelijk materiaal van een patiënt te testen, waarbij de resolutie afhangt van wat aangebracht is op de array. De resolutie van deze technieken neemt steeds verder toe. Recent zijn SNP-gebaseerde opsporingstechnieken beschikbaar gekomen. Dit maakt het ni*et al*leen mogelijk om naar kopie-verschillen te zoeken, maar ook naar verlies van heterozygositeit (diversiteit in het erfelijk materiaal) of naar niet-Mendeliaanse overerving te kijken. Gezien het feit dat het toepassen van karyotypering en genoombrede technieken met een hoge resolutie aanvankelijk relatief duur waren, is in **hoofdstuk II-3** een alternatieve manier van testen voorgesteld. Deze houdt in dat met behulp van MLPA, de plaatsen op het erfelijk materiaal getest worden, waarvan bekend is dat ze frequent veranderingen laten zien (bij een bepaalde studiepopulatie), alvorens genoombreed getest wordt. Karyotypering wordt alleen verricht voor een geselecteerde patiëntengroep die bij MLPA en genoombrede technieken geen verandering liet zien of voor het uitsluiten van een Robertsoniaanse translocatie (versmelting tussen de centromeren van twee chromosomen, die geen functionele korte arm hebben). Ondertussen zijn de kosten van array onderzoek substantiëel gedaald, waardoor de MLPA stap vóór het uitvoeren van array gebaseerde technieken niet meer noodzakelijk is.

Op basis van de resultaten die beschreven staan in **hoofstuk III-1** kan geconcludeerd worden dat verschillende technieken, zoals MAPH/MLPA-, FISH analyse en array gebaseerde technieken, elkaar aanvullen in plaats van dat ze 'concurrenten' zijn. In dit hoofdstuk wordt duidelijk dat de verschillende aspecten van een gecompliceerde

herrangschikking op een chromosoom slechts gedefiniëerd kon worden door het toepassen van meerdere technieken. Deze complexe herrangschikking bleek te bestaan uit een deletie en een duplicatie in het 22q11 gebied op twee verschillende chromosomen 22, gecombineerd met een tweede deletie die verderop op de lange arm van het chromosoom was gelocaliseerd. Daarnaast werd op basis de bevindingen in deze studie beargumenteerd dat een deletie dichtbij het centromeer van chromosoom 22 (het Cat-eye syndroom gerelateerd gebied), waarschijnlijk niet gerelateerd is aan een klinisch beeld.

In **hoofdstuk III-2** beschrijven wij hoe de toepassing van een hoge resolutie techniek (array-CGH) heeft geleid tot de identificatie van de oorzaak van het Peters Plus syndroom, een zeldzame ernstige aandoening. Dit is de eerste autosomaal recessieve aandoening die is opgelost door toepassing van array-CGH.

Hoofdstuk III-3 beschrijft het inzoomen van een gebied op de korte arm van chromosoom 16 dat verantwoordelijk is voor het ATR-16 syndroom (Alpha Thalassemie Retardatie syndroom, waarvan de oorzaak op het 16^{e} chromosoom is gelegen). Met behulp van 3 kleuren MLPA werd het ATR-16 gerelateerde gebied nader gespecificeerd. Aangezien de klinische kenmerken van de ATR-16 patiënten weinig specifiek zijn, wordt aangeraden om bij een patiënt met een verstandelijke beperking en bloedarmoede een eenvoudig hematologische test te laten verrichten. In geval van een microcytaire hypochrome anemie (specifieke vorm van bloedarmoede) met een normaal ijzer gehalte kan gericht moleculair diagnostisch onderzoek (MLPA) naar ATR-16 worden aangevraagd.

De toepassing van vier hoge resolutie technieken voor de identificatie van de breekpunten in vier verschillende patiënten met overlappende deleties op de korte arm van chromosoom 2 is beschreven in **hoofdstuk III-4**. De resultaten van de verschillende technieken waren vergelijkbaar. Door de toepassing van de nieuwe hoge resolutie technieken wordt de resolutie van de chromosoom analyse sterk verbeterd. Echter, de eerste publicaties benadrukken het frequente voorkomen van kleine kopie-verschillen bij gezonde mensen (Iafrate *et al*. 2004; Sebat *et al*. 2004; Redon *et al*. 2006). Met behulp van array-CGH en SNP arrays werd vastgesteld dat geveer 12% van het humane genoom 'onschuldige' kopie-verschillen kan bevatten. Dit is nog maar een klein deel van de variatie die in het humane genoom wordt aangetroffen bij vergelijking op sequentie-niveau. Recent is de volgorde van de bouwstenen van het erfelijk materiaal van één persoon gepubliceerd, namelijk die van de Nobelprijswinnaar James Watson. Dit onderzoek leverde 600.000 niet eerder gerapporteerde veranderingen op. Dit illustreert dat genoombreed sequencen (het bepalen van de volgorde van het gehele erfelijk

materiaal van de mens) het probleem van de interpretatie van de resultaten bij mensen met een aandoening exponentiëel zal vergroten ten opzichte van de 'onbekende' veranderingen waar we nu mee geconfronteerd worden. Het is daarom van zeer groot belang om eerst veel kennis op te doen over de variaties in het erfelijk materiaal bij grote groepen gezonde mensen en daarnaast over variaties die voorkomen in patiënten met een goed gedefiniëerd klinisch beeld (Ropers 2007).

Bij vele patiënten wordt nu een oorzaak gevonden voor hun verstandelijke beperking, waar dit vroeger niet mogelijk was. Om echter alle gegevens, die door de nieuwe technieken beschikbaar komen, goed te interpreteren, is veel werk nodig. Uiteindelijk kan onze kennis van het menselijke genoom zodanig toenemen dat wij per bouwsteen of in elk geval per gen weten of dit een rol speelt in de ontwikkeling van ons verstand.

Curriculum Vitae

Naam: Marjolein Kriek
Geboortedatum: 22-11-1973
Geboorte plaats: Leiden (Academisch Ziekenhuis Leiden)

School

Eindexamen atheneum aan het Visser 't Hooft lyceum te Leiden (1992).

Studies

- 18 August 2000
  Behalen van de artsenbul aan de Universiteit Leiden.
- 17 september 2002
  Doctoraal examen van de studie Biomedische Wetenschappen aan de Universiteit Leiden.

Wetenschappelijk onderzoek

1995 Zes maanden stage bij vakgroep Moleculaire Carcinogenese aan Universiteit Leiden o.l.v. Prof. Dr van der Eb en Dr Zantema.
Titel onderzoek:
"*Association of proteins influenced by the Adenovirus E1A oncoprotein".*

1998 Drie maanden stage bij vakgroep Klinische Epidemiologie in het L.U.M.C. o.l.v. Prof. Dr Roosendaal en Drs Sramek.
Titel onderzoek:
"Mortality in carriers of Hemophilia".
Dit onderzoek leidde tot een tweede auteurschap in de Lancet.

2001 Eindvakstage Biomedische Wetenschappen (9 maanden)
bij de vakgroep Humane en Klinische Genetica o.l.v. Prof. Breuning
Titel onderzoek:
"*Screening for mutations in mentally retarded patients using MAPH".*
Dit onderzoek vormde de basis van het huidige proefschrift.

2002 Begonnen aan promotie onderzoek getiteld; “The human genome; you gain some, you lose some”, onder leiding van Prof. M.H.Breuning, Prof. G-J B. Van Ommen en dr. J.T. den Dunnen: Aanvankelijk als AGNIO, vanaf 1 januari 2003 is dit omgezet in een AGIKO traject op basis van ZONMW-subsidie (AGIKO-fellowship 940-37-032).

Klinische ervaring

2000 Half jaar als AGNIO gewerkt op de afdeling Klinische Genetica (LUMC)

1 april 2005 tot heden
In opleiding tot klinisch geneticus op de afdeling Klinische Genetica (LUMC)

List of publications

2002

White S, Kalf M, Liu Q, Villerius M, Engelsma D, Kriek M, Vollebregt E, Bakker B, van Ommen GJ, Breuning MH *et al*. Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet. 2002 Aug;71(2):365-74.*

2003

Sramek A, Kriek M, Rosendaal FR. Decreased mortality of ischaemic heart disease among carriers of haemophilia.
Lancet. 2003 Aug 2;362(9381):351-4

2004

Kriek M, White SJ, Bouma MC, Dauwerse HG, Hansson KB, Nijhuis JV, Bakker B, van Ommen GJ, den Dunnen JT, Breuning MH. Genomic imbalances in mental retardation. *J Med Genet. 2004 Apr;41(4):249-55*

White SJ, Vink GR, Kriek M, Wuyts W, Schouten J, Bakker B, Breuning MH, den Dunnen JT. Two-color multiplex ligation-dependent probe amplification: detecting genomic rearrangements in hereditary multiple exostoses.
Hum Mutat. 2004 Jul;24(1):86-92.

2006

Rosenberg C, Knijnenburg J, Bakker E, Vianna-Morgante AM, Sloos W, Otto PA, Kriek M, Hansson K, Krepischi-Santos AC, Fiegler H, Carter NP, Bijlsma EK, van Haeringen A, Szuhai K, Tanke HJ. Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents.
J Med Genet. 2006 Feb;43(2):180-6.

Kriek M, White SJ, Szuhai K, Knijnenburg J, van Ommen GJ, den Dunnen JT, Breuning MH. Copy number variation in regions flanked (or unflanked) by duplicons among patients with developmental delay and/or congenital malformations; detection of reciprocal and partial Williams-Beuren duplications.
Eur J Hum Genet. 2006 Feb;14(2):180-9

van der Knaap MS, Kriek M, Overweg-Plandsoen WC, Hansson KB, Madan K, Starreveld JS, Schotman-Schram P, Barkhof F, Lesnik Oberstein SA. Cerebral white matter abnormalities in 6p25 deletion syndrome.
AJNR Am J Neuroradiol. 2006 Mar;27(3):586-8

Kriek M, Szuhai K, Kant SG, White SJ, Dauwerse H, Fiegler H, Carter NP, Knijnenburg J, den Dunnen JT, Tanke HJ, Breuning MH, Rosenberg C. A complex rearrangement on chromosome 22 affecting both homologues; haplo-insufficiency of the Cat eye syndrome region may have no clinical relevance.
Hum Genet. 2006 Aug;120(1):77-84.

Lesnik Oberstein SA, Kriek M, White SJ, Kalf ME, Szuhai K, den Dunnen JT, Breuning MH, and Hennekam RC. Peters Plus Syndrome Is Caused by Mutations in *B3GALTL,* a Putative Glycosyltransferase.
Am J Hum Genet. 2006 Aug; 79(3):562-6.

Rosenberg C, Krepischi-Santos ACV, Knijnenburg J, Kok F, Otto PA, Tanke HJ, Kriek M, Zangrande Vieira LC, Nascimento RMP, Vianna-Morgante AM. X-chromosome segmental imbalances as a cause of recessive mental retardation syndromes.
J Med Genet. 2006 Feb;43(2):180-6.

2007

Kant SG, Kriek M, Walenkamp MJE, Hansson KBM, van Rhijn A, Clayton-Smith J, Wit JM, Breuning MH. Tall stature and duplication of the insulin-like growth factor I receptor gene.
Eur J Med Genet. 2007 Jan-Feb;50(1):1-10.

Kriek M, Konijnenburg J, White SJ, Rosenberg C, den Dunnen JT, van Ommen GJ, Tanke HJ, Breuning MB, Szuhai K. Diagnosis of genetic abnormalities in developmentally delayed patients: A new strategy combining MLPA and Array-CGH.
Am J Med Genet A. 2007 Mar 15;143(6):610-4.

Harteveld CL, Kriek M, Bijlsma EK, Erjavec Z, Balak D, Phylipsen M, Voskamp A, di Capua E, White SJ and Giordano PC.Telomeric deletions of 16p causing alpha-thalassemia and mental retardation characterized by multiplex ligation-dependent probe amplification.
Human Genet. 2007 Jun 28; [Epub ahead of print]

Kriek M, Ruivenkamp CAL, Ariyurek Y, Kalf ME, Knijnenburg J, van Haeringen A, Genuardi M, Rosenberg C, Sanders SR., White SJ, Szuhai K, Breuning MH, den Dunnen JT. Comparison of four genome-wide platforms using overlapping interstitial 2p alterations.
Submitted

References

Aitman TJ, Dong R, Vyse TJ, Norsworthy PJ, Johnson MD, Smith J, Mangion J, Roberton-Lowe C, Marshall AJ, Petretto E, Hodges MD, Bhangal G, Patel SG, Sheehan-Rooney K, Duda M, Cook PR, Evans DJ, Domin J, Flint J, Boyle JJ, Pusey CD, Cook HT (2006) Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* **439**:851-855

Amos-Landgraf JM, Ji Y, Gottlieb W, Depinet T, Wandstrat AE, Cassidy SB, Driscoll DJ, Rogan PK, Schwartz S, Nicholls RD (1999) Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. *Am J Hum Genet* **65**:370-386

Anderlid BM, Schoumans J, Anneren G, Sahlen S, Kyllerman M, Vujic M, Hagberg B, Blennow E, Nordenskjold M (2002) Subtelomeric rearrangements detected in patients with idiopathic mental retardation. *Am J Med Genet* **107**:275-284

Armengol L, Pujana MA, Cheung J, Scherer SW, Estivill X (2003) Enrichment of segmental duplications in regions of breaks of synteny between the human and mouse genomes suggest their involvement in evolutionary rearrangements. *Hum Mol Genet* **12**:2201-2208

Armour JA, Sismani C, Patsalis PC, Cross G (2000) Measurement of locus copy number by hybridisation with amplifiable probes. *Nucl Acids Res* **28**:605-609

Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE (2002a) Recent segmental duplications in the human genome. *Science* **297**:1003-1007

Bailey JA, Liu G, Eichler EE (2003) An Alu transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet* **73**:823-834

Bailey JA, Yavor AM, Viggiano L, Misceo D, Horvath JE, Archidiacono N, Schwartz S, Rocchi M, Eichler EE (2002b) Human-specific duplication and mosaic transcripts: the recent paralogous structure of chromosome 22. *Am J Hum Genet* **70**:83-100

Baker E, Hinton L, Callen DF, Altree M, Dobbie A, Eyre HJ, Sutherland GR, Thompson E, Thompson P, Woollatt E, Haan E (2002) Study of 250 children with idiopathic mental retardation reveals nine cryptic and diverse subtelomeric chromosome anomalies. *Am J Med Genet* **107**:285-293

Bayes M, Magano LF, Rivera N, Flores R, Perez Jurado LA (2003) Mutational mechanisms of Williams-Beuren syndrome deletions. *Am J Hum Genet* **73**:131-151

Bi W, Park SS, Shaw CJ, Withers MA, Patel PI, Lupski JR (2003) Reciprocal crossovers and a positional preference for strand exchange in recombination events resulting in deletion or duplication of chromosome 17p11.2. *Am J Hum Genet* **73**:1302-1315

Biesecker LG (2002) The end of the beginning of chromosome ends. *Am J Med Genet* **107**:263-266

Blonden LAJ, Grootscholten PM, Den Dunnen JT, Bakker E, Abbs SJ, Bobrow M, Boehm C *et al*. (1991) 242 breakpoints in the 200-kb deletion-prone P20 region of the DMD-gene are widely spread. *Genomics* **10**:631-639

Bocian E, Helias-Rodzewicz Z, Suchenek K, Obersztyn E, Kutkowska-Kazmierczak A, Stankiewicz P, Kostyk E, Mazurczak T (2004) Subtelomeric rearrangements: results from FISH studies in 84 families with idiopathic mental retardation. *Med Sci Monit* **10**:CR143-CR151

Bonifacio S, Centrone C, Da Prato L, Scordo MR, Estienne M, Torricelli F (2001) Use of primed in situ labeling (PRINS) for the detection of telomeric deletions associated with mental retardation. *Cytogenet Cell Genet* **93**:16-18

Borgione E, Giudice ML, Galesi O, Castiglia L, Failla P, Romano C, Ragusa A, Fichera M (2001) How

microsatellite analysis can be exploited for subtelomeric chromosomal rearrangement analysis in mental retardation. *J Med Genet* **38**:E1

Boue A, Boue J (1977) [Role of chromosome abnormalities in reproduction failures]. *J Gynecol Obstet Biol Reprod*(Paris) **6**:5-21

Breuning MH, Dauwerse HG, Fugazza G, Saris JJ, Spruit L, Wijnen H, Tommerup N, van der Hagen CB, Imaizumi K, Kuroki Y, . (1993) Rubinstein-Taybi syndrome caused by submicroscopic deletions within 16p13.3. *Am J Hum Genet* **52**:249-254

Bruder CE, Hirvela C, Tapia-Paez I, Fransson I, Segraves R, Hamilton G, Zhang XX *et al.* (2001) High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH. *Hum Mol Genet* **10**:271-282

Bugge M, Bruun-Petersen G, Brondum-Nielsen K, Friedrich U, Hansen J, Jensen G, Jensen PK, Kristoffersson U, Lundsteen C, Niebuhr E, Rasmussen KR, Rasmussen K, Tommerup N (2000) Disease associated balanced chromosome rearrangements: a resource for large scale genotype-phenotype delineation in man. *J Med Genet* **37**:858-865

Bundey S, Thake A, Todd J (1989) The recurrence risks for mild idiopathic mental retardation. *J Med Genet* **26**:260-266

Butler MG (1995) High resolution chromosome analysis and fluorescence in situ hybridization in patients referred for Prader-Willi or Angelman syndrome. *Am J Med Genet* **56**:420-422

Buzhov BT, Lemmers RJ, Tournev I, Dikova C, Kremensky I, Petrova J, Frants RR, Van Der Maarel SM (2005) Genetic confirmation of facioscapulohumeral muscular dystrophy in a case with complex D4Z4 rearrangments. *Hum Genet* **116**:262-266

Carr DH (1971) Chromosomes and abortion. *Adv Hum Genet* **2**:201-257

Caspersson T, Lomakka G, Zech L (1972) The 24 fluorescence patterns of the human metaphase chromosomes - distinguishing characters and variability. *Hereditas* **67**:89-102

Chance PF, Abbas N, Lensch MW, Pentao L, Roa BB, Patel PI, Lupski JR (1994) Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Hum Mol Genet* **3**:223-228

Chen KS, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault AC, Lee CC, Lupski JR (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat Genet* **17**:154-163

Cheung J, Estivill X, Khaja R, MacDonald JR, Lau K, Tsui LC, Scherer SW (2003a) Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biol* **4**:R25

Cheung J, Estivill X, Khaja R, MacDonald JR, Lau K, Tsui LC, Scherer SW (2003b) Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biol* **4**:R25

Cheung VG, Nowak N, Jang W, Kirsch IR, Zhao S, Chen XN, Furey TS *et al.* (2001) Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* 409:953-958

Clarkson B, Pavenski K, Dupuis L, Kennedy S, Meyn S, Nezarati MM, Nie G, Weksberg R, Withers S, Quercia N, Teebi AS, Teshima I (2002) Detecting rearrangements in children using subtelomeric FISH and SKY. *Am J Med Genet* **107**:267-274

Coe BP, Ylstra B, Carvalho B, Meijer GA, MacAulay C, Lam WL (2007) Resolving the resolution of array CGH. *Genomics*

Colleaux L, Rio M, Heuertz S, Moindrault S, Turleau C, Ozilou C, Gosset P, Raoult O, Lyonnet S, Cormier-Daire V, Amiel J, Le Merrer M, Picq M, de Blois MC, Prieur M, Romana S, Cornelis

F, Vekemans M, Munnich A (2001) A novel automated strategy for screening cryptic telomeric rearrangements in children with idiopathic mental retardation. *Eur J Hum Genet* **9**:319-327

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

Dauwerse JG, Wiegant JCAG, Raap AK, Breuning MH, Van Ommen GJB (1992) Multiple colors by fluorescence in situ hybridization using ratio-labelled DNA probes create a molecular karyotype. *Hum Mol Genet* **1**:593-598

De Vries BB, Pfundt R, Leisink M, Koolen DA, Vissers LE, Janssen IM, Reijmersdal S, Nillesen WM, Huys EH, Leeuw N, Smeets D, Sistermans EA, Feuth T, Ravenswaaij-Arts CM, van Kessel AG, Schoenmakers EF, Brunner HG, Veltman JA (2005) Diagnostic genome profiling in mental retardation. *Am J Hum Genet* 77:606-616

De Vries BB, van den Ouweland AM, Mohkamsing S, Duivenvoorden HJ, Mol E, Gelsema K, van Rijn M, Halley DJ, Sandkuijl LA, Oostra BA, Tibben A, Niermeijer MF (1997) Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group. *Am J Hum Genet* **61**:660-667

Delach JA, Rosengren SS, Kaplan L, Greenstein RM, Cassidy SB, Benn PA (1994) Comparison of high resolution chromosome banding and fluorescence in situ hybridization (FISH) for the laboratory evaluation of Prader-Willi syndrome and Angelman syndrome. *Am J Med Genet* **52**:85-91

Den Dunnen JT, Bakker E, Klein-Breteler EG, Pearson PL, Van Ommen GJB (1987) Direct detection of more than 50% Duchenne muscular dystrophy mutations by field-inversion gels. *Nature* **329**:640-642

Edelmann L, Pandita RK, Morrow BE (1999) Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am J Hum Genet* **64**:1076-1086

Edelmann L, Pandita RK, Spiteri E, Funke B, Goldberg R, Palanisamy N, Chaganti RS, Magenis E, Shprintzen RJ, Morrow BE (1999) A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum Mol Genet* **8**:1157-1167

Eichler EE (2001a) Recent duplication, domain accretion and the dynamic mutation of the human genome. *Trends Genet* **17**:661-669

Eichler EE (2001b) Segmental duplications: what's missing, misassigned, and misassembled--and should we care? *Genome Res* **11**:653-656

Eichler EE (2006) Widening the spectrum of human genetic variation. *Nat Genet* **38**:9-11

Eichler EE, Lu F, Shen Y, Antonacci R, Jurecic V, Doggett NA, Moyzis RK, Baldini A, Gibbs RA, Nelson DL (1996) Duplication of a gene-rich cluster between 16p11.1 and Xq28: a novel pericentromeric-directed mechanism for paralogous genome evolution. *Hum Mol Genet* **5**:899-912

Einfeld SL (1984) Clinical assessment of 4500 developmentally delayed individuals. *J Ment Defic Res* **28** (Pt 2):129-142

Elwood JH, Darragh PM (1981) Severe mental handicap in Northern Ireland. *J Ment Defic Res* **25**:147-155

Engels H, Ehrbrecht A, Zahn S, Bosse K, Vrolijk H, White S, Kalscheuer V, Hoovers JM, Schwanitz G, Propping P, Tanke HJ, Wiegant J, Raap AK (2003) Comprehensive analysis of human subtelomeres with combined binary ratio labelling fluorescence in situ hybridisation. *Eur J Hum Genet* **11**:643-651

Ensenauer RE, Adeyinka A, Flynn HC, Michels VV, Lindor NM, Dawson DB, Thorland EC, Lorentz CP, Goldstein JL, McDonald MT, Smith WE, Simon-Fayard E, Alexander AA, Kulharya AS, Ketterling RP, Clark RD, Jalal SM (2003) Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet* **73**:1027-1040

Fan YS, Siu VM, Jung JH, Xu J (2000) Sensitivity of multiple color spectral karyotyping in detecting small interchromosomal rearrangements. *Genet Test* **4**:9-14

Fan YS, Zhang Y, Speevak M, Farrell S, Jung JH, Siu VM (2001) Detection of submicroscopic aberrations in patients with unexplained mental retardation by fluorescence in situ hybridization using multiple subtelomeric probes. *Genet Med* **3**:416-421

Flint J, Knight S (2003) The use of telomere probes to investigate submicroscopic rearrangements associated with mental retardation. *Curr Opin Genet Dev* **13**:310-316

Flint J, Wilkie AO (1996) The genetics of mental retardation. *Br Med Bull* **52**:453-464

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

Florijn RJ, Blonden LAJ, Vrolijk H, Wiegant J, Vaandrager JW, Baas F, Den Dunnen JT, Tanke HJ, Van Ommen GJB, Raap AK (1995) High-resolution FISH for genomic DNA mapping and colour bar-coding of large genes. *Hum Mol Genet* **4**:831-836

Ford M, Fried M (1986) Large inverted duplications are associated with gene amplification. *Cell* **45**:425-430

Francke U (1999) Williams-Beuren syndrome: genes and mechanisms. *Hum Mol Genet* **8**:1947-1954

Francke U, Ochs HD, De Martinville B, Giacalone J, Lindgren V, Disteche C, Pagon RA, Hofker MH, Van Ommen GJB, Pearson PL, Wedgwood R (1985) Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy,chronic granulomatous disease, retinitis pigmentosa and McLeod syndrome. *Am J Hum Genet* **37**:250-267

Fredman D, White SJ, Potter S, Eichler EE, Den Dunnen JT, Brookes AJ (2004) Complex SNP-related sequence variation in segmental genome duplications. *Nat Genet* **36**:861-866

Friedman JM, Baross A, Delaney AD, Ally A, Arbour L, Asano J, Bailey DK *et al.* (2006) Oligonucleotide microarray analysis of genomic imbalance in children with mental retardation. *Am J Hum Genet* **79**:500-513

Gibbons B, Datta P, Wu Y, Chan A, Al Armour J (2006) Microarray MAPH: accurate array-based detection of relative copy number in genomic DNA. *BMC Genomics* **7**:163

Giglio S, Calvari V, Gregato G, Gimelli G, Camanini S, Giorda R, Ragusa A, Guerneri S, Selicorni A, Stumm M, Tonnies H, Ventura M, Zollino M, Neri G, Barber J, Wieczorek D, Rocchi M, Zuffardi O (2002) Heterozygous submicroscopic inversions involving olfactory receptor-gene clusters mediate the recurrent t(4;8)(p16;p23) translocation. *Am J Hum Genet* **71**:276-285

Giles RH, Petrij F, Dauwerse HG, Den Hollander AI, Lushnikova T, Van Ommen GJB, Goodman RH, Deaven LL, Doggett NA, Peters DJ, Breuning MH (1997) Construction of a 1.2-Mb contig surrounding, and molecular analysis of, the human CREB-binding protein (CBP/CREBBP) gene on chromosome 16p13.3. *Genomics* **42**:96-114

Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, Murthy KK, Rovin BH, Bradley W, Clark RA, Anderson SA, O'connell RJ, Agan BK, Ahuja SS, Bologna R, Sen L, Dolan MJ, Ahuja SK (2005) The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **307**:1434-1440

Gribble SM, Prigmore E, Burford DC, Porter KM, Ng BL, Douglas EJ, Fiegler H, Carr P, Kalaitzopoulos D, Clegg S, Sandstrom R, Temple IK, Youings SA, Thomas NS, Dennis NR, Jacobs PA, Crolla JA, Carter NP (2005) The complex Nature of constitutional *de novo* apparently balanced translocations in patients presenting with abnormal phenotypes. *J Med Genet* **42**:8-16

Groot PC, Mager WH, Henriquez NV, Pronk JC, Arwert F, Planta RJ, Eriksson AW, Frants RR (1990) Evolution of the human alpha-amylase multigene family through unequal, homologous, and inter- and intrachromosomal crossovers. *Genomics* **8**:97-105

Gustavson KH, Holmgren G, Blomquist HK (1987) Chromosomal aberrations in mildly mentally retarded children in a northern Swedish county. *Ups J Med Sci Suppl* **44**:165-168

Hall H, Hunt P, Hassold T (2006) Meiosis and sex chromosome aneuploidy: how meiotic errors cause aneuploidy; how aneuploidy causes meiotic errors. *Curr Opin Genet Dev* **16**:323-329

Harada N, Hatchwell E, Okamoto N, Tsukahara M, Kurosawa K, Kawame H, Kondoh T, Ohashi H, Tsukino R, Kondoh Y, Shimokawa O, Ida T, Nagai T, Fukushima Y, Yoshiura K, Niikawa N, Matsumoto N (2004) Subtelomere specific microarray based comparative genomic hybridisation: a rapid detection system for cryptic rearrangements in idiopathic mental retardation. *J Med Genet* **41**:130-136

Harteveld CL, Voskamp A, Phylipsen M, Akkermans N, Den Dunnen JT, White SJ, Giordano PC (2005) Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha- and beta-thalassaemia characterised by high resolution multiplex ligation-dependent probe amplification. *J Med Genet* **42**:922-931

Heath KE, Day IN, Humphries SE (2000) Universal primer quantitative fluorescent multiplex (UPQFM) PCR: a method to detect major and minor rearrangements of the low density lipoprotein receptor gene. *J Med Genet* **37**:272-280

Helias-Rodzewicz Z, Bocian E, Stankiewicz P, Obersztyn E, Kostyk E, Jakubow-Durska K, Kutkowska-Kazmierczak A, Mazurczak T (2002) Subtelomeric rearrangements detected by FISH in three of 33 families with idiopathic mental retardation and minor physical anomalies. *J Med Genet* **39**: e53

Herr A, Grutzmann R, Matthaei A, Artelt J, Schrock E, Rump A, Pilarsky C (2005) High-resolution analysis of chromosomal imbalances using the Affymetrix 10K SNP genotyping chip. *Genomics* **85**:392-400

Herrmann BG, Barlow DP, Lehrach H (1987) A large inverted duplication allows homologous recombination between chromosomes heterozygous for the proximal t complex inversion. *Cell* **48**:813-825

Higgs DR, Hill AV, Bowden DK, Weatherall DJ, Clegg JB (1984) Independent recombination events between the duplicated human alpha globin genes; implications for their concerted evolution. *Nucleic Acids Res* **12**:6965-6977

Hollox EJ, Atia T, Cross G, Parkin T, Armour JA (2002) High throughput screening of human subtelomeric DNA for copy number changes using multiplex amplifiable probe hybridisation (MAPH). *J Med Genet* **39**:790-795

Horvath JE, Schwartz S, Eichler EE (2000) The mosaic structure of human pericentromeric DNA: a strategy for characterizing complex regions of the human genome. *Genome Res* **10**:839-852

Hulley BJ, Hummel M, Wenger SL (2003) Screening for cryptic chromosomal abnormalities in patients with mental retardation and dysmorphic facial features using telomere FISH probes. *Am J Med Genet* A **117**:302-303

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

Inoue K, Dewar K, Katsanis N, Reiter LT, Lander ES, Devon KL, Wyman DW, Lupski JR, Birren B (2001) The 1.4-Mb CMT1A duplication/HNPP deletion genomic region reveals unique genome architectural features and provides insights into the recent evolution of new genes. *Genome Res* **11**:1018-1033

Jacobs PA, Baikie AG, Court Brown WM, Strong JA (1959) The somatic chromosomes in mongolism. *Lancet* **1**:710

Jalal SM, Harwood AR, Sekhon GS, Pham LC, Ketterling RP, Babovic-Vuksanovic D, Meyer RG, Ensenauer R, Anderson MH, Jr., Michels VV (2003) Utility of subtelomeric fluorescent DNA probes for detection of chromosome anomalies in 425 patients. *Genet Med* **5**:28-34

Ji Y, Eichler EE, Schwartz S, Nicholls RD (2000) Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res* **10**:597-610

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

Knight SJ, Horsley SW, Regan R, Lawrie NM, Maher EJ, Cardy DL, Flint J, Kearney L (1997) Development and clinical application of an innovative fluorescence in situ hybridization technique which detects submicroscopic rearrangements involving telomeres. *Eur J Hum Genet* **5**:1-8

Knight SJ, Regan R, Nicod A, Horsley SW, Kearney L, Homfray T, Winter RM, Bolton P, Flint J (1999) Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* **354**:1676-1681

Komura D, Shen F, Ishikawa S, Fitch KR, Chen W, Zhang J, Liu G, Ihara S, Nakamura H, Hurles ME, Lee C, Scherer SW, Jones KW, Shapero MH, Huang J, Aburatani H (2006) Genome-wide detection of human copy number variations using high-density DNA oligonucleotide arrays. *Genome Res* **16**:1575-1584

Koolen DA, Nillesen WM, Versteeg MH, Merkx GF, Knoers NV, Kets M, Vermeer S, van Ravenswaaij CM, de Kovel CG, Brunner HG, Smeets D, De Vries BB, Sistermans EA (2004) Screening for subtelomeric rearrangements in 210 patients with unexplained mental retardation using multiplex ligation dependent probe amplification (MLPA). *J Med Genet* **41**:892-899

Koolen DA, Vissers LE, Pfundt R, de Leeuw N, Knight SJ, Regan R, Kooy RF, Reyniers E, Romano C, Fichera M, Schinzel A, Baumer A, Anderlid BM, Schoumans J, Knoers NV, van Kessel AG, Sistermans EA, Veltman JA, Brunner HG, De Vries BB (2006) A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* **38**:999-1001

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG (2004) Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nat Genet* **36**:631-635

Kriek M, White SJ, Szuhai K, Knijnenburg J, van Ommen GJ, Den Dunnen JT, Breuning MH (2006) Copy number variation in regions flanked (or unflanked) by duplicons among patients with developmental delay and/or congenital malformations; detection of reciprocal and partial Williams-Beuren duplications. *Eur J Hum Genet* **14**:180-189

Kurahashi H, Shaikh T, Takata M, Toda T, Emanuel BS (2003) The constitutional t(17;22): another translocation mediated by palindromic AT-rich repeats. *Am J Hum Genet* **72**:733-738

Kurahashi H, Shaikh TH, Hu P, Roe BA, Emanuel BS, Budarf ML (2000) Regions of genomic instability on 22q11 and 11q23 as the etiology for the recurrent constitutional t(11;22). *Hum Mol Genet* **9**:1665-1670

Kuwano A, Mutirangura A, Dittrich B, Buiting K, Horsthemke B, Saitoh S, Niikawa N, Ledbetter SA, Greenberg F, Chinault AC, . (1992) Molecular dissection of the Prader-Willi/Angelman syndrome region (15q11-13) by YAC cloning and FISH analysis. *Hum Mol Genet* **1**:417-425

Lam AC, Lam ST, Lai KK, Tong TM, Chau TC (2006) High rate of detection of subtelomeric aberration by using combined MLPA and subtelomeric FISH approach in patients with moderate to severe mental retardation. *Clin Biochem* **39**:196-202

Lamont MA, Dennis NR (1988) Aetiology of mild mental retardation. *Arch Dis Child* **63**:1032-1038

Landegent JE, Jansen in dW, van Ommen GJ, Baas F, de Vijlder JJ, Van Duijn P, Van der PM (1985) Chromosomal localization of a unique gene by non-autoradiographic in situ hybridization. *Nature* **317**:175-177

Laurendeau I, Bahuau M, Vodovar N, Larramendy C, Olivi M, Bieche I, Vidaud M, Vidaud D (1999) TaqMan PCR-based gene dosage assay for predictive testing in individuals from a cancer family with INK4 locus haploinsufficiency. *Clin Chem* **45**:982-986

Lauritsen JG, Jonasson J, Therkelsen AJ, Lass F, Lindsten J, Petersen GB (1972) Studies on spontaneous abortions. Fluorescence analysis of abnormal karyotypes. *Hereditas* **71**:160-163

Leana-Cox J, Pangkanon S, Eanet KR, Curtin MS, Wulfsberg EA (1996) Familial DiGeorge/velocardiofacial syndrome with deletions of chromosome area 22q11.2: report of five families with a review of the literature. *Am J Med Genet* **65**:309-316

Lejeune J, Turpin R, Gautier M (1959) [Mongolism; a chromosomal disease (trisomy).]. *Bull Acad Natl Med* **143**:256-265

Li R, Zhao ZY (2004) Two subtelomeric chromosomal deletions in forty-six children with idiopathic mental retardation. *Chin Med J* (Engl.) **117**:1414-1417

Liu G, Zhao S, Bailey JA, Sahinalp SC, Alkan C, Tuzun E, Green ED, Eichler EE (2003) Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome. *Genome Res* **13**:358-368

Locke DP, Archidiacono N, Misceo D, Cardone MF, Deschamps S, Roe B, Rocchi M, Eichler EE (2003) Refinement of a chimpanzee pericentric inversion breakpoint to a segmental duplication cluster. *Genome Biol* **4**:R50

Lucci-Cordisco E, Zollino M, Baglioni S, Mancuso I, Lecce R, Gurrieri F, Crucitti A, Papi L, Neri G, Genuardi M (2005) A novel microdeletion syndrome with loss of the MSH2 locus and hereditary non-polyposis colorectal cancer. *Clin Genet* **67**:178-182

Lupski JR, Montes De Oca-Luna R, Slaugenhaupt S, Pentao L, Guzzetta V, Trask B, Saucedo-Cardenas O, Barker DF, Killian JM, Garcia CA, Chakravarti A, Patel PI (1991) DNA duplication associated with Charcot-Marie-Tooth disease type 1a. *Cell* **66**:219-232

Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* **14**:417-422

Lupski JR (2007) Structural variation in the human genome. *N Engl J Med* **356**:1169-1171

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

McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ, Altshuler DM (2006) Common deletion polymorphisms in the human genome. *Nat Genet* **38**:86-92

McDonald AD (1973) Severely retarded children in Quebec: prevalence, causes, and care. *Am J Ment Defic* **78**:205-215

Mefford HC, Trask BJ (2002) The complex structure and dynamic evolution of human subtelomeres. *Nat Rev Genet* **3**:91-102

Menten B, Maas N, Thienpont B, Buysse K, Vandesompele J, Melotte C, de Ravel T, Van Vooren S, Ba-

likova I, Backx L, Janssens S, De Paepe A, De Moor B, Moreau Y, Marynen P, Fryns JP, Mortier G, Devriendt K, Speleman F, Vermeesch JR (2006) 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* **43**:625-633

Miller SA, Dykes DD, PoleskY HF (1988) A simple salting out procedure for extracting DNA from human nucleated Cells. *Nucleic Acids Res* **16**:1215

Ming JE, Geiger E, James AC, Ciprero KL, Nimmakayalu M, Zhang Y, Huang A, Vaddi M, Rappaport E, Zackai EH, Shaikh TH (2006) Rapid detection of submicroscopic chromosomal rearrangements in children with multiple congenital anomalies using high density oligonucleotide arrays. *Hum Mutat* **27**:467-473

Miyake N, Shimokawa O, Harada N, Sosonkina N, Okubo A, Kawara H, Okamoto N, Kurosawa K, Kawame H, Iwakoshi M, Kosho T, Fukushima Y, Makita Y, Yokoyama Y, Yamagata T, Kato M, Hiraki Y, Nomura M, Yoshiura K, Kishino T, Ohta T, Mizuguchi T, Niikawa N, Matsumoto N (2006) BAC array CGH reveals genomic aberrations in idiopathic mental retardation. *Am J Med Genet* A **140**:205-211

Morris CA, Thomas IT, Greenberg F (1993) Williams syndrome: autosomal dominant inheritance. *Am J Med Genet* **47**:478-481

Murray A, Youings S, Dennis N, Latsky L, Linehan P, McKechnie N, Macpherson J, Pound M, Jacobs P (1996) Population screening at the FRAXA and FRAXE loci: molecular analyses of boys with learning difficulties and their mothers. *Hum Mol Genet* **5**:727-735

Nederlof PM, Robinson D, Abuknesha R, Wiegant J, Hopman AH, Tanke HJ, Raap AK (1989) Three-color fluorescence in situ hybridization for the simultaneous detection of multiple nucleic acid sequences. *Cytometry* **10**:20-27

Nederlof PM, Van Der Flier S, Wiegant J, Raap AK, Tanke HJ, Ploem JS, Van der ploeg M (1990) Multiple fluorescence in situ hybridization. *Cytometry* **11**:126-131

Nobile C, Toffolatti L, Rizzi F, Simionati B, Nigro V, Cardazzo B, Patarnello T, Valle G, Danieli GA (2002) Analysis of 22 deletion breakpoints in dystrophin intron 49. *Hum Genet* **110**:418-421

Novelli A, Ceccarini C, Bernardini L, Zuccarello D, Caputo V, Digilio MC, Mingarelli R, Dallapiccola B (2004) High frequency of subtelomeric rearrangements in a cohort of 92 patients with severe mental retardation and dysmorphism. *Clin Genet* **66**:30-38

Palomares M, Delicado A, Lapunzina P, Arjona D, Aminoso C, Arcas J, Martinez BA, Fernandez L, Lopez P, I (2006) MLPA vs multiprobe FISH: comparison of two *Methods* for the screening of subtelomeric rearrangements in 50 patients with idiopathic mental retardation. *Clin Genet* **69**:228-233

Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, Cheung SW, Shen RM, Barker DL, Gunderson KL (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res* **16**:1136-1148

Pentao L, Wise CA, Chinault AC, Patel PI, Lupski JR (1992) Charcot-Marie-Tooth type 1A duplication appears to arise from recombination at repeat sequences flanking the 1.5 Mb monomer unit. *Nat Genet* **2**:292-300

Peoples R, Franke Y, Wang YK, Perez-Jurado L, Paperna T, Cisco M, Francke U (2000) A physical map, including a BAC/PAC clone contig, of the Williams-Beuren syndrome--deletion region at 7q11.23. *Am J Hum Genet* **66**:47-68

Petrij F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RC, Masuno M, Tommerup N, van Ommen GJ, Goodman RH, Peters DJ. (1995) Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376**:348-351

Pickard BS, Hollox EJ, Malloy MP, Porteous DJ, Blackwood DH, Armour JA, Muir WJ (2004) A 4q35.2 subtelomeric deletion identified in a screen of patients with co-morbid psychiatric illness and mental retardation. *BMC Med Genet* **5**:21

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* **20**:207-211

Popp S, Schulze B, Granzow M, Keller M, Holtgreve-Grez H, Schoell B, Brough M, Hager HD, Tariverdian G, Brown J, Kearney L, Jauch A (2002) Study of 30 patients with unexplained developmental delay and dysmorphic features or congenital abnormalities using conventional cytogenetics and multiplex FISH telomere (M-TEL) integrity assay. *Hum Genet* **111**:31-39

Potocki L, Bi W, Treadwell-Deering D, Carvalho CM, Eifert A, Friedman EM, Glaze D, Krull K, Lee JA, Lewis RA, Mendoza-Londono R, Robbins-Furman P, Shaw C, Shi X, Weissenberger G, Withers M, Yatsenko SA, Zackai EH, Stankiewicz P, Lupski JR (2007) Characterization of Potocki-Lupski Syndrome (dup(17)(p11.2p11.2)) and Delineation of a Dosage-Sensitive Critical Interval That Can Convey an Autism Phenotype. *Am J Hum Genet* **80**:633-649

Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboa K, Kashork CD, Shaffer LG, Lupski JR (2000) Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* **24**:84-87

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

Raap AK, Florijn RJ, Blonden LAJ, Wiegant J, Vaandrager JW, Vrolijk H, Den Dunnen JT, Tanke HJ, Van Ommen GJB (1996) FiberFISH as a DNA mapping tool. *Methods* **9**:67-73

Raap AK, Tanke HJ (2006) COmbined Binary RAtio fluorescence in situ hybridiziation (COBRA-FISH): development and applications. Cytogenet. *Genome Res* **114**:222-226

Rauch A, Hoyer J, Guth S, Zweier C, Kraus C, Becker C, Zenker M, Huffmeier U, Thiel C, Ruschendorf F, Nurnberg P, Reis A, Trautmann U (2006) Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet* A **140**:2063-2074

Rauch A, Ruschendorf F, Huang J, Trautmann U, Becker C, Thiel C, Jones KW, Reis A, Nurnberg P (2004) Molecular karyotyping using an SNP array for genomewide genotyping. *J Med Genet* **41**:916-922

Rauen KA, Albertson DG, Pinkel D, Cotter PD (2002) Additional patient with del(12)(q21.2q22): further evidence for a candidate region for cardio-facio-cutaneous syndrome? *Am J Med Genet* **110**:51-56

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H *et al.* (2006) Global variation in copy number in the human genome. *Nature* **444**:444-454

Reiter LT, Hastings PJ, Nelis E, De Jonghe P, Van Broeckhoven C, Lupski JR (1998) Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. *Am J Hum Genet* **62**:1023-1033

Reiter LT, Murakami T, Koeuth T, Pentao L, Muzny DM, Gibbs RA, Lupski JR (1996) A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. *Nat Genet* **12**:288-297

Ried T, Mahler V, Vogt P, Blonden LAJ, Van Ommen GJB, Cremer T, Cremer M (1990) Direct carrier

detection by in situ suppression hybridization with cosmid clones of the Duchenne/Becker muscular dystrophy locus. *Hum Genet* **85**:581-586

Riegel M, Baumer A, Jamar M, Delbecque K, Herens C, Verloes A, Schinzel A (2001) Submicroscopic terminal deletions and duplications in retarded patients with unclassified malformation syndromes. *Hum Genet* **109**:286-294

Rio M, Molinari F, Heuertz S, Ozilou C, Gosset P, Raoul O, Cormier-Daire V, Amiel J, Lyonnet S, Le Merrer M, Turleau C, de Blois MC, Prieur M, Romana S, Vekemans M, Munnich A, Colleaux L (2002) Automated fluorescent genotyping detects 10% of cryptic subtelomeric rearrangements in idiopathic syndromic mental retardation. *J Med Genet* **39**:266-270

Rodriguez-Revenga L, Badenas C, Sanchez A, Mallolas J, Carrio A, Pedrinaci S, Barrionuevo JL, Mila M (2004) Cryptic chromosomal rearrangement screening in 30 patients with mental retardation and dysmorphic features. *Clin Genet* **65**:17-23

Rooms L, Reyniers E, van Luijk R, Scheers S, Wauters J, Ceulemans B, Van Den EJ, Van Bever Y, Kooy RF (2004a) Subtelomeric deletions detected in patients with idiopathic mental retardation using multiplex ligation-dependent probe amplification (MLPA). *Hum Mutat* **23**:17-21

Rooms L, Reyniers E, van Luijk R, Scheers S, Wauters J, Kooy RF (2004b) Screening for subtelomeric rearrangements using genetic markers in 70 patients with unexplained mental retardation. *Ann Genet* **47**:53-59

Rooms L, Reyniers E, Wuyts W, Storm K, van Luijk R, Scheers S, Wauters J, Van Den EJ, Biervliet M, Eyskens F, van Goethem G, Laridon A, Ceulemans B, Courtens W, Kooy RF (2006) Multiplex ligation-dependent probe amplification to detect subtelomeric rearrangements in routine diagnostics. *Clin Genet* **69**:58-64

Ropers HH (2007) New perspectives for the elucidation of genetic disorders. *Am J Hum Genet* **81**:199-207

Rosenberg C, Florijn RJ, Blonden LAJ, Van Ommen GJB, Den Dunnen JT (1995) High resolution DNA fiber FISH on yeast artificial chromosomes: direct visualization of replication forks. *Nat Genet* **10**:477-479

Rosenberg C, Knijnenburg J, Bakker E, Vianna-Morgante AM, Sloos W, Otto PA, Kriek M, Hansson K, Krepischi-Santos AC, Fiegler H, Carter NP, Bijlsma EK, Van Haeringen A, Szuhai K, Tanke HJ (2006) Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. *J Med Genet* **43**:180-186

Rosenberg MJ, Killoran C, Dziadzio L, Chang S, Stone DL, Meck J, Aughton D, Bird LM, Bodurtha J, Cassidy SB, Graham JM, Jr., Grix A, Guttmacher AE, Hudgins L, Kozma C, Michaelis RC, Pauli R, Peters KF, Rosenbaum KN, Tifft CJ, Wargowski D, Williams MS, Biesecker LG (2001) Scanning for telomeric deletions and duplications and uniparental disomy using genetic markers in 120 children with malformations. *Hum Genet* **109**:311-318

Rossi E, Piccini F, Zollino M, Neri G, Caselli D, Tenconi R, Castellan C, Carrozzo R, Danesino C, Zuffardi O, Ragusa A, Castiglia L, Galesi O, Greco D, Romano C, Pierluigi M, Perfumo C, Di Rocco M, Faravelli F, Dagna BF, Bonaglia M, Bedeschi M, Borgatti R (2001) Cryptic telomeric rearrangements in subjects with mental retardation associated with dysmorphism and congenital malformations. *J Med Genet* **38**:417-420

Samonte RV, Eichler EE (2002b) Segmental duplications and the evolution of the primate genome. *Nat Rev Genet* **3**:65-72

Sanders SR, Dawson AJ, Vust A, Hryshko M, Tomiuk M, Riordan D, Prasad C (2003) Interstitial deletion of chromosome 2p16.2p21. *Clin Dysmorphol* **12**:183-185

Schoumans J, Ruivenkamp C, Holmberg E, Kyllerman M, Anderlid BM, Nordenskjold M (2005) Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). *J Med Genet* **42**:699-705

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

Schrock E, Veldman T, Padilla-Nash H, Ning Y, Spurbeck J, Jalal S, Shaffer LG, Papenhausen P, Kozma C, Phelan MC, Kjeldsen E, Schonberg SA, O'Brien P, Biesecker L, du MS, Ried T (1997) Spectral karyotyping refines cytogenetic diagnostics of constitutional chromosomal abnormalities. *Hum Genet* **101**:255-262

Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B *et al*. (2007) Strong Association of *De novo* Copy Number Mutations with Autism. *Science*

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M (2004) Large-scale copy number polymorphism in the human genome. *Science* **305**:525-528

Shaffer LG, Lupski JR (2000) Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Annu Rev Genet* **34**:297-329

Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, Emanuel BS (2000) Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* **9**:489-501

Sharp AJ, Hansen S, Selzer RR, Cheng Z, Regan R, Hurst JA, Stewart H, Price SM, Blair E, Hennekam RC, Fitzpatrick CA, Segraves R, Richmond TA, Guiver C, Albertson DG, Pinkel D, Eis PS, Schwartz S, Knight SJ, Eichler EE (2006) Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* **38**:1038-1042

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, Oseroff VV, Albertson DG, Pinkel D, Eichler EE (2005) Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* **77**:78-88

Shaw CJ, Bi W, Lupski JR (2002) Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. *Am J Hum Genet* **71**:1072-1081

Shaw-Smith C, Pittman AM, Willatt L, Martin H, Rickman L, Gribble S, Curley R, Cumming S, Dunn C, Kalaitzopoulos D, Porter K, Prigmore E, Krepischi-Santos AC, Varela MC, Koiffmann CP, Lees AJ, Rosenberg C, Firth HV, de Silva R, Carter NP (2006) Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet*

She X, Jiang Z, Clark RA, Liu G, Cheng Z, Tuzun E, Church DM, Sutton G, Halpern AL, Eichler EE (2004) Shotgun sequence assembly and recent segmental duplications within the human genome. *Nature* **431**:927-930

She X, Liu G, Ventura M, Zhao S, Misceo D, Roberto R, Cardone MF, Rocchi M, Green ED, Archidiacano N, Eichler EE (2006) A preliminary comparative analysis of primate segmental duplications shows elevated substitution rates and a great-ape expansion of intrachromosomal duplications. *Genome Res* **16**:576-583

Shuber AP, Grondin VJ, Klinger KW (1995) A simplified procedure for developing multiplex PCRs. *Genome Res* **5**:488-493

Sismani C, Armour JA, Flint J, Girgalli C, Regan R, Patsalis PC (2001) Screening for subtelomeric chromosome abnormalities in children with idiopathic mental retardation using multiprobe telomeric FISH and the new MAPH telomeric assay. *Eur J Hum Genet* **9**:527-532

Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH (2003) Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat Genet* **33**:466-468

Slater HR, Bailey DK, Ren H, Cao M, Bell K, Nasioulas S, Henke R, Choo KH, Kennedy GC (2005) High-Resolution Identification of Chromosomal Abnormalities Using Oligonucleotide Arrays Containing 116,204 SNPs. *Am J Hum Genet* 77:709-726

Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, Law S, Myambo K, Palmer J, Ylstra B, Yue JP, Gray JW, Jain AN, Pinkel D, Albertson DG (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* **29**:263-264

Snijders AM, Nowee ME, Fridlyand J, Piek JM, Dorsman JC, Jain AN, Pinkel D, van Diest PJ, Verheijen RH, Albertson DG (2003) Genome-wide-array-based comparative genomic hybridization reveals genetic homogeneity and frequent copy number increases encompassing CCNE1 in fallopian tube carcinoma. *Oncogene* **22**:4281-4286

Snijders AM, Pinkel D, Albertson DG (2003) Current status and future prospects of array-based comparative genomic hybridisation. *Brief Funct Genomic Proteomic* **2**:37-45

Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* **20**:399-407

Somerville MJ, Mervis CB, Young EJ, Seo EJ, del Campo M, Bamforth S, Peregrine E, Loo W, Lilley M, Perez-Jurado LA, Morris CA, Scherer SW, Osborne LR (2005) Severe expressive-language delay related to duplication of the Williams-Beuren locus. *N Engl J Med* **353**:1694-1701

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**:503-517

Speicher MR, Gwyn BS, Ward DC (1996) Karyotyping human chromosomes by combinatorial multifluor FISH. *Nat Genet* **12**:368-375

Stankiewicz P, Park SS, Inoue K, Lupski JR (2001) The evolutionary chromosome translocation 4;19 in Gorilla gorilla is associated with microduplication of the chromosome fragment syntenic to sequences surrounding the human proximal CMT1A-REP. *Genome Res* **11**:1205-1210

Stoppa-Lyonnet D, Carter PE, Meo T, Tosi M (1990) Clusters of intragenic Alu repeats predispose the human C1 inhibitor locus to deleterious rearrangements. *Proc Natl Acad Sci U.S.A.* **87**:1551-1555

Tanke HJ, Wiegant J, van Gijlswijk RP, Bezrookove V, Pattenier H, Heetebrij RJ, Talman EG, Raap AK, Vrolijk J (1999) New strategy for multi-colour fluorescence in situ hybridisation: COBRA: COmbined Binary RAtio labelling. *Eur J Hum Genet* 7:2-11

Taylor JS, Raes J (2004) Duplication and divergence: the evolution of new genes and old ideas. *Annu Rev Genet* **38**:615-643

Telenius H, Carter NP, Bebb CE, Nordenskjold M, Ponder BA, Tunnacliffe A (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* **13**:718-725

Ting JC, Ye Y, Thomas GH, Ruczinski I, Pevsner J (2006) Analysis and visualization of chromosomal abnormalities in SNP data with SNPscan. *BMC Bioinformatics* 7:25

Trask BJ, Friedman C, Martin-Gallardo A, Rowen L, Akinbami C, Blankenship J, Collins C, Giorgi D, Iadonato S, Johnson F, Kuo WL, Massa H, Morrish T, Naylor S, Nguyen OT, Rouquier S, Smith

T, Wong DJ, Youngblom J, van den EG (1998) Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. *Hum Mol Genet* **7**:13-26

Turner G, Webb T, Wake S, Robinson H (1996) Prevalence of fragile X syndrome. *Am J Med Genet* **64**:196-197

Tyson C, Harvard C, Locker R, Friedman JM, Langlois S, Lewis ME, Van Allen M, Somerville M, Arbour L, Clarke L, McGilivray B, Yong SL, Siegel-Bartel J, Rajcan-Separovic E (2005) Submicroscopic deletions and duplications in individuals with intellectual disability detected by array-CGH. *Am J Med Genet* A **139**:173-185

Ullmann R, Turner G, Kirchhoff M, Chen W, Tonge B, Rosenberg C, Field M, Vianna-Morgante AM, Christie L, Krepischi-Santos AC, Banna L, Brereton AV, Hill A, Bisgaard AM, Muller I, Hultschig C, Erdogan F, Wieczorek G, Ropers HH (2007) Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* **28**:674-682

Urban Z, Helms C, Fekete G, Csiszar K, Bonnet D, Munnich A, Donis-Keller H, Boyd CD (1996) 7q11.23 deletions in Williams syndrome arise as a consequence of unequal meiotic crossover. *Am J Hum Genet* **59**:958-962

Valentijn LJ, Bolhuis PA, Zorn I, Hoogendijk JE, Van Den Bosch N, Hessels GW, Stanton VP, Husman DE, Fischbeck KH, Ross DA, Nicholson GA, Meershoek EJ, Dauwerse HG, Van Ommen GJB, Baas F (1992) The peripheral myelin gene PMP22/GAS3 is duplicated in Charcot-Marie-Tooth disease type 1A. *Nat Genet* **1**:166-170

van Geel M, Dickson MC, Beck AF, Bolland DJ, Frants RR, Van Der Maarel SM, De Jong PJ, Hewitt JE (2002) Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics* **79**:210-217

van Karnebeek CD, Koevoets C, Sluijter S, Bijlsma EK, Smeets DF, Redeker EJ, Hennekam RC, Hoovers JM (2002) Prospective screening for subtelomeric rearrangements in children with mental retardation of unknown aetiology: the Amsterdam experience. *J Med Genet* **39**:546-553

van Ommen GJ (2005) Frequency of new copy number variation in humans. *Nat Genet* **37**:333-334

van Ommen GJ, Verkerk JM, Hofker MH, Monaco AP, Kunkel LM, Ray P, Worton R, Wieringa B, Bakker E, Pearson PL (1986) A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X-chromosome. *Cell* **47**:499-504

van Overveld PG, Lemmers RJ, Deidda G, Sandkuijl L, Padberg GW, Frants RR, Van Der Maarel SM (2000) Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity. *Hum Mol Genet* **9**:2879-2884

Veltman JA (2006) Genomic microarrays in clinical diagnosis. Curr.Opin.Pediatr. 18:598-603

Veltman JA, Schoenmakers EF, Eussen BH, Janssen I, Merkx G, van Cleef B, van Ravenswaaij CM, Brunner HG, Smeets D, van Kessel AG (2002) High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization. *Am J Hum Genet* **70**:1269-1276

Vissers LE, De Vries BB, Osoegawa K, Janssen IM, Feuth T, Choy CO, Straatman H, Van D, V, Huys EH, Van Rijk A, Smeets D, Ravenswaaij-Arts CM, Knoers NV, Van DB, I, De Jong PJ, Brunner HG, van Kessel AG, Schoenmakers EF, Veltman JA (2003) Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* **73**:1261-1270

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, Schoenmakers EF, Brunner HG, Veltman JA, van Kessel AG (2004) Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* **36**:955-957

Vissers LE, Veltman JA, van Kessel AG, Brunner HG (2005) Identification of disease genes by whole genome CGH arrays. *Hum Mol Genet* **14** Suppl 2:R215-R223

Waldman AS, Liskay RM (1988) Dependence of intrachromosomal recombination in mammalian Cells on uninterrupted homology. *Mol Cell Biol* **8**:5350-5357

Walter S, Sandig K, Hinkel GK, Mitulla B, Ounap K, Sims G, Sitska M, Utermann B, Viertel P, Kalscheuer V, Bartsch O (2004) Subtelomere FISH in 50 children with mental retardation and minor anomalies, identified by a checklist, detects 10 rearrangements including a *de novo* balanced translocation of chromosomes 17p13.3 and 20q13.33. *Am J Med Genet* A **128**:364-373

Weiss MM, Snijders AM, Kuipers EJ, Ylstra B, Pinkel D, Meuwissen SG, van Diest PJ, Albertson DG, Meijer GA (2003) Determination of amplicon boundaries at 20q13.2 in tissue samples of human gastric adenocarcinomas by high-resolution microarray comparative genomic hybridization. *J Pathol* **200**:320-326

White S, Kalf M, Liu Q, Villerius M, Engelsma D, Kriek M, Vollebregt E, Bakker B, van Ommen GJ, Breuning MH, Den Dunnen JT (2002) Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet* **71**:365-374

White SJ, Vink GR, Kriek M, Wuyts W, Schouten J, Bakker B, Breuning MH, Dunnen JT (2004) Two-color multiplex ligation-dependent probe amplification: Detecting genomic rearrangements in hereditary multiple exostoses. *Hum Mutat* **24**:86-92

Wiegant J, Kalle W, Mullenders L, Brookes S, Hoovers JM, Dauwerse JG, van Ommen GJ, Raap AK (1992) High-resolution in situ hybridization using DNA halo preparations. *Hum Mol Genet* **1**:587-591

Wilke K, Duman B, Horst J (2000) Diagnosis of haploidy and triploidy based on measurement of gene copy number by real-time PCR. *Hum Mutat* **16**:431-436

Yobb TM, Somerville MJ, Willatt L, Firth HV, Harrison K, MacKenzie J, Gallo N, Morrow BE, Shaffer LG, Babcock M, Chernos J, Bernier F, Sprysak K, Christiansen J, Haase S, Elyas B, Lilley M, Bamforth S, McDermid HE (2005) Microduplication and triplication of 22q11.2: a highly variable syndrome. *Am J Hum Genet* **76**:865-876

Yunis JJ (1976) High resolution of human chromosomes. *Science* **191**:1268-1270

Zhang L, Lu HH, Chung WY, Yang J, Li WH (2005) Patterns of segmental duplication in the human genome. *Mol Biol Evol* **22**:135-141

Zweier C, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J, Reardon W, Saraiva J, Cabral A, Gohring I, Devriendt K, de Ravel T, Bijlsma EK, Hennekam RC, Orrico A, Cohen M, Dreweke A, Reis A, Nurnberg P, Rauch A (2007) Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am J Hum Genet* **80**:994-1001

Appendix

1. MAPH / array-CGH request form
2. Colour pictures

N.B.: Please, send this form as attachment to
K.Szuhai@lumc.nl and M.Kriek@lumc.nl

Patient for MAPH and/or Array-CGH screening

Date of birth:	…… / …… / ……..	
DNA number /Isolation number:	D.. / ….	D1. ……
Gender:	M / F	
Severe developmental delay	YES / NO	
Mild developmental delay	YES / NO	
Dysmorphic features:	YES / NO	
Please, list:		
(Multiple) Congenital Abnormalities (MCA)	YES / NO	
Please, list		
Heart defects present:	YES / NO	
Positive family history:	YES / NO	
If yes, please specify		
Consanguinity:	YES / NO	
Perinatal onset growth retardation:	YES / NO	
Previously tested for:		
• Karyotyping:	YES / NO	P-number:
• Fragile X	YES / NO	
• microdeletion syndrome	YES / NO	
Outcome:		
MAPH screening :	YES / NO	
CGH screening:	YES / NO	
Responsible clinician:	Date:	

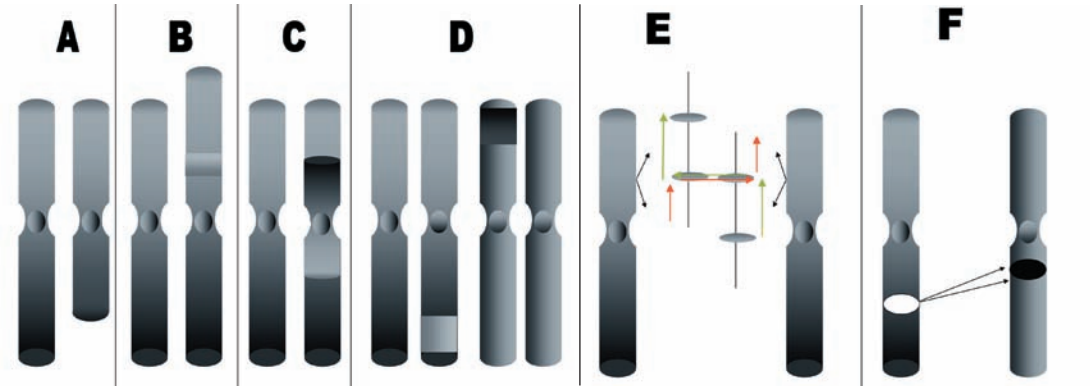
Colour pictures

Chapter I

p. 14 and 15

Figure 1. Deletion, duplication, inversion and balanced translocation.

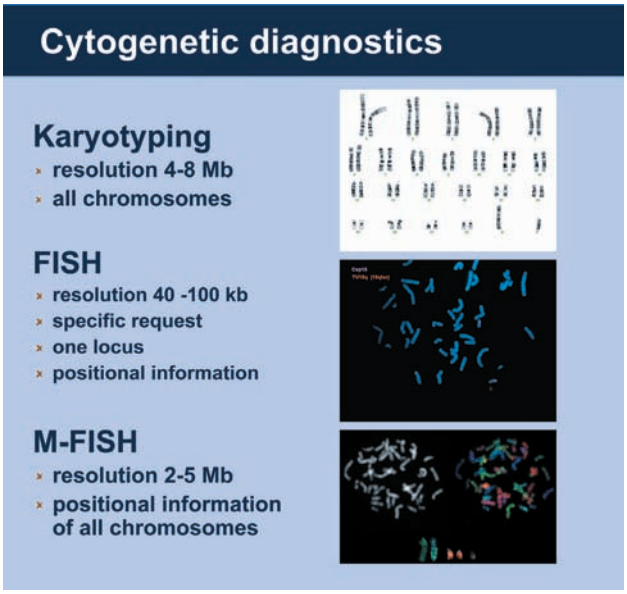
Figure 2. Non-allelic homologous recombination and insertions.



A. Part of the long arm of the right chromosome is missing. The loss of genomic material is called a deletion.
B. A part of the short arm of the chromosome is present twice (right). This extra material is called a duplication. As the duplicated region is localised within the chromosome, this duplication is called an interstitial duplication.
C. The amount of genetic material in part C of this picture is similar to the unaffected left chromosome. However, a part of the chromosome is inverted. As the centromere is localised within the invertion, this situation is called a pericentromeric inversion.
D. Again the amount of genetic material is normal, however, a part of the information of the dark grey chromosome has been transported to the light grey chromosome and vice versa. This is called a balanced translocation.
E. Non allelic homologous recombination. The two alleles of a chromosome contain regions that are highly homologous (e.g. segmental duplications, low copy repeats or duplicons). The presence of these segmental duplications can result in misalignment of these regions and subsequently in non allelic homologous recombination. The green arrow shows the origin of a duplication of the region present between two highly homologous regions, whereas the red arrow indicates the origin of a deletion.
F. In this situation a part of the left chromosome is inserted in another chromosome. This is called an insertion.

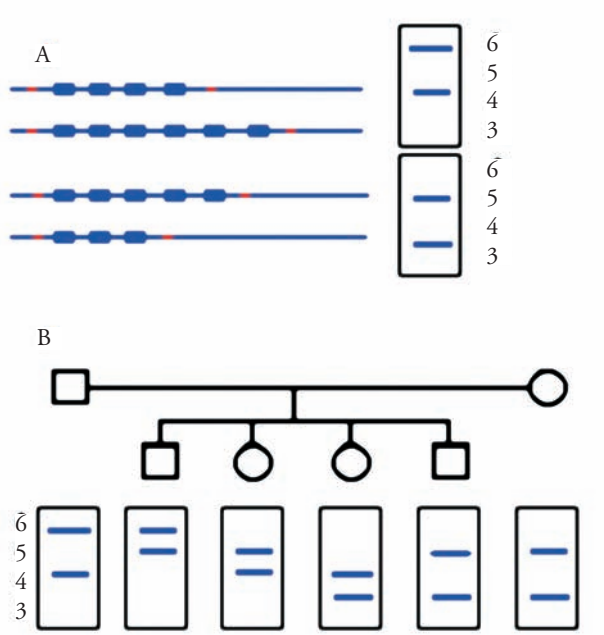
p. 25

Figure 3. Current standard cytogenetic diagnostic tools and their characteristics.



p. 31

Figure 4. Identification of the parental origin of an allele.

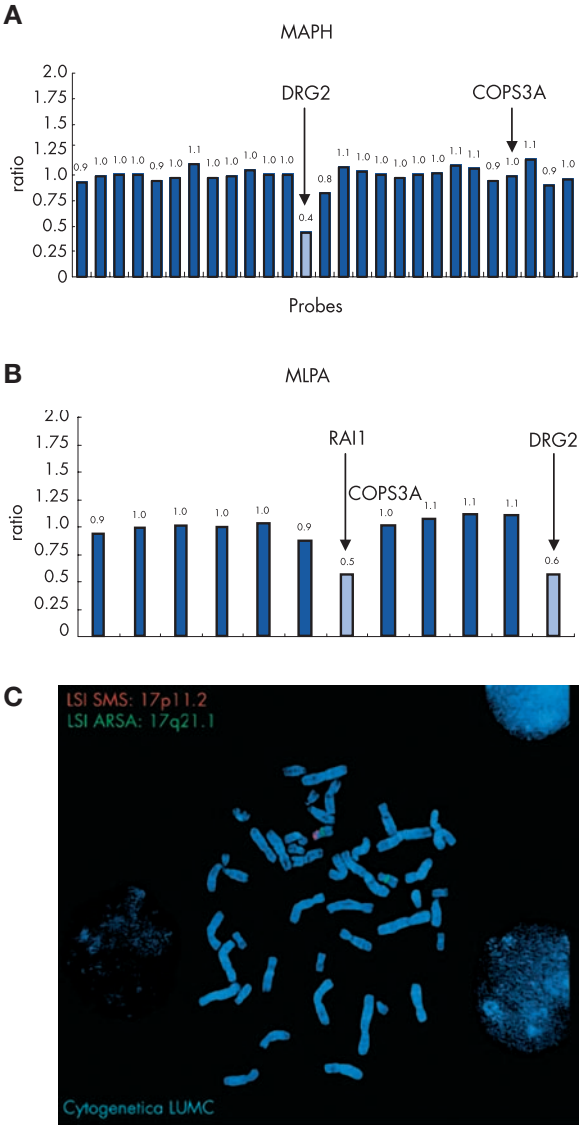


A. Different VNTR lengths in both parents present on a specific region in the human genome. B.
B. One of the children has the identical combination of VNTR lengths as one of its parents. Uniparental disomy (of genetic material from the parent with identical VNTR lengths) or a deletion present at the allele inherited from the 'other' parent should be considered. Picture derived from www.geninfo.no.

Chapter II-1

p. 52

Figure 1. Results of case 3.

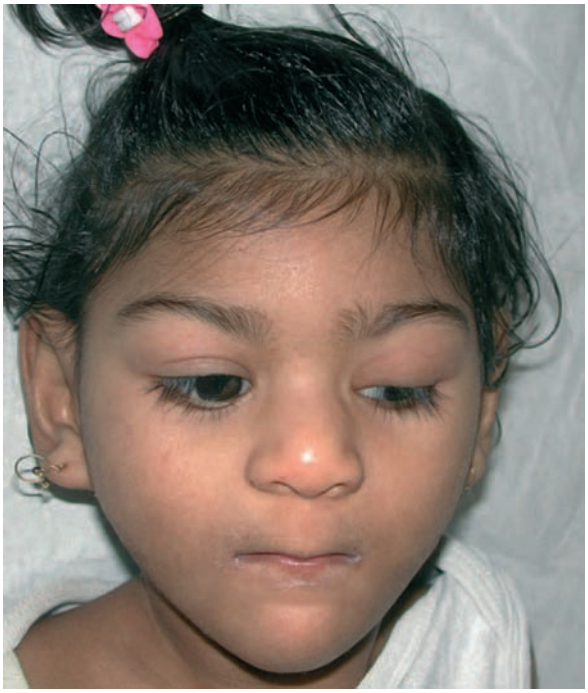


The plots correspond to the MAPH results showing (A) a deletion of the DRG2 gene, two normal copies of COPS3A (RAI1 not present), and the MLPA results; and (B) a deletion of RAI1, a deletion of DRG2, and a normal ratio of COPS3A. (C) The additional FISH analysis using the LSI-SMS probe specific for the Smith Magenis chromosomal region shows a normal signal on the short arm of only one copy of chromosome 17.

p. 54

Figure 3. Facial dysmorphism of case 6.

Note the microcephaly, ptosis of the left eye, flat philtrum, and thin upper lip.



Chapter III-1

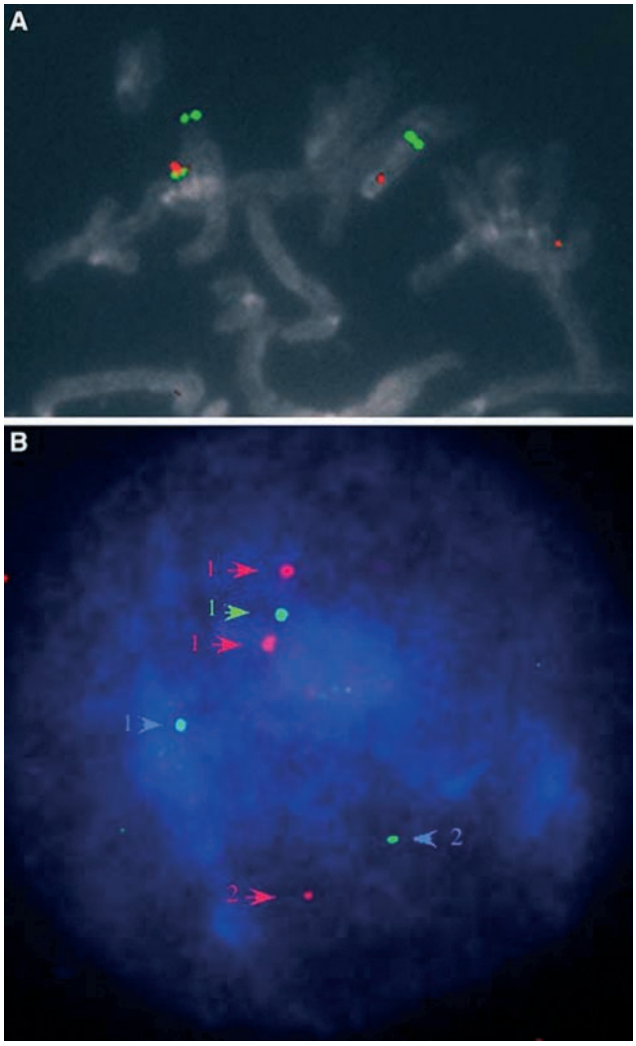
p. 100

Figure 2 Picture of the proband.

Note the microcephaly, myotonic facial expression, the proptosis of the eyes and the prominent simple ears.

p. 106

Figure 6. FISH analysis of chromosome 22.

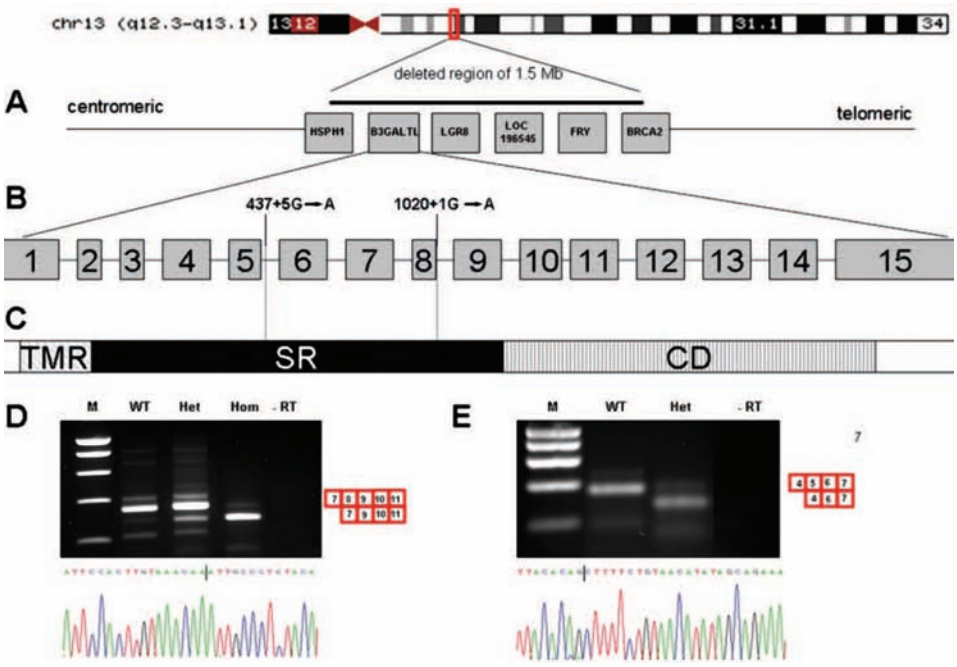


A A partial metaphase of the patient, hybridised with the telomere probe RP11-3018K1 (*green*; chromosome region 22q13), N25 (*red*; VCFS/DGS region) and RP11-66F9 (*green*; CES region). On the right chromosome, green signals of RP11-3018K1 (telomeric side of chromosome 22) and a red signal N25 corresponding to the VCFS/DGS region are present; however, the signal of RP11-66F9 is lacking, indicating a deletion of the CES region. On the left chromosome, in addition to the green signals of RP11-3018K1, a red signal corresponding to the VCFS/DGS regions and a green signal corresponding to the CES region are both present. These latest two signals are partly overlapping. On this chromosome, the signal of N25 is stronger than the signal on the right chromosome, suggesting a duplication of the VCFS/DGS region. These findings are confirmed by the result of the interphase nucleus depicted in part b of this figure. B The different chromosomes 22 are marked 1 and 2. The signal of LSI ARSA, corresponding to the telomeric side of chromosome 22, is indicated with a *blue arrow*. The *red arrow* indicates the N25 signal (corresponding to the VCFS/DGS region), which is duplicated in chromosome 22 nr.1 (two red signals). The *green arrow* indicates the signal of RP11-66F9 (corresponding to the CES region). This signal is missing on chromosome 22 nr.2, demonstrating the deletion of the CES region.

Chapter III-2

p. 115

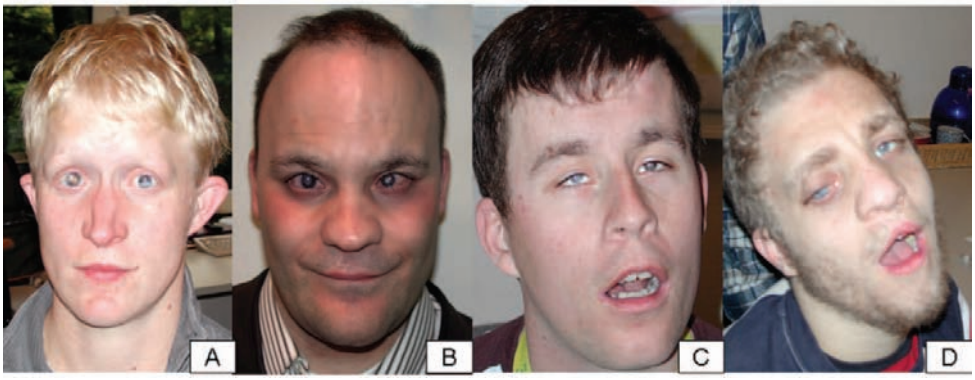
Figure 1. Overview of the location of the mutations in the *B3GALTL* gene and the results of the RT-PCR of RNA isolated from fibroblasts.



A, Genes present in the 1.5-Mb deletion found in two brothers with Peters Plus syndrome. *B,* 15 exons of the *B3GALTL* gene, with the localization of the mutations. *C,* B3GALTL protein, which consists of a transmembrane region (TMR), a stem region (SR), and a catalytic domain (CD). Both mutations (c.1020 1GrA and c.437 5GrA) are located in the stem region. *D,* Result of the nested RT-PCR of exons 7–11 of the *BGALTL* gene, with RNA derived from myoblasts (WT), RNA from fibroblasts of a father heterozygous for the c.1020 1GrA mutation (Het), and RNA from fibroblasts of his affected son with c.1020 1GrApat/delmat (Hom). The patient shows a smaller band compared with the WT band, which indicates a skip of exon 8. Sequence analysis of this band is shown. The vertical line indicates the end of exon 7 and the beginning of exon 9. The RT-PCR of the father shows, in addition to the WT band, a skipped product with much less intensity. *E,* Result of the RT-PCR encompassing exons 4–7 of the *BGALTL* gene, with RNA derived from lymphocytes of a control individual (WT) and a patient with a c.1020 1GrAmat/c.437 5GrApat genotype (Het). In addition to a faint WT band, the patient shows a smaller product that lacks exon 5. The sequence analysis of this smaller band confirms the skip of exon 5.

p. 116

Figure 2. Facial features of four patients with Peters Plus syndrome.

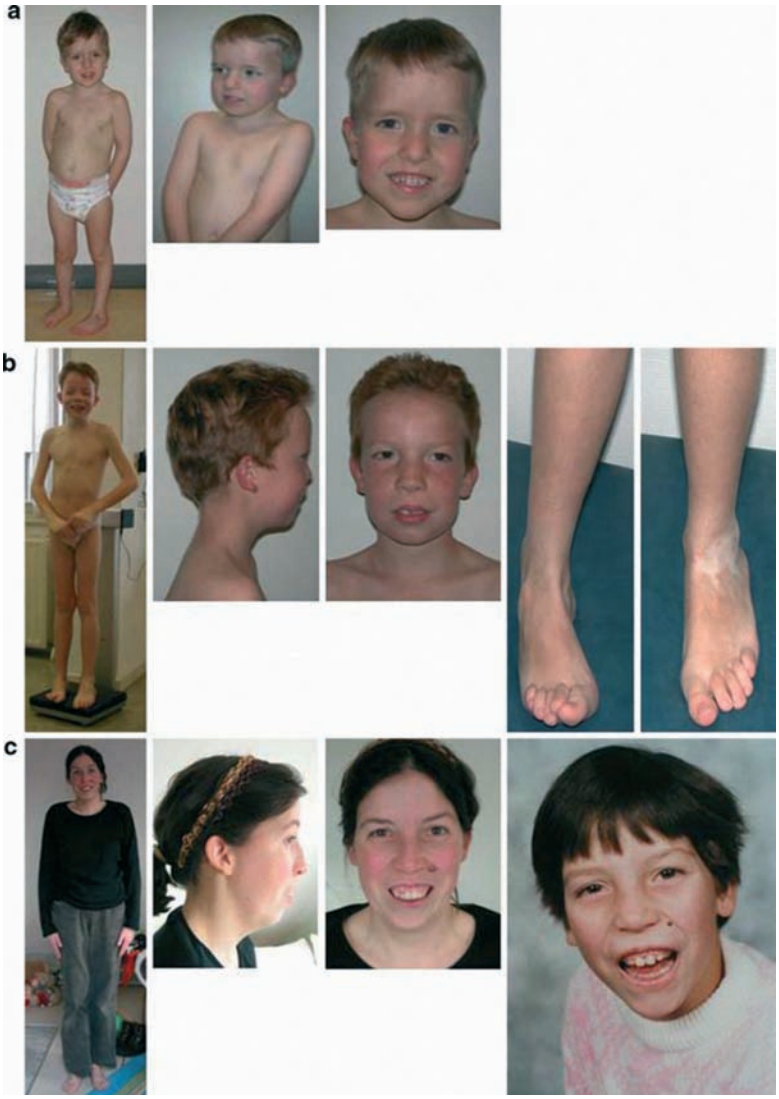


Patients A and C are homozygous for the c.1020+1G→A mutation. Patient B has the c.1020+1G→A_{mat}/c.437+5G→A_{pat} genotype, and patient D has the c.1020+1G→A_{pat}/del_{mat} genotype. Note the Peters anomaly of the eyes, the long face, and the Cupid's bow shape of the upper lip in all patients. Patients B and D have a repaired cleft lip and/or palate. Patient A is female; the rest are male.

Chapter III-3

p. 126

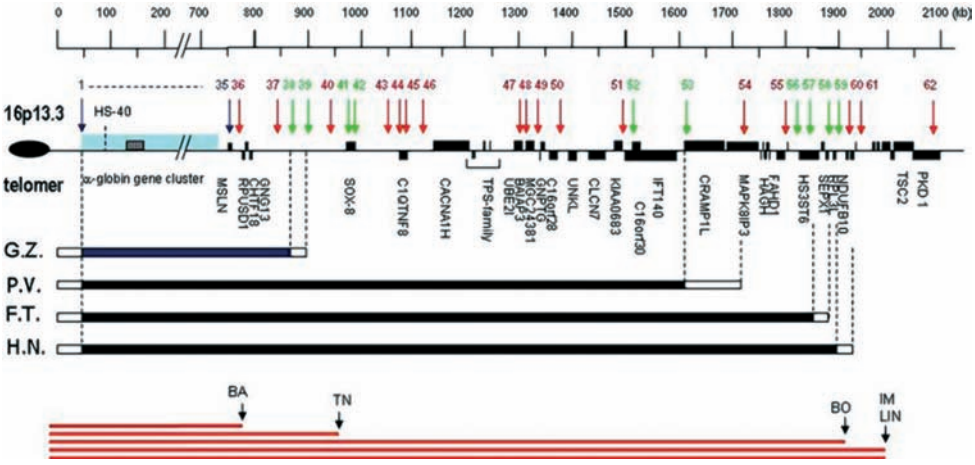
Figure 1. Three unrelated patients.



H.N. (a), P.V. (b) and F.T. (c) showing a mild mental retardation (IQ 50–60), a severe delay in active language ability, some typical facial features like downslanted palpebral fissures, mild hypertelorism, a broad nasal tip and small ears and a short neck with webbing, which is most pronounced in a and b. Patient H.N. and P.V. both show pectus carinatum. This was also observed for patient F.T. (not shown). H.N. also has an operated clubfoot on the left, while patient P.V.'s right foot is turned inside, the other foot showing a cafe´-au-lait spot. c Patient F.T. has a short neck and small ears. On the outer right a photograph is shown of the patient at age 11. The karyotype was normal in all patients and hematological analysis showed a persistent microcytic hypochromic anemia without iron deficiency

p. 129

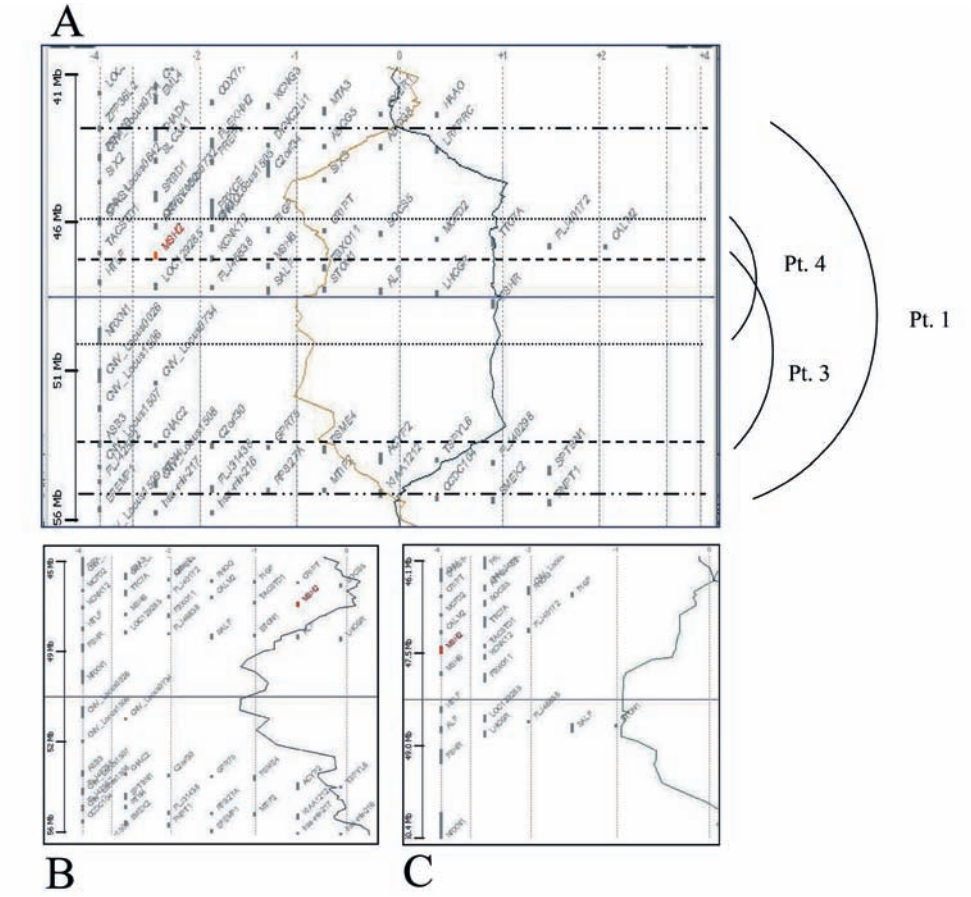
Figure 2 Schematic presentation of short arm of chromosome 16 (16p13.3), showing a 2 Mb region from the telomere containing the α-globin gene cluster up to the *TSC* and *PKD* genes.



The *arrows* and *numbers* represent the location of the probes. The deletions found by MLPA are shown as *bars* below the figure. Large deletions previously described are indicated as *red bars.*

p. 147

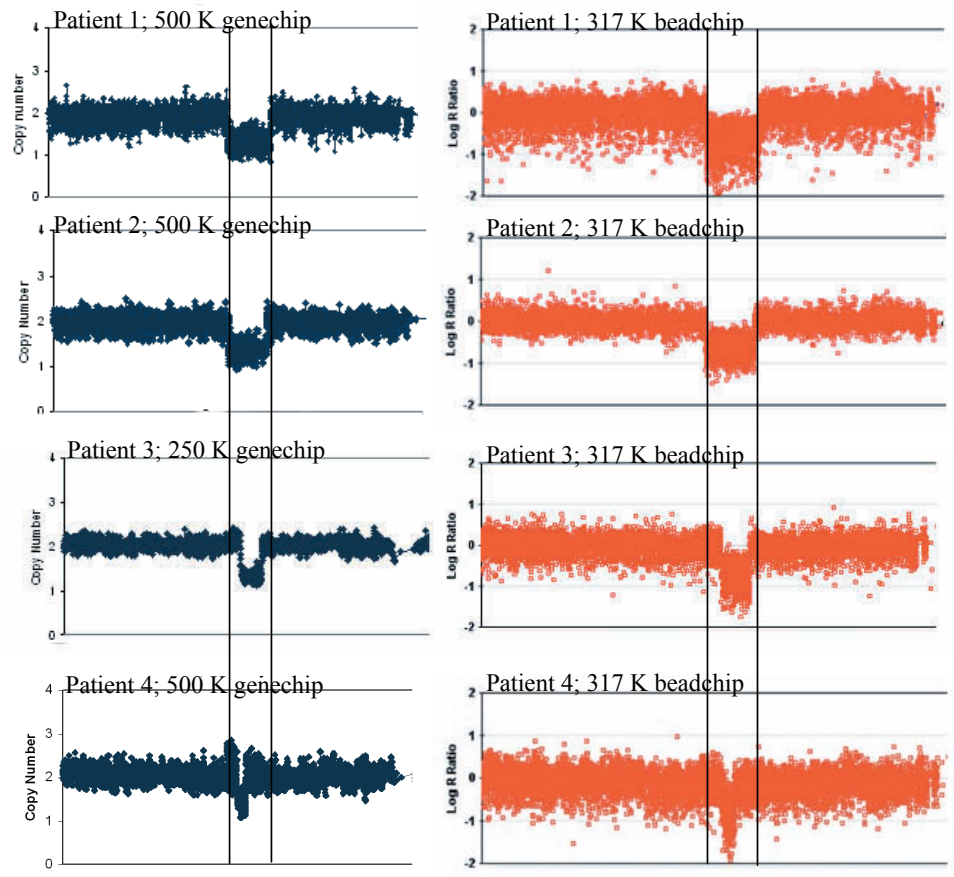
Figure 1. The result of patient 1 using the Agilent microarray platform.



(A) Due to the dye swap, the deletion is depicted in two colors resulting in a symmetrical profile pattern. All genes localized within the deleted region are visualized using the Agilent software tool. The deleted region of patient 3 (B) and patient 4 (C) are outlined by the dashed line and the dotted line, respectively. For the size of the deletion see table 2. Pt= patient.

p. 150

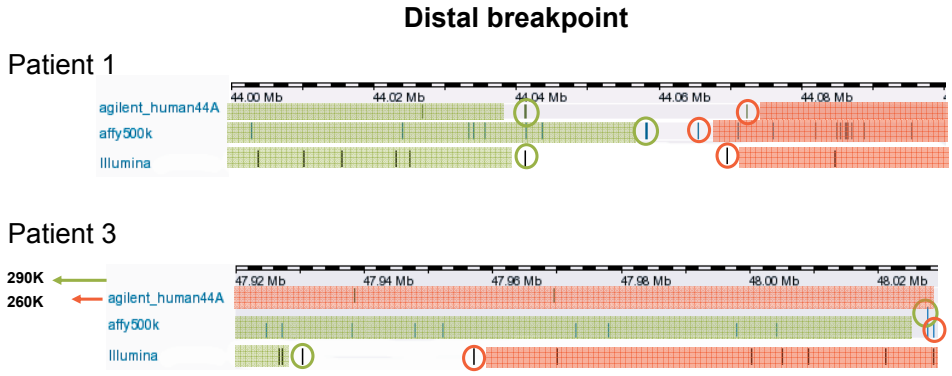
Figure 2. The interstitial 2p deletion of the four patients analysed by Affymetrix genechip (left) and the beadchip of Illumina (right).



The deletions of the different patients are shown separately. Patient 3 was only analysed using 250K NspI genechip. A normal copy number of two is represented by a copy number between 1.6 and 2.4 for the Affymetrix genechip or by a LogR ratio between –0.3 and +0.3 for the beadchip of Illumina. The vertical lines represent the size of the largest deletion. In general, the variation of the data points obtained by the beadchip is larger than that of the genechip. Especially in patient 3, the difference in variation is remarkable.

p. 152

Figure 3. Overview of the distal breakpoints of patient 1 and patient 3 defined by Agilent, the Affymetrix genechip and the beadchip of Illumina.



The deleted region is depicted in red, whereas regions showing two copies are depicted in green. A green circle represents the last data point that showed a normal copy of two. A red circle represents the first data point that showed a deletion.
The number of data points per platform is comparable at the location of the distal breakpoint of patient 1 and 3. In patient 1, the breakpoint mapping of all platforms is concordant. In contrast, there is a huge difference in breakpoint mapping in patient 3. According to the results obtained by Agilent platform, the distal breakpoint of the deletion is located 290-260K outside the most distal point of the picture (47,92 Mb) (green and red arrow). The results of the Affymetrix platform show that the deleted region starts more proximally at ~48.03Mb (black arrow). The beadchip of Illumina defines the distal breakpoint of the deletion between these two points.