Emerging patterns of cryptic chromosomal imbalances in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of the literature

Björn Menten, Nicole Maas, Bernard Thienpont, Karen Buysse, Jo Vandesompele, Cindy Melotte, Thomy de Ravel, Steven Van Vooren, Irina Balikova, Liesbeth Backx, Sandra Janssens, Anne De Paepe, Bart De Moor, Yves Moreau, Peter Marynen, Jean-Pierre Fryns, Geert Mortier, Koen Devriendt, Frank Speleman and Joris R Vermeesch

J. Med. Genet. published online 20 Feb 2006;
doi:10.1136/jmg.2005.039453

JMG Online First, published on February 20, 2006 as 10.1136/jmg.2005.039453

Emerging patterns of cryptic chromosomal imbalances in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of the literature

B Menten*, N Maas*, B Thienpont, K Buysse, J Vandesompele, C Melotte, T de Ravel, S Van Vooren, I Balikova, L Backx, S Janssens, A De Paepe, B De Moor, Y Moreau, P Marynen, J-P Fryns, G Mortier, K Devriendt, F Speleman, J R Vermeesch

*Both authors contributed equally to this work.

B Menten, K Buysse, J Vandesompele, S Janssens, A De Paepe, G Mortier, F Speleman, Center for Medical Genetics, Ghent University, Ghent; Belgium

N Maas, B Thienpont, C Melotte, T de Ravel, I Balikova, L Backx,, J-P Fryns, K Devriendt, J R Vermeesch, Center for Human Genetics, University Hospital Gasthuisberg, Leuven; Belgium

S Van Vooren, B De Moor, Y Moreau, ESAT-SISTA, University of Leuven, Leuven, Belgium
P Marynen, Flanders Interuniversity Institute for Biotechnology (VIB4), Department of Human Genetics, University of Leuven, Leuven, Belgium

Key-words: Array CGH, Mental retardation, Chromosomal mosaicism

Correspondence to: J.R. Vermeesch, Center for Human Genetics, Herestraat 49, 3000 Leuven, Belgium, Joris.Vermeesch@uz.kuleuven.ac.be

Background: Chromosomal abnormalities are a major cause of mental retardation and multiple congenital anomalies (MCA/MR). Screening for these chromosomal imbalances has mainly been performed by standard karyotyping. Previous array CGH studies on selected patients with chromosomal phenotypes and normal karyotypes suggested an incidence of 10-15% previously unnoticed *de novo* chromosomal imbalances.
Methods: Here we report on array CGH screening of a series of 140 patients (the largest series published thus far) with idiopathic mental retardation and multiple congenital anomalies (MCA/MR) but normal karyotypes.
Results: Submicroscopic chromosomal imbalances were detected in 20% (28/140) patients and included 18 deletions, 7 duplications and 3 unbalanced translocations. Seventeen from twenty four imbalances were confirmed *de novo* and 19 were assumed to be causal. Excluding subtelomeric imbalances, our study identified 11 (8%) clinically relevant interstitial submicroscopic imbalances. Taking into consideration this and previously reported studies, array CGH screening with a resolution of at least 1 Mb, has been performed on 432 patients with MCA/MR. Most imbalances are non-recurrent and spread across the genome.
Conclusions: In at least 8.8% (38/432) of these patients *de novo* intrachromosomal alterations have been identified. Hence, array CGH should be considered as an essential aspect of the genetic analysis of patients with MCA/MR. In addition, in our study 3 patients were mosaic for a structural chromosome rearrangement. One of these patients had monosomy 7 in as little as 7% of the cells, illustrating that array CGH allows the detection of low grade mosaicisims.

INTRODUCTION

Chromosomal abnormalities are a major cause of mental retardation and congenital malformations. Many chromosomal defects are readily detected by standard or high resolution karyotyping. However, at best, the resolution of cytogenetic analysis is limited to about 5 to 10 Mb. It has long been assumed that a considerable proportion of patients with multiple congenital anomalies and mental retardation (MCA/MR) have submicroscopic chromosomal imbalances, not detectable by routine karyotyping. Such hidden abnormalities were detected at the subtelomeric regions in ~5% of these patients.[1][2][3][4] Following the introduction of the principle of array CGH,[5][6] genome wide high resolution analysis for DNA copy number alterations became feasible. In analogy with karyotyping, genome wide array CGH has been termed molecular karyotyping.[7][8][9] The first papers by Vissers et al.[10] and Shaw-Smith et al.[11] reported as much as 15-24% of segmental aneusomies in patients with idiopathic mental retardation and dysmorphism. A few additional studies reported detection rates between 10-25%. [12][13][14] To evaluate the clinical relevance of a chromosomal imbalance, there is the need to collect genotype and phenotype information in a large number of patients. This will enable the determination of the incidence and the genomic distribution of disease causing imbalances and may reveal the underlying mechanisms causing chromosomal imbalances.
In this study we report array CGH data on a new series of 140 patients and review the findings of 292 previously reported patients in order to determine the overall incidence and clinical relevance of each of these chromosomal imbalances. In addition, we provide the first evidence that array CGH screening allows the detection of low grade mosaicism for chromosomal aberrations.

MATERIALS AND METHODS

Selection of patients

This was a collaborative study between the genetic teams of Leuven and Gent. Patients were selected for this study by clinical geneticists from both teams. The study was approved by the institutional review board and appropriate informed consent was obtained from human subjects. Individuals with mental handicap without known etiological (or cause), but in whom a chromosomal aberration was suspected because of the association with one or more major congenital malformation (such as congenital heart defect, cleft palate, brain malformation ...) and/or dysmorphism (i.e. three or more minor anomalies). Ages varied between 1 and 62 with a mean of 13.1 years of age. The number of males and females are about equal. All patients had a normal karyotype on G-banding analysis at ISCN +550. The presence of a subtelomeric abnormality was excluded by FISH or MLPA in 31/140 patients. Genomic DNA from each patient was either isolated from blood lymphocytes or cultured fibroblasts. When consented, full phenotypic descriptions of patients with anomalies have been submitted to the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER: http://www.sanger.ac.uk/PostGenomics/decipher/)

Array CGH

BAC arrays were developed from the 1 Mb clone set of the Sanger Institute which contains 3431 BAC and PAC clones as previously described.[8][15][16] In short, BAC and PAC DNA was isolated from 1 ml bacterial cultures was amplified by two rounds of DOP-PCR using an amino-linked primer in the second PCR [15] , and purified on Multiscreen purification plates (Millipore). Purified aminolinked PCR products were spotted in duplicate or triplicate at a concentration of 250 ng/µl on 3-D CodeLink Bioarray System slides (Amersham Biosciences, Piscataway, NJ) with a Lucidea spotter (Amersham Biosciences) or a QArrayMini spotter (Genetix). 300 ng DNA was labeled by a random prime labeling system (BioPrime Array CGH Genomic Labeling System, Invitrogen) using Cy3- and Cy5-labeled dCTPs (Amersham Biosciences). Probe concentration and labeling efficiencies were measured with the Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Rockland, DE). Following labeling, hybridization and washing of the slides, arrays were scanned at 532 nm and 635 nm using a GenePix 4000B scanner (Molecular Devices Corporation Union City, CA) or a GMS 418 scanner (MWG).

Image and data analysis

The scan images were processed with Imagene software (Biodiscovery, El Segundo, CA) and further analyzed with an in house developed and freely available software tool arrayCGHbase (http://medgen.ugent.be/arrayCGHbase/).[16] In short, spot intensities were corrected for local background and only spots with signal intensities at least 1.5 times above background were included in the analysis. Where useful, further normalization of the data was achieved by 2D Lowess normalization using the Bioconductor software.[17] Following this normalization, the values of the

duplicates/triplicates on the array and the duplicate experiments were averaged and a $\log_2$ value was calculated. If signal intensity ratios among replicate spots deviated more than twice the overall standard deviation of all intensity ratios, the spot was not further analyzed. At least 95% of the spotted clones fulfilled these quality criteria. The experiment was only scored succesfull if the standard deviation (SD) of the $\log_2$ of the overall spot intensity ratios was lower than 0.096. Typically, this SD value for a combined experiment is between 0.035 and 0.06. Clones that have been identified in previous control hybridizations and other studies as being polymorphic were excluded from the analysis [8][18] Of the 3431 targets on the array, 57 autosomal and 8 X chromosomal clones are considered to be polymorphic"

Two or more flanking targets exceeding a value of the mean of plus or minus four times the standard deviation (SD) of the $\log_2$ of all intensity ratios for that hybridization experiment were further investigated to confirm the presence or absence of a genomic imbalance. Single targets showing hybridization intensity ratios exceeding a value of $\pm |\log_2(3/2) - 2*SD|$ were also further validated. Validation was performed by metaphase FISH for all potential deletions and both metaphase and interphase FISH analysis for all potential duplications larger than 2 Mb in size. Real time quantitative PCR was used to confirm duplications smaller than 2 Mb in size. If two or more flanking clones the $\log_2$ of the combined intensity ratios exceeded the threshold value of 4xSD, FISH or real time quantitative-PCR experiments always confirmed the presence of a chromosomal imbalance. If the intensity ratio exceeded $\pm|\log_2(3/2)-2*SD|$ at only one isolated clone in both experiments, a false positive rate of one every 7 patients is observed.

FISH

Labeling of the DOP amplified BAC DNA that was used for spotting the arrays was performed by DOP-PCR on a thermocycler (GeneAmp9700, Applera, Nieuwekerk a/d Ijzer, The Netherlands). The reactions were performed in a total volume of 50 µl containing 5 µl of 15 µM DOP 1, 2, 3 primermix, 5 µl of 10X PCR buffer w/o $MgCl_2$, which is specially designed for use with Platinum® *Taq* DNA polymerase (Invitrogen, Carlsbad, CA), 5 µl 50 mM $MgCl_2$. For the dNTPs 1 µl 10 mM dATP, dCTP, dGTP each, 0.7 µl 10 mM dTTP, 1 µl of 1 mM SpectrumGreen™-, or SpectrumOrange™-dUTP (Vysis, Abbott laboratories, IL) or 5 µl 10X dNTP mixture containing 1 mM biotin-14-dCTP, 1 mM dCTP, 2 mM dATP, 2 mM dGTP, 2 mM dTTP in 10 mM Tris-HCl (pH 7.5), 1 mM Na2 EDTA (Bioprime DNA Labeling System, Invitrogen) was used. Platinum® *Taq* DNA polymerase (Invitrogen) (0.5 µl), 2 µl of the DOP amplified BAC DNA and H_2O to 50 µl were added. After initial denaturation at 95°C for 10 minutes, the reaction was as follows: 35 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 1 min and a final extension step of 72°C for 10 min.

Purification of the PCR product was performed with the Qiaquick 8 PCR purification kit (Qiagen NV, Venlo, The Netherlands) using QIAvac 6S vacuum according to instructions of the suppliers.

In addition to region specific BAC clones used for validation of array CGH results in patients with suspected imbalance, a chromosome 7 centromere specific probe was used for analysis of patient 19 with suspected monosomy 7 mosaicism (see Results). A total of 200 cells were screened for this patient and a control sample by two independent observers.

Before FISH, cells were air dried on slides and pretreated with pepsin followed by fixation with a 1% free formaldehyde solution and subsequently dehydrated with ethanol. After hybridization O/N at 37°C, the slides were washed for one minute in 0.4 x SSC/0.3% NP40 solution at 72°C, one minute at 2 x SSC/0.1% NP40 solution at RT and one minute at 2X SSC. The cells were counterstained with DAPI and the slides were mounted in Vectashield mounting medium (Vector Laboratories, Burlingame, CA). The signal was visualized by digital imaging microscopy with Cytovision capturing software (Applied Imaging, Santa Clara, CA). FISH was performed as described.[19].

Real time quantitative PCR (RTQ-PCR)

The oligonucleotides were selected by using PrimerExpress 2.0.0 ABI Prism oligo design software (Applied Biosystems, Lennik, Belgium). A penalty score lower than 150 was used to analyze the selected oligonucleotides further. The primers and amplicon were separately checked to exclude any repetitive sequences by using the BLAST program from the NCBI browser (http://www.ncbi.nlm.nih.gov/BLAST/) and the repeatmasker program (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker).

RTQ-PCR was performed using the qPCR mastermix Plus for SYBR Green I without UNG (Eurogentec, Liege, Belgium) according to the manufacturers instructions. The final volume of 25 µl contained 0,5 mM of both forward and reverse primers, 12,5 µl of 2x reaction buffer and 5 µl DNA

solution in the range of 2 to 50 ng per reaction. Total genomic DNA from human blood was purified by using an automated version of the purification protocol using Chemagic Magnetic Separation (Chemagen Biopolymer Technologie AG, Baesweiler, Germany).

PCR was carried out in triplicate from each fraction using 50 °C for 2 min and 95 °C for 10 min, followed by 40 cycles of: 95 °C for 15 s and 60 °C for 60 sec. A 81 basepair DNA fragment within the p53 gene (forward: 5'-CCC-AAG-CAA-TGG-ATG-ATT-TGA-3' and reverse: 5'-GAG-CTT-CAT-CTG-GAC-CTG-GGT-3') was used as a control amplicon (Eurogentec). Serial 5-fold dilutions of this target ranging from 100 ng to 0.16 ng per experiment served as a standard quantitation curve.

Real-time quantitative PCR was performed with the locus-specific oligonucleotides of interest on an ABI PRISM 7000 Sequence Detection System (SDS) according to the manufacturers instruction manual (Applied Biosystems, Lennik, Belgium). The amplification results and the melting curve were analyzed with the ABI Prism 7000 SDS Software version 1.1 (Applied Biosystems, Lennik, Belgium). The DNA levels were normalized to the gene p53 and relative differences were calculated according to the relative quantitation method.[20]

RESULTS

Array CGH findings in 140 patients with unexplained MCA/MR

140 patients with unexplained mental retardation and features suggestive of a chromosomal anomaly (e.g. a major malformation or multiple minor anomalies) were analysed on a 1 Mb BAC array. The DNA from each patient was labeled and hybridized with label swap versus the DNA of two other MCA/MR patients, rather than using a "normal" reference sample. Dye swap hybridizations for three patients in three hybridizations, reduces by half the number of experiments and cost per patient sample. This approach may be counter-intuitive and seem inappropriate in a diagnostic setting. However, *the* ideal reference genome is non-existent due to large scale copy number variations between the genomes of different "normal" individuals.[21][22] To mask benign copy number variation (CNVs), other groups have used pooled DNA of 7-10 different male or females as reference material [11][12][13][14]. For frequently occurring CNVs intensity ratios will be reduced. In case a CNV is present in 50% of the population, the intensity ratio difference at this locus would be reduced by half. Rather than improving the outcome, this result complicates data interpretation. One disadvantage of using patients as reference in three hybridizations could be that similar imbalances in two or three of the patients would result in equal intensity ratios for the affected region and potentially mask imbalances. However, the finding that the recurrence of a similar chromosomal imbalance in two patients with idiopathic MCA/MR is less than 1% (see below), makes the risk that a similar imbalance would occur in two and three independent patients lower than respectively $1/10^4$ and $1/10^6$.

In 28 patients (20.1%) a chromosomal imbalance was detected. An overview of all imbalances is shown in figure 1 and array CGH profiles for aberrant chromosomes are presented as supplementary information. Table 1 summarizes the genotype and phenotype of these 28 patients. For 8 patients the imbalance spanned over more than 5 clones (> 5 Mb in size), for 10 patients between 2 and 5 clones (1 - 8 Mb in size) and for 10 patients the imbalance was only a single clone (< 3 Mb). In two patients there was evidence of mosaicism for a structural chromosomal aberration and in one patient a low grade mosaicism for chromosome 7 monosomy was detected (see below). In 17/24 of the patients where the parents could be investigated the chromosomal imbalance was *de novo* by either FISH (deletions or duplications larger than 3 Mb) or qPCR (small duplications). While none of the imbalances smaller than 5 Mb could be detected by high resolution karyotyping, three large deletions (patients 7, 12 and 15) and two mosaics (patients 14 and 18) became apparent after retrospective analysis of the karyotype. Eight imbalances (5.7%) involved a subtelomeric region.

All *de novo* alterations can be considered causal for the MCA/MR phenotype observed in the patients. For 4/28 patients the parents were not available for genotyping. One of these (patient 1) had a large deletion on 1p36.2 spanning multiple clones. Since the observed phenotype in this patient resembles that of patients with known 1p terminal deletions, this imbalance was considered causal. For patients 3, 19 and 20 only one or two clones were abnormal making the causal relationship between genotype and phenotype difficult to determine.

For 7/28 patients the imbalance (three duplications and four deletions) was inherited from one of the parents. These parents were phenotypically normal with the exception of the father of patient 27 who presented with mild learning disabilities and the mother of patient 7 who was similarly affected as the daughter. Patient 27 presented with cleft-lip and palate, mild learning difficulties and a truncus arteriosus. A duplication on chromosome 22q11.2 was detected in this boy and his father. In view of previous reports describing 22q11.2 duplications (including inherited from normal parents), we assume a direct relationship between the 22q11.2 duplication and the observed phenotype in this patient. In

patient 10 and one of two imbalances in patient 7 have been listed as polymorphic in the Toronto polymorphism database.[21] In patient 7, the larger deletion on chromosome 5 spanning between 6.8

Table 1 Summary of copy number changes detected by array CGH, short clinical description and parental analysis

CASE	CLINICAL DETAILS*	MOLECULAR KARYOTYPE	PARENTS	N° CLONES	SIZE (MB)**	FLANKING CLONES
1	retinal dystrophy, growth retardation, short fingers, lowset ears, epicathic folds	46,XY.arr cgh del(1)(p36.23p36.32)	not determined	5	4,6-8,1	RP4-785P20, RP11-338N10
2	microcephaly, ventricular septal defect, large corneae, midface hypoplasia, presacral groove	46,XY.arr cgh del(1)(p36.31p36.32)	de novo	3	2,8-6,0	RP4-785P20, RP11-49J3
3	short stature, microcephaly, strabismus, unilateral renal agenesis, simple ears	46,XX.arr cgh del(1)(q21.1q21.1)	not determined	2	0,9-4,0	RP11-533N14, RP11-301M17
4	epilepsy, brachydactyly type E, scoliosis, absence of some toenails, synophrys	46,XX.arr cgh der(2)t(2;22)(q37;q13)	de novo	chr 2: 4 chr 22: 3	4,0-4,7 1,6-2,9	chr2: RP11-556H17, RP11-15L18 chr22: cN75H12, RP5-925J7
5	seizures, spasticity, hypotonia, hypoplastic cerebellum and brainstem, Dandy-Walker malformation	46,XY.arr cgh dup(3)(p12.2p12.2)	inherited (pat)	3	0,5-2,1	RP11-425D6, RP11-359D24
6	coarse facial features, Dandy-Walker malformation, wide pontine cisterns, right cerebellar lobe atrophy, hirsutism, pigmented nevi	46,XX.arr cgh del(3)(p12.1)	inherited (pat)	1	0,1-2,5	RP11-474M18
7	pectus excavatum, sacral dimples, recurrent infections, sparse eyebrows, small high nasal bridge	46,XY.arr cgh del(5)(q34q35.1)del(15)(q13.1)	Inherited (mat)¶ inherited (pat) ***	chr 5: 8 chr 15: 2	6,9-11,8 0,8-3,5	chr5:RP11-505G12, RP11-420L4, chr15:RP11-408F10, RP11-38E12
8	tetralogy of Fallot, double outlet right ventricle, hypertelorism, high and broad forehead, brachycephaly	46,XX.arr cgh del(5)(q35.1q35.1)	de novo	1	0,2-2,8	RP11-20O22
9	VSD, absent thumbs, growth retardation, hydronephrosis, preductal coarctation of aorta	46,XX.arr cgh del(7)(pterqter) .ish 46,XX[92]/45,XX.-7[8]	de novo	212	158	CTB-164D18,RP4-764O12
10	axial hypotonia, short stature, stereotypic movements, hypertelorism, strabismus	46,XY.arr cgh del(8)(q24.23q24.23)	inherited (pat)***	1	0,2-1,5	RP11-17M8
11	chondrodysplasia punctata brachytelephalangic type, obesity, short stature, small deeply set nose, hypotonia	46,Y.arr cgh der(X)t(X;9)(p22.32;p23)	de novo	chr 9: 16 chr X: 5	chr 9:13,0-13,9 chr X: 5,4-6,9	chr 9: RP11-187K14, GS1-77L23 chr X: RP11-60N3, CTB-98C4
12	short stature, microcephaly, VSD, preductal coarctation of aorta, midface hypoplasia	46,XX.arr cgh der(9)t(9;20)(q34.3;q13.33)	de novo	chr 9: 4 chr 20: 7	chr 9: 3,1-4,7 chr 20: 3,5-4,8	del:RP11-399H11, GS1-135I17, dup:RP5-836E13, CTB-81F12
13	hypotonia, spasticity, abdominal muscle hypoplasia, fine hair, macroglossia	46,XX.arr cgh del(9)(q34q34)	de novo	1	0,1-0,6	GS1-135I17
14	valvular pulmonary stenosis, cleft uvula, epilepsy, hypoplastic corpus callosum, hypoplastic genitalia	46,XX.arr cgh del(11)(q22.3q23.3)[66] / der(9)t(9;11)(qter;q21),del(11)(q22.3q23.3)[33]	de novo	chr 9: 6 chr 11: 50	del: 8,5-10,2 dup: 40,4-41,1	del:RP11-531F16, RP11-114K7, dup:RP11-685N10, RP11-469N6
15	carpal synostosis, macrocephaly, strabismus, oral frenulae, autistic behaviour	46,XX.arr cgh del(10)(q25.1q26.11)	de novo	10	8,2-10,3	RP11-271I13, RP11-355F22
16	broad thumbs, nasal speech, strabismus, deep hoarse voice, trigonocephaly	46,XX.arr cgh dup(13)(q31.3q33.1) .ish 46,XX[40]/46,XX dup(13)(q31.3q33.1)[60]	de novo	14	12,2-13,9	RP11-388D4, RP11-564N10
17	microbrachycephaly, almond shaped eyes, wide nasal bridge, large mouth, synophrys	46,XX.arr cgh del(15)(q22.2q22.2)	inherited (mat)	1	0,2-3,2	RP11-231A23
18	dysplastic ears, median cleft palate, small penis, brachycephaly, unilateral preauricular fistula	46,XY.arr cgh dup(16)(p13.2p13.3) .ish der(22)t(16;22)(p13.2p13.3;p21)	de novo	9	7,4-8,3	RP11-433P17, RP11-148F10
19	generalized hypotonia, scoliosis, congenital heart disease, short stature, brachycephaly	46,XY.arr cgh dup(17)(p13.3p13.3)	not determined	1	0,1-1,7	RP11-135N5
20	camptodactyly, ectropion, hypoplastic cerebellar hemispheres and vernis, hypertelorism, genital hypoplasia	46,XY.arr cgh del(17)(p12p12)	not determined	1	0,1-2,3	RP1-27J12
21	microcephaly, long eyelashes, long columella, deep presacral groove, lacrimal duct stenosis	46,XX.arr cgh del(17)(q11.2q11.2)	de novo	1	0,1-1,9	RP11-474K4
22	psychiatric disorder, macrocephaly	46,XX.arr cgh del(17)(q23.2q24.1)	de novo	3	1,1-4,2	RP11-115N5, RP11-74H8
23	small stature, narrow thorax, macrocephaly, downslanting palpebral fissures, prominent maxillary incisors	46,XY.arr cgh del(18)(q12.3q12.3)	de novo	2	1,4-4,6	RP11-486C18, RP11-463D17
24	joint laxity, scoliosis, hyperelastic skin, webbed neck, beaked nose	46,XX.arr cgh dup(20)(q13.13q13.2)	de novo	2	0,7-2,7	RP5-1071L10, RP5-994O24
25	myopia, nasal peech, cleft uvula, pulmonary stenosis, strabismus	46,XX.arr cgh del(22)(q12.2q12.2)	de novo	2	0,7-2,2	CTA-57G9, RP1-76B20
26	hypotonia, adduction of thumbs, claw toes, cut syndactyly fingers 3/4, dorsiflexion of the wrists	46,XX.arr cgh del(22)(q13.33q13.33)	de novo	3	1,4-1,9	CTA-722E9, CTB-99K24
27	cleft lip and palate, truncus arteriosus type I, short neck, ptosis, uteronephrosis,	46,XX.arr cgh dup(22)(q11.21 q11.21)	inherited (pat)	1	0,1-4,2	XX-91c
28	epilepsy, microcephaly, abdominal situs inversus, VSD, hypotonia	46,XY.arr cgh dup(X)(p21.3p21.3)	inherited (mat)	2	0,3-1,2	RP11-37E19, RP6-27C10

All patients presented with mental retardation. Only the five most relevant dysmorphic features are retained in the table. Full phenotypic descriptions together with the genotype data can be viewed in Ensembl(http://www.ensembl.org/index.html)* via the Decipher DAS server.

** sizes of the aberrations are shown from a minimal to maximal size in Megabases

*** This imbalance has already proven to be polymorphism

¶ Mother has the same phenotype including MR. Further family could not be investigated.

and 11.8 Mb was also present in the similarly affected mother. Hence, this deletion is likely to be causal for the phenotype. In patient 5, the duplicated region in the healthy father and son contains only a single gene, the glycogen branching enzyme (GBE1), dosage effect for this gene seems a rather unlikely cause. In patients 6, 17 and 28 single clone imbalances are inherited and the causal relationship between genotype and phenotype remain to be determined.

In summary, we consider that at least 19 out of the 28 observed imbalances are causal for the MCA/MR in the patients.

Cytogenetic features of (low grade) mosaic chromosomal imbalances

A further interesting observation in this study was the finding of three mosaics. In patient 16, array CGH revealed increased average intensity ratios for a 12 Mb region compatible with a duplication spanning the long arm of chromosome 13 from band 13q31.3 to 13q33.1 (Fig. 2a). The average $\log_2$ of the intensity ratio values of the abnormal clones was 0.38. Since the theoretical intensity ratio of a duplication is $\log_2$ (3/2) or 0.58, the estimated mosaicism level is 0.38/0.58 or 65%. FISH analysis confirmed the duplication to be present in 60% of cultured lymphocytes (Fig. 2b).

In patient 14, standard array CGH revealed a 5 Mb deletion at 11q22.1-23.1. FISH with clone RP11-87N22 confirmed the deletion at the 11q22.1 locus in all cells. 40 clones flanking this deletion (14 proximal and 26 distal to the deleted segment) showed intensity ratios with a mean of 0.21 suggesting a duplication of the adjacent region at 11q21-qter in approximately 35% of the cells (Fig. 2c). FISH with clone RP11-744N12 located within this presumed duplicated region showed a translocation of 11q21-qter onto chromosome 9 in 6% of the cells, in contrast to the estimated 35% (Fig. 2d). Since this FISH analysis was performed on lymphocytes following stimulation with phytohemaglutinin, and DNA used for array CGH was extracted from uncultured lymphocytes, we assumed that culturing resulted in clonal selection of the normal cells. FISH on uncultured lymphocytes confirmed this hypothesis and showed three signals of RP11-744N12 in as much as 25% of the nuclei of uncultured lymphocytes.

Array CGH analysis on patient 9 revealed an average intensity ratio of -0,0496 for the clones from chromosome 7 (Fig. 2e). The level of mosaicism is calculated to be 5%. Interphase FISH analysis by two independent observers using a centromere 7 specific probe revealed a single signal in 10.5% of the nuclei of peripheral white blood cells of the patient while in a control sample a single signal was only observed in 3.5% of the nuclei. The difference between these two proportions is statistically significant ($p<0.01$) thus confirming the presence of the monosomy in approximately 7% of the patient's white blood cells. This finding can probably be explained by the presence of a (pre)-malignant clone in this patient.

Literature review of MCA/MR patients with submicroscopic imbalances

In order to obtain insight into the incidence, characteristics and genomic distribution of imbalances detected by array CGH in MCA/MR patients, all published genomic imbalances were reviewed (Fig. 1 and table 2).[10][11][12][13][14] From a total of 192 patients screened by arrays at a ~1 Mb resolution, 41 (21%) imbalances were detected of which at least 20 (10%) were *de novo*. 113/192 patients were screened for subtelomeric imbalances before array CGH. The number of interstitial imbalances is 35 (18%) of which at least 17 de novo (8.8%). In addition, de Vries et al. analyzed 100 patients previously shown not to carry subtelomeric imbalances using an array covering the full genome and detected *de novo* alterations in 10 patients.[12] Also 5 imbalances were likely to be causal, but parents were not available for analysis. Of these 15 imbalances, 5 were smaller than 1 Mb.

Table 2. Literature review: summary of intrachromosomal copy number changes detected by array CGH

Paper	number of patients (*)	intrachromosomal			n° of targets on array
		De novo	Fam	Unknown	
Vissers L et al.[10]	20 (0)	2	2	1	3569
Shaw-Smith et al.[11]	50 (41)	7	5	0	~ 3500
Rosenberg et al.[13]	81(0)	4	7	3	~ 3500
Schoumans et al.[14]	41(41)	4	0	0	2600
this study	140 (31)	11	7	3	~ 3500
Total	**332**	**28**	**21**	**7**	
De Vries et al.[12]	100	10	0	5	32447

(*) Number of patients on which subtelomeric inbalances have been excluded before array CGH was performed

Figure 1 shows that the imbalances are more or less scattered across the genome and appear mostly randomly distributed over all chromosomes. Some chromosomal regions appear nonrandom involved. Interstitial aberrations at chromosome 1p36 were detected in two patients in the present study and in 3 published array CGH cases. Hence, in addition to the 1p36 terminal deletion syndrome, considered to be the most common subtelomeric microdeletion syndrome,[23] also interstitial subtelomeric deletions appear to be common. At two loci (1q21.1 and 5q35.1) both a duplication and a deletion were observed. Possibly, these sites may mark novel microdeletion syndromes caused by recurrent non-homologous recombination in low copy repeats. Of particular interest is the finding of a familial duplication on 22q11.2 in this study as well as in three previous reported cases (2 *de novo* and 1 case of unknown origin), further suggesting the recurrent nature of this duplication and the variable phenotypic effect.

DISCUSSION

This study represents the largest series of patients reported that have been screened for chromosomal imbalances with a 1 Mb resolution BAC array. In total of 140 patients 28 chromosomal imbalances (20%) were detected. These include 7 duplications, 18 deletions and 3 unbalanced translocations. In order to determine the causal role of these chromosomal aberrations, parents were investigated in 24 out of 28 patients. In addition, the Toronto database of normal variants was consulted. About three quarters (17/24) of the observed chromosomal aberrations were *de novo* and not reported before as a normal variant. In one patient for which the parents could not be tested, available phenotypic data for similar published cases indicated that the genotype could explain the observed phenotype and in one patient with inherited deletion the mother was equally affected. This brings the total of clinically relevant imbalances to 19. Taking into account these data and excluding those subtelomeric imbalances that could have been detected by FISH or MLPA/MAPH analysis, our study has identified 11 (8%) clinically relevant imbalances undetectable by karyotyping and subtelomeric screening. This is in accordance with previous findings of 10% to 15% causal interstitial submicroscopic imbalances in patients with MCA/MR.[10][11][12][13][14] Imbalances identified thus far in MCA/MR patients have been positioned on the human genome map in order to assess their genomic distribution and to detect overlapping regions. This map further confirms that most imbalances are scattered across the genome.

From our data and data from the literature it has become clear that the clinical application of array CGH poses new challenges. While it is assumed that *de novo* alterations result in the observed phenotype, only the recurrent association of imbalances with specific phenotypic features will reinforce this causal relationship. Hence, it will be essential to collect genotypic and phenotypic information on a large number of MCA/MR patients. In contrast to *de novo* alterations, many chromosomal imbalances are inherited. Although it is likely that frequently occurring genomic copy number variations (CNVs) may not have major disease causing phenotypic effects, rare variants such as the six familial inherited imbalances detected in this study, should be evaluated with care. In particular, imbalances of regions which are recurrently involved in familial transmission from a normal parent to affected children will pose specific problems for genetic counseling as illustrated by the 22q11.2 duplication. This is in line with previous observations that 22q11 duplications result in diverse phenotypes from normal over mild to severe and sharing a tendency for velopharyngeal insufficiency with DiGeorge/VCFS but having other distinctive characteristics as well.[24][25] The 22q11 duplication syndrome may hallmark a novel paradox encountered by molecular karyotyping as the causal relationship between a chromosomal anomaly and an associated phenotype becomes blurred. Hence, imbalances inherited from phenotypical normal parents may contribute to the phenotype through variable penetrance and/or expressivity, epigenetic effects or by uncovering a recessive mutation on the non-deleted allele. To understand the involvement of these variations in the observed phenotypes, it will be necessary not only to collect benign variation in the genome and collect information on *de novo* imbalances associated with disease phenotypes, but also to collect both genotype and phenotype information from patients with familial inherited imbalances and phenotypically normal parents. To start this data collection, both genotype and phenotype data from all patients who consented was submitted at the DECIPHER database (http://www.sanger.ac.uk/Postgenomics/decipher/).

Segmental chromosomal imbalances in mosaic state are causal in several MCA/MR syndromes.[26] The present study illustrates that array CGH may detect segmental chromosomal imbalances which may be overlooked in standard karyotyping when a small number of cells is

analyzed or when the abnormality is too small to arouse suspicion. A remarkable observation in one of the mosaics was that phytohemaglutinin stimulation of lymphocytes and subsequent short culture apparently induced a selective growth advantage of the normal cells. Clearly, such culture effects can bias the final cytogenetic observations as was observed in patient 14. Presently a theoretical model is being developed which should enhance the sensitivity for the detection of low grade mosaicism. Clearly, the presence of a large deletions present in as little as 5% of cells can easily be detected. The ability to detect low grade mosaics will allow the detection of chromosomal aneuploidies in highly contaminated specimens such as aborted fetuses[27] and in the analysis of tumors and leukemias[28].

In all reports, including this study, the number of deletions (57) was higher than duplications (24). This may have both a technical and a biological component. Technically, most threshold algorithms may favor more false negatives for duplication events as compared with deletion events. Most threshold algorithms determine cut-offs for both deletions and duplications at equal distance from the mean of all intensity ratios. Since the intensity ratios for chromosomal deletions are more distant from the mean (ratio of ½) as compared with the intensity ratios observed for duplications (ratio of 3/2), inevitably, there is a higher chance that some duplications may be missed. Secondly, there may be a biological bias. Duplications generally result in a milder phenotype. Therefore, there may be a selection bias in this patient population. In addition, the frequency of random duplication events in the human genome may be lower than the frequency of deletion events. Van Ommen[29] estimated the frequency of deletion events to be 1 in every 8 births and the duplication frequency 1 in every 50 births. This suggests the number of deletion events to be about six fold higher than the number of duplication events. In patients with MCA/MR, deletions outnumber duplications by approximately twofold.

In conclusion, we confirm that a high percentage of hitherto idiopathic MCA/MR is caused by submicroscopic chromosomal imbalances. Consequently, screening of selected patients with normal karyotypes seems desirable and feasible. The availability of commercial platforms and improved hybridization schemes resulting in reduction of costs for these analyses opens the way for implementation of array CGH in routine diagnostic analysis. At present it remains unclear which will be the optimal resolution of the array to screen MCA/MR patients. Higher resolution arrays may reveal higher numbers of small chromosomal imbalances. However, the finding of only 10% *de novo* imbalances in a cohort of 100 patients by a full coverage array may indicate that higher resolution not necessarily increases the diagnostic yield. More studies using high resolution arrays are needed to compare incidence of small imbalances in different patient populations. Nevertheless, using a 1 Mb resolution array, some imbalances smaller than 1 Mb are being missed. In addition, the false positive rate may be lowered, especially if the identification of imbalances is based on intensity alterations of three or more aberrant flanking clones.[12] Considering the large percentage of inherited chromosomal imbalances, establishing both benign copy number variations in the human genome as well as developing a comprehensive morbid map of the human genome will be of major importance in order to understand which imbalances are causative.

ACKNOWLEDGEMENTS

The authors wish to thank the MicroArray Facility, Flanders Interuniversity Institute for Biotechnology (VIB) for their help in the spotting of the arrays and the Mapping Core and Map Finishing groups of the Wellcome Trust Sanger Institute for initial clone supply and verification. This work was made possible by grants G.0200.03 from the FWO, OT/O2/40, GOA/2006/12 and Centre of Excellence SymBioSys (Research Council K.U.Leuven EF/05/007) from the University of Leuven and GOA-grant 12051203 from the University of Ghent.

Jo Vandesompele is supported by a grant from the Flemish Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT). Karen Buysse is a Research Assistant of the Research Foundation - Flanders (FWO - Vlaanderen). This text presents research results of the Belgian program of Interuniversity Poles of attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming (IUAP).

REFERENCES

1 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.
2 Knight SJ, Horsley SW, Regan R, Lawrie NM, Maher EJ, Cardy DL, Flint J, Kearney L. Development and clinical application of an innovative fluorescence in situ hybridization technique which detects submicroscopic rearrangements involving telomeres. *Eur J Hum Genet* 1997;**5**:1-8.
3 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.
4 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. Screening for subtelomeric rearrangements in 210 patients with unexplained mental retardation using multiplex ligation dependent probe amplification (MLPA). *J Med Genet* 2004;**41**:892-9.
5 Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997;**20**:399-407.
6 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.
7 Rauch A, Ruschendorf F, Huang J, Trautmann U, Becker C, Thiel C, Jones KW, Reis A, Nurnberg P. Molecular karyotyping using an SNP array for genomewide genotyping. *J Med Genet* 2004;**41**:916-22.
8 Vermeesch JR, Melotte C, Froyen G, Van Vooren S, Dutta B, Maas N, Vermeulen S, Menten B, Speleman F, De Moor B, Van Hummelen P, Marynen P, Fryns JP, Devriendt K. Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis. *J Histochem Cytochem* 2005;**53**:413-22.
9 Vermeesch JR. From chromosomes to molecular karyotyping. *Eur J Med Genet* 2005;**48**:211-3.
10 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, van Ravenswaaij-Arts CM, Knoers NV, van der Burgt I, de Jong PJ, Brunner HG, Geurts van Kessel A, Schoenmakers EF, Veltman JA. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* 2003;**73**:1261-70.
11 Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, Firth H, Sanlaville D, Winter R, Colleaux L, Bobrow M, Carter NP. 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**:241-8.
12 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. Diagnostic genome profiling in mental retardation. *Am J Hum Genet* 2005;**77**:606-16.
13 Rosenberg C, Knijnenburg J, Bakker E, Vianna-Morgante A, Sloos WC, Otto PA, Kriek M, Hansson K, Krepisch-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*. Published Online First: 24 June 2005. doi: 10.1136/jmg.2005.032268
14 Schoumans J, Ruivenkamp C, Holmberg E, Kyllerman M, Anderlid BM, Nordenskjold M. Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). *J Med Genet* 2005;**42**:699-705.
15 Fiegler H, Carr P, Douglas EJ, Burford DC, Hunt S, Scott CE, Smith J, Vetrie D, Gorman P, Tomlinson IP, Carter NP. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* 2003;**36**:361-74.
16 Menten B, Pattyn F, De Preter K, Robbrecht P, Michels E, Buysse K, Mortier G, De Paepe A, Van Vooren S, Vermeesch J, Moreau Y, De Moor B, Vermeulen S, Speleman F, Vandesompele J. arrayCGHbase: an analysis platform for comparative genomic hybridization microarrays. *BMC Bioinformatics* 2005;**6**:124.
17 Dudoit S, Gentleman RC, Quackenbush J. Open source software for the analysis of microarray data. *Biotechniques* 2003;**Suppl**:45-51.

18 Bauters M, Van Esch H, Marynen P, Froyen G. X chromosome array-CGH for the identification of novel X-linked mental retardation genes. *Eur J Med Genet* 2005;**48**:263-75.

19 Van Roy N, Laureys G, Cheng NC, Willem P, Opdenakker G, Versteeg R, Speleman F. 1;17 translocations and other chromosome 17 rearrangements in human primary neuroblastoma tumors and cell lines. *Genes Chromosomes Cancer* 1994;**10**:103-14.

20 Ginzinger DG. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol* 2002;**30**:503-12.

21 Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. Detection of large-scale variation in the human genome. *Nat Genet* 2004;**36**:949-51.

22 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**:525-8.

23 Battaglia A. Del 1p36 syndrome: a newly emerging clinical entity. *Brain Dev* 2005;**27**:358-61.

24 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**:1027-40.

25 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. Microduplication and triplication of 22q11.2: a highly variable syndrome. *Am J Hum Genet* 2005;**76**:865-76.

26 Schinzel A. Catalogue of unbalanced chromosome aberrations in man. Berlin-New York: Walter de Gruyter 2001.

27 Schaeffer AJ, Chung J, Heretis K, Wong A, Ledbetter DH, Lese MC. Comparative genomic hybridization-array analysis enhances the detection of aneuploidies and submicroscopic imbalances in spontaneous miscarriages. *Am J Hum Genet* 2004;**74**:1168-74.

28 Garnis C, Coe BP, Lam SL, MacAulay C, Lam WL. High-resolution array CGH increases heterogeneity tolerance in the analysis of clinical samples. *Genomics* 2005;**85**:790-3.

29 van Ommen GJ. Frequency of new copy number variation in humans. *Nat Genet* 2005;**37**:333-4.

LEGEND TO THE FIGURES

Figure 1: Overview of all published interstitial submicroscopic imbalances detected by array-CGH in patients with MCA/MR. Microdeletions and duplications identified in this study are represented by respectively a red and a green bar. Microdeletions and duplications identified by previous array CGH studies.[10][11][12][13][14] are indicated by respectively the orange and the blue bars. Polymorphic variants from de Vries et al. are not shown.

Figure 2: Cytogenetic analysis of patient 16 (panels A-B), patient 14 (panel C-D) with segmental chromosomal mosaicisms, and patient 9 (panel E) with a mosiacism monosomy of chromosome 7. (A) Partial molecular karyotype enlarging the ratio profiles for chromosome 13; in the X axis clones are ordered from the centromere to the q-arm telomere and the Y axis shows the $\log_2$ transformed intensity ratios at each locus. Red lines indicate the threshold for clone deletion or duplication (± 4*SD). (B) FISH with PAC 1091O16 confirms the duplication at 13q32The duplication was present in 60% of the cultured lymphocytes. (C) Partial molecular karyotype enlarging the ratio profiles for chromosome 11. In the X axis clones are ordered from the p-arm telomere to the q-arm telomere and the Y axis shows the $\log_2$ transformed intensity ratios at each locus. Red lines indicate the threshold for clone deletion or duplication (± 4*SD). (D) The duplication at 11q24.3 was confirmed with clone BAC 744N12 and was the result of a translocation between 11q and 9q. FISH on cultured and uncultured lymphocytes showed the duplication to be present in respectively 6% and 25% of the cells. (E) Molecular karyotype showing the ratio profiles for the chromosomes 1 to 22, X, and Y. Between the two vertical lines, chromosome 7 is positioned, and shows $\log_2$ transformed intensity ratios with an average of -0.05.