



Fondata nel 1241

University of Siena

Ph.D in Medical Genetics

Interpretation of molecular imbalances detected
by array-CGH analysis

Filomena Tiziana Papa

Supervisor: Prof. Alessandra Renieri

Academic Year 2009-2010

Ph.D dissertation board

Prof. Alessandra Renieri

Professor of Medical Genetics, University of Siena, Italy

Prof. Mario De Marchi

Professor of Medical Genetics, University of Torino, Italy

Prof. Francesca Gensini

Professor of Medical Genetics, University of Firenze, Italy

INDEX

1. INTRODUCTION	8
1.1 Historical perspective	9
1.2 Array - CGH Methodologies	13
1.3 Applications of array-CGH in clinical genetics.....	15
1.3.1 Application in known and unknown chromosomal syndrome.....	16
1.3.2 Application to identify new genetic defects in known syndrome.....	17
1.3.3 Application to detect the prevalence of mosaicism.....	18
1.3.4 Application to ascertain the frequency of copy number variants across the genome.	19
1.4 CNV in Autism and RETT syndrome.....	20
1.5 DECIPHER and Human Copy Number Variations database	22
2.RATIONALE, AIM and OUTLINE	24
3.MATERIALS and METHODS	27
3.1 Patients collection.....	28
3.2 Array-based CGH	28
3.2.1 Samples preparation.....	28
3.2.2 Human oligonucleotides array.....	28
3.3 Real-time quantitative PCR.....	30
3.4 Multiplex Ligation-dependent Probe Amplification (MLPA).....	31
4. RESULTS	32
4.1 Novel deletions/duplications	33

4.2 CNV in Autism and RETT syndrome	40
4.2.1 Genetic susceptibility factors in ASD patients	41
4.2.2 Genetic modifier factor	56
4.3 Human Copy Number Variations Database	69
5. DISCUSSION, CONCLUSIONS and FUTURE PERSPECTIVES	74
5.1 DISCUSSION.....	75
5.2 CONCLUSIONS AND FUTURE PERSPECTIVES.....	83
6. REFERENCES	84
<i>Curriculum Vitae</i>	90

...with love to my family

Acknowledgements

I would like to thank Prof. Alessandra Renieri for believing in me and giving me the opportunity to do my abroad experience. Her teaching and trust helped me believe in my possibilities and follow my goals.

Thanks to ...

... Chiara who introduced me to genetic and cared me at the beginning of my no-student-world-work.

... Rossella, pioneer of array technique and Eleni, Mafalda and Viria, we enjoyed working together!

... all my colleagues and friends of *Medical Genetics* staff, in particular Mirella, Francesca A., Francesca M. for their patience and Marzia, Annabella, Laura and Cinzia my favorite medical doctors.

... Rosangela who after a "stormy" start is became my inseparable family.

... Katia who always cheered me up.

... Giulio who always supported me.

... my new Lexington friends who in these last months helped me not feeling lonely.

A special thank to my "roommate" Maria Antonietta, I miss our conversation in the night!

The experience of life changes your own personality and sometime carries you far from your hearth and home, but you know you can always find friends who wait for you, listen to you and enjoy with you in every part of world we are. Grazia, Francesca, Nicoletta, Monica ... thanks, I am so glad to have friends like you!

1. INTRODUCTION

1. INTRODUCTION

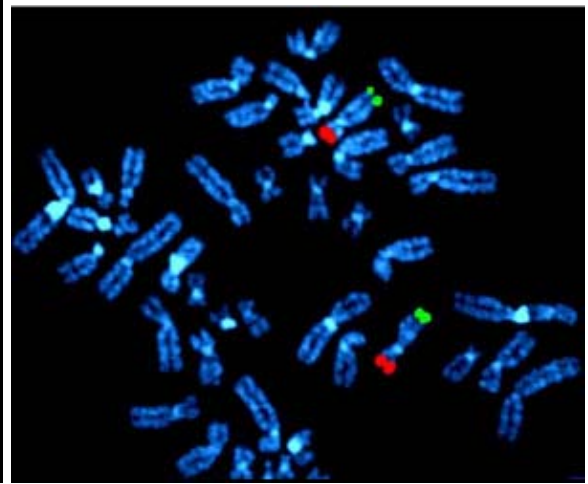
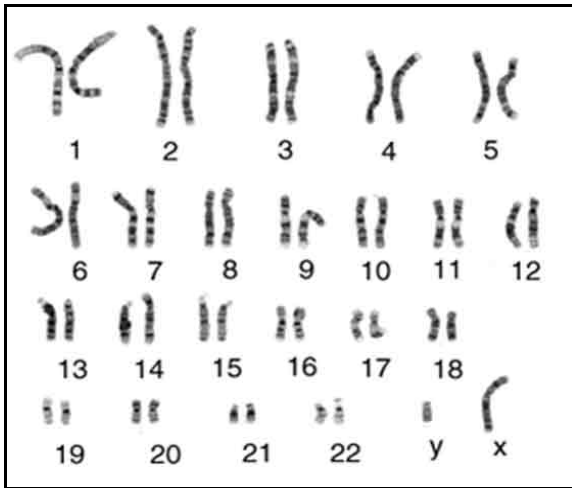
1.1 Historical perspective

Since karyotyping became a routine technique in clinical genetics, mental impairment, with or without other abnormalities, has often been found to be associated with chromosome rearrangements. However, in the majority of the patients, the *G* banded karyotype is normal, and in about half of them no obvious cause for the impairment is found. More recently, cryptic chromosome rearrangements have been reported in patients with an apparently normal karyotype and an unexplained abnormal phenotype. The best characterized rearrangements are the recurrent microdeletion syndromes, such as the Miller-Dieker lissencephaly and DiGeorge syndrome. In addition, subtelomeric imbalances of variable sizes cause intellectual disability in 5-7% of these cytogenetically "normal" cases. In the last few years, microarray-based comparative genomic hybridization (aCGH) analysis has become available, and appears to be a robust tool for detecting genomic imbalances in patients, with a much higher resolution than permitted by cytogenetic analyses based on chromosome banding. [1]

Conventional *G*-banded karyotyping remains the gold standard, but it is time-consuming (about 10-14 days is required to obtain the result), labor-intensive and the limitations to detect imbalances larger than 5-10 Mb. To overcome these limitations, rapid fluorescent in situ hybridization (FISH) and multiplex ligation-dependent probe amplification (MLPA) and micro-array platforms have been developed and introduced to this field to detect such cryptic genomic aberrations [2] (Fig. 1).

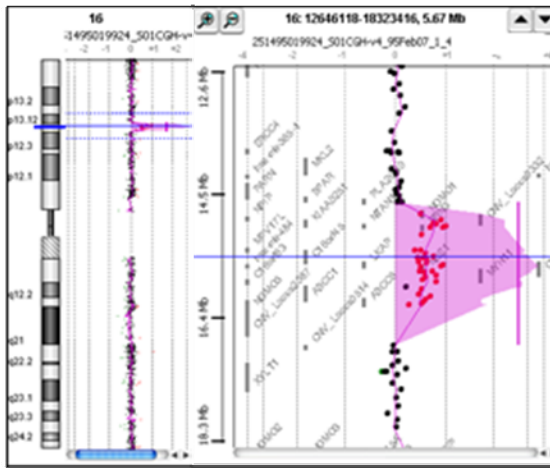
In the 1990s the introduction of molecular cytogenetic techniques into the clinical laboratory setting represented a major advance in the ability to detect known syndromes and identify chromosomal rearrangements of unknown

origin. FISH, which is the annealing of fluorescently labelled locus-specific probes to their complementary sequences in the genome, allowed the detection of specific microdeletion syndromes. [3] There are currently a number of commercially available FISH probes for the most common disorders and this method is still predominantly used when the clinical phenotype is suggestive of a particular disorder. Several other FISH-based methods, including spectral karyotyping (SKY), multicolour FISH (m-FISH), and comparative genomic hybridization (CGH) have proven extremely useful in the identification of unknown chromosomal material. SKY and m-FISH rely mainly on the principle of differentially labelling each chromosome using a unique combination of fluorochromes and are especially beneficial for identifying the origin and content of supernumerary marker chromosomes (SMCs) and complex chromosome rearrangements (CCRs) that involve more than two chromosomes. CGH was originally introduced for the cytogenetic analysis of solid tumors, which can be difficult to culture and involves the differential labeling of DNA from a test sample and a reference sample. The fluorescently labelled reactions are combined and hybridized to metaphase spreads from chromosomally normal individuals. Gains and losses of the genome in the test sample relative to the control sample are represented as ratios that are quantified from digital image analysis. This method allows the investigation of the whole genome and is very useful for determining the origin of unknown genetic material, such as SMCs and other unbalanced rearrangements. However, CGH does not detect balanced rearrangements, the resolution is on the order of 5-10 Mb and consequently many genomic disorders cannot be detected. [4]



A.

B.



C.

Fig.1 | A. G banded karyotype. B. Fluorescence in situ hybridization (FISH) of metaphase human chromosomes. C. aCGH result.

The need to screen the whole genome at a resolution that surpassed the existing technologies led to the implementation of microarray based CGH. The principle is very similar to that employed for traditional CGH, where two differentially labelled specimens are co-hybridized in the presence of Cot1 DNA. However, instead of metaphase spreads, the hybridization targets are DNA substrates immobilized on a glass slide. [5] [6] [7] Subsequently, the arrays are scanned and the resultant data are analyzed by software that computes the log₂ ratios for a variety of copy number differences between a patient and reference sample (Fig. 2).

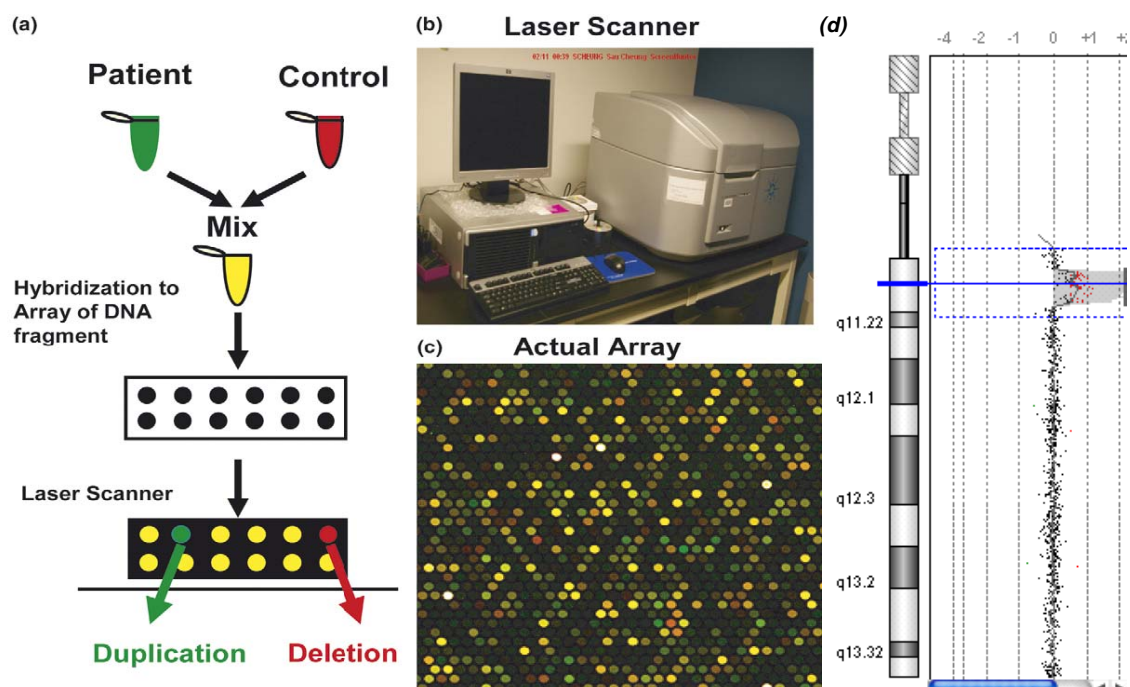


Fig. 2 | Schematic representation of an array-CGH experiment. a) Test and reference DNA are differentially labelled, co-precipitated and hybridised to an array. b) and c) After wash procedures, the slides are analysed through a scanner and fluorescence intensities of each probe are determined. d) After imaging processing and data normalization, the log₂ ratios of the probes are plotted as a function of chromosomal position. Probes with a value of zero represent equal fluorescence intensity ratio between sample and reference. Each dot represents a single probe spotted on the array. In this representation, copy number loss shift the ratio to the left and copy number gains shift the ratio to the right.

Consequently, aCGH is an entirely molecular technique with a cytogenetic application and represents a hybrid method that requires the expertise of both specialties. The current limitations of the technology include the inability to detect balanced chromosome rearrangements and the equivocal nature of copy number alterations of unknown significance that may be identified. Nevertheless, it is being used routinely in the clinical setting with a normal chromosome result in cases of intellectual disability and/or multiple congenital anomalies (ID/MCA); as a result the diagnostic yield in this patient group has increased considerably. [4]

1.2 Array - CGH Methodologies

Two major types of array targets are currently being utilized. Initially, bacterial artificial chromosomes (BACs) were the array target of choice. [6] However, more recently, oligonucleotide arrays have been adopted due to the increased genome coverage they afford. The design of both array types was made possible by the availability of the complete map and sequence of the human genome. The BAC arrays may contain DNA isolated from large insert clones that range in size from 150-200 Kb, spotted directly onto the array or may employ the spotting of PCR products amplified from the BAC clones. [8] These arrays are generally very sensitive and results can be directly validated with FISH using the BAC DNA as a probe. However, production of BAC DNA is labor-intensive and the resolution is limited to 50-100 Kb, even on a whole genome tiling path array. [9] Oligonucleotide arrays offer a flexible format with the potential for very high resolution and customization. Several different platforms are available for oligonucleotide arrays, some of which were adapted from genome wide SNP-based oligonucleotide markers and others that were created from a library of virtual probes that span the genome, and consequently can be constructed to have extremely high resolution. [10] Both BAC and

oligonucleotide arrays have been used successfully to detect copy number changes in patients with ID/MCA and autism. A number of different array design approaches have been taken for diagnostic purposes. A targeted array is one that contains specific regions of the genome, such as the subtelomeres and those responsible for known microdeletion/microduplication syndromes, but does not have probes that span the whole genome. [11] [12] [13] This type of array was initially used for clinical applications in postnatal specimens but has also been implemented for prenatal specimens with an abnormal ultrasound finding or for general screening purposes. [14] [15] [16] A whole genome or tiling path array offers full genome coverage with a resolution that is dependent on the spacing of the probes. For clinical testing the resolution generally involves a spacing of 50 Kb to 1 Mb between adjacent probes on the array often with additional coverage at the subtelomeric regions. [17] [18] The enhanced coverage of whole genome arrays identifies an additional 5% of abnormalities when compared to a targeted array. [19] [20] For research purposes, very high density oligonucleotide whole genome arrays and region specific custom arrays have been instrumental in defining new syndromes, detecting target gene deletions and characterizing breakpoint regions. [21] [22] [23]

1.3 Applications of array-CGH in clinical genetics

aCGH is a highly effective technique that is entering routine clinical use much faster than other microarray technologies. Indeed, compared to, for example, expression microarrays, aCGH enjoys several technical advantages: (1) genomic DNA samples are less prone to degradation than mRNA samples, (2) genomic DNA samples show much less variation between biological replicates than mRNA samples, and (3) interpretation of chromosomal imbalances is much easier than that of expression fingerprints. These advantages explain why aCGH for the diagnosis of constitutional anomalies is progressing faster towards the clinic than expression microarrays for the prediction of clinical outcome (e.g., in cancer), [24] for which a few applications are now entering clinical practice. aCGH have revolutionized cytogenetic diagnostics and, in turn, the clinical management of patients with developmental delays and multiple congenital anomalies [25] [26] [27] aCGH has given the clinician a greater appreciation of variability in the clinical presentation of many well-described conditions, [28] [29] has identified numerous previously unrecognized chromosomal syndromes with relatively mild phenotypes [30] [31] [32] [33] [34] [35] and has refined critical regions for established genetic defects. [36] aCGH allowed to appreciate the prevalence of mosaicism in individuals with developmental delay and ascertain the unexpected frequency of copy number variants (CNVs) across the genome. [37] Furthermore, the application of aCGH has created a paradigm shift in genetics that has moved the description and discovery of genetic conditions from the "phenotype-first" approach, in which patients exhibiting similar clinical features are identified prior to the discovery of an underlying etiology, to a "genotype-first" approach, in which a collection of individuals with similar copy-number imbalances can be examined for common clinical features. [38] Moreover we performed aCGH to detect recurrent genomic alterations in tumor tissue, underlying genetic susceptibility factors in a consistent number of

autism spectrum disorders (ASD) patients, to identify new genetic determinants in known syndrome.

1.3.1 Application in known and unknown chromosomal syndrome

"Recognizable syndromes" are recognizable because they exhibit, to the trained clinician, a constellation of signs and symptoms that arouse sufficient suspicion to cause the clinician to order a test that will confirm the clinical diagnosis.

The advent of molecular tools such as aCGH, allows us to define the rearrangements in a more detailed and comprehensive manner. A report has highlighted the usefulness of aCGH to characterize the Angelman syndrome/Prader-Willi syndrome (AS/PWS) region. [15] This age-old pattern of medical practice creates a loop that includes the patient, the clinician, and the laboratory and, in doing so, reinforces these recognizable features, cements them to the syndrome, and makes the clinician more confident in his/her diagnostic skills. With the application of aCGH to individuals with nonspecific developmental delay (DD) and/or intellectual disability (ID), with or without dysmorphic features (DF), it is now clear that many recognizable microdeletions and microduplication syndromes have a much wider spectrum of clinical presentation than was previously appreciated. [39] [25] A more complete understanding of the full clinical spectrum of these disorders will be achieved as the use of aCGH in the clinic becomes more prevalent and as correlations of these clinical findings with the genomic lesions are made. Existing website resources such as DECIPHER (see 1.5). [40] Many recognizable microdeletion syndromes are caused by nonallelic homologous recombination (NAHR) mediated by flanking low-copy repeat (LCR) sequences. [41] Interchromosomal and interchromatid NAHR between LCRs in direct orientation result in reciprocal duplication and deletion, whereas intrachromatid NAHR only creates deletion.

[42] This mechanism predicts that the prevalence of the reciprocal duplication product in the population should be 1:2 to the frequency of deletion. However, duplications have not been observed until fairly recently, likely because, in general, individuals with duplications tend to have a milder phenotype than those with the complementary deletions [43] [44] [45] [46] [47] and this milder phenotype may not lead to clinical investigation. [48] [49] The introduction of aCGH in clinical practice has virtually eliminated all the technical impediments of traditional cytogenetics and FISH and allowed the detection of such conditions with relative-but not complete-independence from the clinician's diagnostic judgment. Therefore, recent reviews of cohorts of patients ascertained with aCGH showed that the frequency of these duplications is much higher than heretofore appreciated. As aCGH becomes the primary method of testing individuals with even mild DD/ID, the frequency of microduplications at the common microdeletion syndrome loci will likely increase. [37] [50]

Molecular karyotyping is being increasingly applied to delineate novel disease causing microaberrations and related syndromes in patients with ID of unknown aetiology. The use of aCGH has allowed the discovery of many new syndromes in the last few years and has revitalized clinical cytogenetics. [51] [52] [53] [54]

1.3.2 Application to identify new genetic defects in known syndrome

Until recently, the cause of CHARGE syndrome, a sporadic malformation, was unknown. Vissers et coll. localized the gene for CHARGE syndrome by identifying and characterizing microdeletions by aCGH. CHARGE syndrome is a pleiotropic disorder comprising of coloboma, heart defects, choanal atresia, retarded growth and development, genital hypoplasia, ear anomalies and deafness. They tested 18 patients with CHARGE syndrome on a 1 Mb resolution genome-wide BAC array. One *de novo* microdeletion of 4.8 Mb was identified on

8q12. Another *CHARGE* patient originally reported with a balanced chromosome 8 translocation revealed a complex microdeletion partially overlapping with the one encountered in our index patient. No microdeletions were identified in 17 additional *CHARGE* patients tested on a tiling resolution chromosome 8 BAC array. Sequence analysis of nine genes located within the minimal region of deletion overlap revealed causative mutations in *CHD7*, a novel member of the chromodomain helicase DNA-binding gene family, in the majority of *CHARGE* patients without microdeletions. From these results, they concluded that *CHARGE* syndrome is caused by haploinsufficiency of the *CHD7* gene, either by a microdeletion encompassing the *CHD7* gene or by single base changes within this gene. This study showed that aCGH can indeed serve as an effective new approach to localize disease-causing genes. [55]

1.3.3 Application to detect the prevalence of mosaicism

Even though the effect of mosaicism on embryonic development and pregnancy outcome is not entirely clear, mosaic chromosomal imbalances have been shown to affect the development of in vitro-generated preimplantation embryos. [56] However, the detection of mosaicism in only 5% of aneuploid spontaneous miscarriages between 6-20 weeks gestation [57] and in only 12% of viable pregnancies screened by chorionic villus sampling (CVS) [58][59] indicates that the incidence of mosaicism decreases through the first and second trimesters of pregnancy and is even rarer in live births. This dramatic reduction in mosaicism from the early stages of embryonic development through the late stages of clinically established pregnancies suggests that there is significant selection against mosaicism. Nevertheless, detecting low-level mosaicism for clinically significant chromosome abnormalities remained a pressing diagnostic challenge for conventional cytogenetic testing until the advent of aCGH. The first systematic study of mosaicism in a large cohort identified mosaicism in as

little as 3% of the cells on the basis of metaphase counts. [60] aCGH has revealed a surprising prevalence of mosaicism in unselected populations of individuals who are referred for evaluation owing to DD, ID, and DF; the frequency of mosaicism among such individuals may be as high as 8-10%. [25] [60] Perhaps most significant is that aCGH does not require the stimulation of cell cultures by phytohemagglutinin, as classically done in clinical cytogenetics, which may distort the percentage of mosaic cells and inhibit the detection of some mosaic abnormalities by chromosome analysis. [61]

1.3.4 Application to ascertain the frequency of copy number variants across the genome.

Improvements in microarray resolution and hybridization robustness have resulted in the widespread implementation of genomic microarray technologies in medical research and diagnostics. This technology is most effective in detecting genomic deletions and duplications larger than 1 Kb, known as CNVs. Genomic microarrays are commonly used to identify rare, but highly penetrant, and commonly single CNVs in patients suffering from neurological disorders such as autism, [62] [63] schizophrenia [64] [65] and intellectual disability. [50] [66] [67] However CNVs have also been recently recognized as a common form of genomic structural variation: high resolution microarrays and sequencing approaches are able to identify 600-900 CNVs in a single individual. [68] [69] [70] [71] [72] Current clinical interpretation therefore needs to contrast the frequencies of a CNV in affected *versus* unaffected individuals, as well as determining the inheritance of CNVs via parental analysis. [73] [74] The size distributions of CNVs detected is dependent on the technology used. Increases in microarray resolution are revealing both a much higher rate of rare CNVs than previously thought [75] and an increasing number of genomic loci being reported that show variable inheritance and penetrance. There are limitations in

considering CNVs as either benign when common and inherited, or causal when rare and *de novo*. At present up to 5% of the human genome has been shown to vary in large scale copy number in numerous healthy controls [71] [76] and novel CNVs continue to be identified. In particular, CNV regions are enriched in repetitive sequences of near identical DNA known as segmental duplications [77] and are less prone to recombination. [78] The CNVs reported to date are documented in the Database of Genomic Variants (DGVs) (<http://projects.tcag.ca/variation>). During our analysis we found some rearrangements not reported in that database. We built, in collaboration with other Italian laboratory, an Human Copy Number Variations database (See 1.5).

1.4 CNV in Autism and RETT syndrome

The application of high-resolution genome analysis in research and clinical laboratories will uncover the genomic basis of many such disorders and will allow for better correlation of the many known CNVs with specific phenotypes. Although this promises to be a very challenging exercise, [79] [80] much work has already been initiated in cancer, neurological and neuropsychological conditions, infectious diseases, and others, suggesting that the clinical utility and applicability of such investigations cannot be too distant. Recent publications document the contribution of CNV to genetic diversity in humans and human disease. CNV has been shown to contribute to phenotype in model organisms and to important production and disease traits in domesticated livestock species. [81] A significant fraction of CNVs are likely to have functional consequences, due to gene dosage alteration, disruption of genes, positional effects or the uncovering of deleterious alleles. [82]

Autistic disorder is the prototypic form of a group of conditions, the autism spectrum disorders, which share common characteristics including

impairments in reciprocal social interaction and communication, and the presence of restricted and repetitive interests and behaviors. Individuals with an ASD vary greatly in language ability and cognitive development. For example, cognitive development ranges from above average to intellectual disability. [83]

Autism was first described by Leo Kanner in 1943 as a childhood developmental disorder characterized by the "inability to relate themselves in the ordinary way to other people" and by "insistence on sameness". The diagnostic criteria include qualitative impairment of reciprocal social interaction and communicative development and restricted interests and repetitive behaviors. The prevalence of autism is 1:500 with a 4:1 male to female ratio. In addition to autism, ASDs include Asperger syndrome, pervasive developmental disorder not otherwise specified, and a few specific conditions such as Rett syndrome. [84]

Twin and family studies indicate a predominantly genetic basis for ASD susceptibility and provide support for considering these disorders as a clinical spectrum. Concordance of monozygotic twins for autistic disorder is ~ 60%, but rises to ~ 90% when less severe cognitive and social deficits are considered. [85] Additionally, subclinical autistic traits are sometimes observed in first-degree relatives [86] and, to a lesser extent, in the general population. [87] The architecture of ASD genetic risk is complex. [88] Some 5-15% of individuals with an ASD have an identifiable genetic etiology corresponding to known rare single-gene disorders (e.g., fragile X syndrome, tuberous sclerosis 1 (*TSC1*), tuberous sclerosis 2 (*TSC2*), methyl CpG binding protein 2 (*MECP2*) in Rett syndrome and phosphatase and tensin homolog (*PTEN*)) and chromosomal rearrangements (e.g., maternal duplication of 15q11q13). [89] [90]

Rett syndrome RTT is an X-linked dominant neurodegenerative disorder that predominantly affects women. Individuals with RTT show progressive

deficits beginning at 6 to 18 months that include atrophy of verbal and skilled motor abilities, social withdrawal, hand wringing, respiratory difficulties, seizures, and autism spectrum behavior. RTT symptoms are variable and depend on the pattern of X-chromosome inactivation of the mutant allele, 121, 125-127 the nature of the *MECP2* mutation, and epigenetic factors. [91] The syndrome has been the object of extensive investigations, revealing a wide spectrum of clinical phenotypes including the classic form, the early onset seizure variant, the Zappella variant (Z-RTT), the congenital variant, the 'forme fruste' variant, and the late regression variant. [92] The majority of cases of RTT are the result of de novo mutations in *MECP2* gene that can arise in either parent, but more frequently occur in the paternal gamete. Mutations in the this gene account for half of girls with classic RTT. This mutation is prenatally lethal in males. A few surviving RTT male cases have been reported, but these were generally mosaic for *MECP2* mutations. [93] *MeCP2* inhibits transcription by binding DNA methylated at CpG dinucleotides and translation via direct interaction with RNA. Mutations in another gene, *CDKL5* which encodes cyclin-dependent kinase-like 5 and *FOXP1* gene, transcription factors which may play a role in the development of the brain and telencephalon, have also been implicated in a Rett like syndrome phenotype.

1.5 DECIPHER and Human Copy Number Variations database

As we already mentioned before, many patients suffering from developmental disorders harbor submicroscopic deletions or duplications that, by affecting the copy number of dosage-sensitive genes or disrupting normal gene expression, lead to disease. However, many aberrations are novel or extremely rare, making clinical interpretation problematic and genotype-phenotype correlations uncertain. Identification of patients sharing a genomic rearrangement and having phenotypic features in common leads to greater

certainty in the pathogenic nature of the rearrangement and enables new syndromes to be defined. To facilitate the analysis of these rare events, an interactive web-based database has been developed called DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources; <http://www.sanger.ac.uk/PostGenomics/decipher/>) (Wellcome Trust Sanger Inst. 2007). It may facilitate widespread appreciation of such phenotypic variability, which incorporates a suite of tools designed to aid the interpretation of submicroscopic chromosomal imbalance, inversions, and translocations. DECIPHER catalogs common copy-number changes in normal populations and thus, by exclusion, enables changes that are novel and potentially pathogenic to be identified. DECIPHER enhances genetic counseling by retrieving relevant information from a variety of bioinformatics resources. Known and predicted genes within an aberration are listed in the DECIPHER patient report, and genes of recognized clinical importance are highlighted and prioritized. DECIPHER enables clinical scientists worldwide to maintain records of phenotype and chromosome rearrangement for their patients and, with informed consent, share this information with the wider clinical research community through display in the genome browser Ensembl. By sharing cases worldwide, clusters of rare cases having phenotype and structural rearrangement in common can be identified, leading to the delineation of new syndromes and furthering understanding of gene function. [40]

The possible differences in frequency of particular CNVs in different populations and the variety of different CNVs observed in a given study may be significantly limited by the number and ethnic origin of individuals examined. From a collaboration with some Italian laboratories, an interactive web-based database, who collect rearrangements, *de novo* and inherited, not described in DGVs, has been developed called Human Copy Number Variations Database (HCNVs) (<http://dbcnv.oasi.en.it/gvarianti/index.php>).

2.RATIONALE, AIM and OUTLINE

2. RATIONALE, AIM and OUTLINE OF THE STUDY

The objectives of this study were: (a) to discover emerging syndromes in a cohort of patients with ID and MCA, (b) to detect the underlying genetic susceptibility factors in a consistent number of ASD patients, and (c) to explore the hypothesis that CNVs leading to over/underexpression of genes may modulate the phenotype by comparing Rett patient with different mutation. It was hypothesized that accurate characterization of the underlying molecular defect could help predict clinical outcomes, and therefore aid with early, appropriate interventions. Moreover all inherited or *de novo* rearrangements, not reported in DGVs, have been collected in HCNVs database.

By the use of nucleotide array with 44,000 and 99,000 probes and an average resolution of about 100-130 Kb (44K, Agilent) and 50-65 Kb (105K, Agilent), respectively, we have analyzed 696 patients with mild to severe intellectual disability associated with facial dysmorphisms and/or congenital anomalies. We found 426 (61%) negative while in 88 patients (13%) the analysis is still ongoing. We identified an inherited or *de novo* rearrangement in 165 cases (24%) while in 17 cases (2%) we detected novel *de novo* deletions not reported in the literature.

We selected 95 unrelated patients with ASD. In addition to the well known recurrent rearrangements involving the 15q11q13, 16p13 and 22q13 regions, recurrent microdeletions and microduplications at 16p11.2 have been recently identified and are shown to confer susceptibility to ASDs in up to 1% of autistic patients. In fifteen patients (16%) CNVs were considered disease associated, while in 20 patients (21%) CNVs were classified as uncertain cofactor.

We compare two discordant pairs of Rett sisters and four additional discordant pairs of unrelated Rett girls matched by mutation type to identify one major common modifier gene/region.

All patients described in this thesis have been inserted in the DECIPHER database (<http://www.sanger.ac.uk/PostGenomic/decipher/>) and HCNVs database (<http://dbcnv.oasi.en.it/gvarianti/index.php>) permitting us to create collaborations and extend the clinical phenotype associated with new microdeletion/microduplication syndromes .

3.MATERIALS and METHODS

3. MATERIALS & METHODS

3.1 Patients collection

Patients with ID and MCA enrolled in this study have been selected among those attending the Medical Genetics Unit of the University Hospital of Siena. All they were evaluate in genetic counseling and a clinically recognizable condition was excluded a diagnosis of a recognizable syndrome, and all patients.

3.2 Array-based CGH

3.2.1 Samples preparation

Genomic DNA of normal controls was obtained from Promega. Genomic DNA was extracted from peripheral blood using a QIAamp DNA Blood Maxi kit according to the manufacturer protocol (Qiagen, www.qiagen.com). The OD260/280 method on a photometer was employed to determine the appropriate DNA concentration. [94] Patient and control DNA samples were sonicated to produce a homogeneous smear DNA extending from approximately 600 bp to 2 Kb. DNA samples were then purified using the DNA Clean and Concentrator kit (Zymo Research, Orange, CA). Ten micrograms of genomic DNA both from the patient and from the control were sonicated. Test and reference DNA samples were subsequently purify using dedicated columns (DNA Clean and Concentrator, Zymo research, CA92867-4619, USA) and the appropriate DNA concentrations were determine by a DyNA Quant™ 200 Fluorometer (GE Healthcare).

3.2.2 Human oligonucleotides array

Array based CGH analysis was performed using commercially available oligonucleotide microarrays containing about 43,000 60-mer probes with an

estimated average resolution of about 100-130 Kb (Human Genome CGH Microarray 44B Kit, Agilent Technologies) and microarrays containing 99,000 60-mer probes with an estimate average resolution of 50-65 Kb (Human Genome CGH Microarray 105A Kit, Agilent Technologies). Physical positions of the probes correspond to the UCSC genome browser - NCBI build 36, March 2006. (<http://genome.ucsc.edu>). DNA labelling was executed essentially according to the Agilent protocol (Oligonucleotide Array-Based CGH for Genomic DNA Analysis 2.0v) using the Bioprime DNA labelling system (Invitrogen). Genomic DNA (2 µg) was mixed with 20 µl of 2.5X Random primer solution (Invitrogen) and MilliQ water to a total volume of 41 µl. The mix was denatured at 95° C for 7 minutes and then incubated in ice/water for 5 minutes. Each sample was added with 5 µl of 10X dUTP nucleotide mix (1.2 mM dATP, dGTP, dCTP, 0.6 mM dTTP in 10 mM Tris pH 8 and 1 mM EDTA), 2.5 µl of Cy5-dUTP (test sample) or 2.5 µl of Cy3-dUTP (reference sample) and with 1.5 µl of Exo-Klenow (40 U/µl, Invitrogen). Labeled samples were subsequently purified using CyScribe GFX Purification kit (Amersham Biosciences) according to the manufacturer protocol. Test and reference DNA were pooled and mixed with 50 µg of Human Cot I DNA (Invitrogen), 50 µl of Blocking buffer (Agilent Technologies) and 250 µl of Hybridization buffer (Agilent Technologies). Before hybridization to the array the mix was denatured at 95°C for 7 minutes and then pre-associated at 37°C for 30 minutes. Probes were applied to the slide using an Agilent microarray hybridization station. Hybridization was carried out for 24/40 hrs at 65°C in a rotating oven (20 rpm). The array was disassembled and washed according to the manufacturer protocol with wash buffers supplied with the Agilent kit. The slides were dried and scanned using an Agilent G2565BA DNA microarray scanner. Image analysis was performed using the CGH Analytics software v. 3.4.40 with default settings. The software automatically determines the fluorescence intensities of the spots for both fluorochromes performing

background subtraction and data normalization, and compiles the data into a spreadsheet that links the fluorescent signal of every oligo on the array to the oligo name, its position on the array and its position in the genome. The linear order of the oligos is reconstituted in the ratio plots consistent with an ideogram. The ratio plot is arbitrarily assigned such that gains and losses in DNA copy number at a particular locus are observed as a deviation of the ratio plot from a modal value of 1.0.

3.3 Real-time quantitative PCR

Some aCGH data were confirmed by Real-time Quantitative PCR experiments. To design adequate probes in different regions of the human genome, we used an TaqMan Gene Expression Assays by design which provides pre-designed primers-probe set for real-time PCR experiments (Applied Biosystems, <https://products.appliedbiosystems.com>) PCR was carried out using an ABI prism 7000 (Applied Biosystems) in a 96-well optical plate with a final reaction volume of 50 μ l. A total of 100 ng (10 μ l) was dispensed in each of the four sample wells for quadruplicate reactions. Thermal cycling conditions included a prerun of 2 min at 50°C and 10 min at 95°C. Cycle conditions were 40 cycles at 95°C for 15 sec and 60°C for 1 min according to the TaqMan Universal PCR Protocol (ABI). The TaqMan Universal PCR Master Mix and Microamp reaction tubes were supplied by Applied Biosystems. The starting copy number of the unknown samples was determined using the comparative Ct method as previously described. [95]

3.4 Multiplex Ligation-dependent Probe Amplification (MLPA)

MLPA analysis was performed according to the provider's protocol with a specifically designed set of probes for testing critical regions in DiGeorge syndrome (SALSA P023 kit; MRC-Holland, Amsterdam, Netherlands; <http://www.mrc-holland.com>), 1p-deletion syndrome, Williams syndrome, Smith-Magenis syndrome, Miller-Dieker syndrome, DiGeorge syndrome, Prader-Willi syndrome, Alagille syndrome, Saethre-Chotzen syndrome, Sotos syndrome: (SALSA P064B MR1 kit) and subtelomere regions (SALSA P036D subtelomeric primer kit). The ligation products were amplified by PCR using the common primer set with the 6-FAM label distributed by the supplier. Briefly, 100 ng of genomic DNA was diluted with TE buffer to 5 µl, denatured at 98°C for 5 minutes and hybridized with SALSA Probe-mix at 60°C overnight. Ligase-65 mix was then added and ligation was performed at 54°C for 15 minutes. The ligase was successively inactivated by heat, 98°C for 5 minutes. PCR reaction was performed in a 50 µl volume. Primers, dNTP and polymerase were added and amplification was carried out for 35 cycles (30 seconds at 95°C, 30 seconds at 60°C and 60 seconds at 72°C). Amplification products were identified and quantified by capillary electrophoresis on an ABI 310 genetic analyzer, using GENESCAN software (version 3.7) all from Applied Biosystems (Foster City, CA, USA). The peak areas of the PCR products were determined by GENOTYPER software (Applied Biosystems). A spreadsheet was developed in Microsoft™ Excel in order to process the sample data efficiently. Data were normalized by dividing each probe's peak area by the average peak area of the sample. This normalized peak pattern was divided by the average normalized peak pattern of all the samples in the same experiment. [96]

4. RESULTS

4.1 Novel deletions/duplications

**A 3 Mb Deletion in 14q12 Causes Severe Mental Retardation,
Mild Facial Dysmorphisms and Rett-like Features.**

Filomena Tiziana Papa, Maria Antonietta Mencarelli, Rossella Caselli, Eleni Katzaki, Katia Sampieri, Ilenia Meloni, Francesca Ariani, Ilenia Longo, Angela Maggio, Paolo Balestri, Salvatore Grosso, Maria Angela Farnetani, Rosario Berardi, Francesca Mari, and Alessandra Renieri.

Am. J Med. Genet Part A 146A:1994-1998 (2008)

)

Clinical Report
**A 3 Mb Deletion in 14q12 Causes Severe
 Mental Retardation, Mild Facial Dysmorphisms
 and Rett-like Features**

Filomena Tiziana Papa,¹ Maria Antonietta Mencarelli,¹ Rossella Caselli,¹ Eleni Katzaki,¹
 Katia Sampieri,¹ Ilaria Meloni,¹ Francesca Ariani,¹ Ilaria Longo,¹ Angela Maggio,²
 Paolo Balestri,² Salvatore Grosso,² Maria Angela Farnetani,² Rosario Berardi,²
 Francesca Mari,¹ and Alessandra Renieri^{1*}

¹Medical Genetics Molecular Biology Department, University of Siena, Siena, Italy

²Department of Pediatrics, University of Siena, Italy

Received 31 October 2007; Accepted 7 April 2008

The present report describes a 7-year-old girl with a de novo 3 Mb interstitial deletion of chromosome 14q12, identified by oligo array-CGH. The region is gene poor and contains only five genes two of them, *FOXG1B* and *PRKDI* being deleted also in a previously reported case with a very similar phenotype. Both patients present prominent metopic suture, epicanthic folds, bulbous nasal tip, tented upper lip, everted lower lip and large ears and a clinical course like Rett syndrome, including normal perinatal period, postnatal microcephaly, seizures, and severe mental retardation.

FOXG1B (forkhead box G1B) is a very intriguing candidate gene since it is known to promote neuronal progenitor proliferation and to suppress premature neurogenesis and its disruption is reported in a patient with postnatal microcephaly, corpus callosum agenesis, seizures, and severe mental retardation. © 2008 Wiley-Liss, Inc.

Key words: chromosome 14; array-CGH; 14q deletion; Rett-like features; *FOXG1B* gene

How to cite this article: Papa FT, Mencarelli MA, Caselli R, Katzaki E, Sampieri K, Meloni I, Ariani F, Longo I, Maggio A, Balestri P, Grosso S, Farnetani MA, Berardi R, Mari F, Renieri A. 2008. A 3 Mb deletion in 14q12 causes severe mental retardation, mild facial dysmorphisms and Rett-like features. *Am J Med Genet Part A* 146A:1994–1998.

INTRODUCTION

Interstitial deletions of the long arm of chromosome 14 are a relatively rare finding. Reported deletions have ranged from the loss of multiple bands to smaller deletions involving a single band [Kamnasaran et al., 2001; Petek et al., 2003]. The clinical phenotype varies, but some features commonly seen include psychomotor retardation, hypotonia, microcephaly and craniofacial anomalies.

Bisgaard et al. [2006] have recently described a female patient with an interstitial deletion of 14q12 of 2.65–3.5 Mb, identified by metaphase CGH. The girl presented psychomotor retardation, epilepsy, microcephaly and unusual facial features [Bisgaard et al., 2006].

Here we report on the clinical description and the molecular characterization of a patient with a very similar 3 Mb deletion of the long arm of chromosome 14 involving band q12, identified by oligo array-CGH. The clinical phenotype is represented by

severe mental retardation, epilepsy, microcephaly, Rett-like features and mild facial dysmorphisms.

To our knowledge this is the second case reported of a patient with a deletion smaller than 3.5 Mb involving the band 14q12. The small number of genes included in the deleted region (only five) may allow to identify candidate genes responsible for the clinical features.

CLINICAL REPORT

The patient is a 7-year-old girl, only child of healthy, nonconsanguineous parents. At the time of delivery, the mother and the father were 30 years old. No teratogen exposure during pregnancy was

*Correspondence to: Alessandra Renieri, M.D., Ph.D., Medical Genetics Molecular Biology Department, University of Siena, Viale Bracci, 2 53100 Siena, Italy. E-mail: renieri@unisi.it
 DOI 10.1002/ajmg.a.32413

reported. In the third pregnancy trimester routine echography showed intrauterine growth retardation. The girl was born at 39 weeks of gestation by cesarean. Her birth weight was 2730 g (25° cnt), length was 48.5 cm (25° cnt), OFC was 32 cm (10–25° cnt), and the Apgar score was 8 at 1' and 9 at 5'. Psychomotor development was severely impaired since early infancy: head control was acquired at the age of 4 months, no gain of motor milestones or social contact was achieved until now. An almost normal development during the first 6 months of life followed by a period of regression is referred by the mother. At that time, the child experienced an epileptic episode characterized by staring gaze followed by clonic jerking. At the age of 3 years and 10 months, seizures with similar electroclinical pattern re-occurred. EEG recording showed paroxysmal activity in the temporo-centro-occipital areas, more prominent in the temporo-central region bilaterally. Results of a brain MRI at age 1 year and a new recent evaluation were normal. An ophthalmologic examination showed hypermetropia.

Physical examination at 7 years showed: height 102 cm (<<3° cnt), weight 15.4 kg (<<3° cnt), OFC 45.5 cm (<<3° cnt), prominent metopic suture, bilateral epicanthal folds, intermittent exophoria, fullness of periorbital region, depressed nasal bridge, bulbous nasal tip, large ears, thick lips with everted lower lip, prognathism (Fig. 1). She showed generalized hypotonia, mild scoliosis, gastroesophageal reflux, bruxism, scialorrhoea and cold extremities. Constant stereotypic movements of hands and tongue were also evident. Hand movements are asymmetrical and more often not in the midline. They are rocking/rhythmic movements, hand waving, movements of scratching, pinching and patting with fingers. The patient presents also a double elbow crease on the left and a double ankle crease on the right.

Results of the karyotype (320 bands), analysis of telomeric rearrangements, sequence analysis of the *CDKL5* gene and analysis of 15q11q13 region were normal. DHPLC analysis of the *MECP2* gene showed a nonpathogenic variant (c. 1168_1173del) inherited from the father.

Array-CGH Analysis

Array based CGH analysis was performed as previously reported [Pescucci et al., 2007]. We used commercially available oligonucleotide microarrays containing about 99,000 60-mer probes (Human Genome CGH Microarray 105A Kit, Agilent Technologies, Santa Clara, California). The average resolution is about 25 Kb.

Quantitative PCR Analysis

Real-time quantitative PCR was performed to confirm array-CGH data. We used a pre-designed set of



FIG. 1. Images of the patient at age 13 months (a), 2 years and 7 months (b) and 7 years (c and d). Note the bulbous nasal tip, the thick lips with everted lower lip and protruding tongue and the prognathism. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

primers and probe specific for the *FOXG1B* gene (14q12 locus) provided by the Assay-by-Design service (Applied Biosystems, Foster City, CA). Primers and probe were designed in the gene: *FOXG1B* forward primer: 5'-GTCAGCGAGGTGCAATGTG-3'; *FOXG1B* reverse primer: 5'-GGCAAAGCAGTCAC-TATTTAGATACACA-3'; *FOXG1B* TaqMan probe: 5'-TTGACTTCCAAACCTTATATTCT-3' reverse. PCR was carried out as previously described [Pescucci et al., 2007]. The starting copy number of the unknown samples was determined using the comparative Ct method, as reported by Livak [1997].

Array-CGH analysis identified a chromosome 14 interstitial deletion of about 3 Mb [46,XX,del(14)(q12q12)] (Fig. 2). Both breakpoints are mapped in 14q12, with the last present oligonucleotide located in 27.47 Mb and the first deleted in 27.55 Mb. The distal breakpoint is located between 30.45 Mb, last oligonucleotide deleted, and 30.47 Mb, first oligonucleotide present (according to UCSC Genome Browser, <http://genome.ucsc.edu>, on Human May 2004 Assembly). Only five genes are included in the region: *FOXG1B*, *PRKD1*, *SCFD1*, *COCH* and *STRN3*.

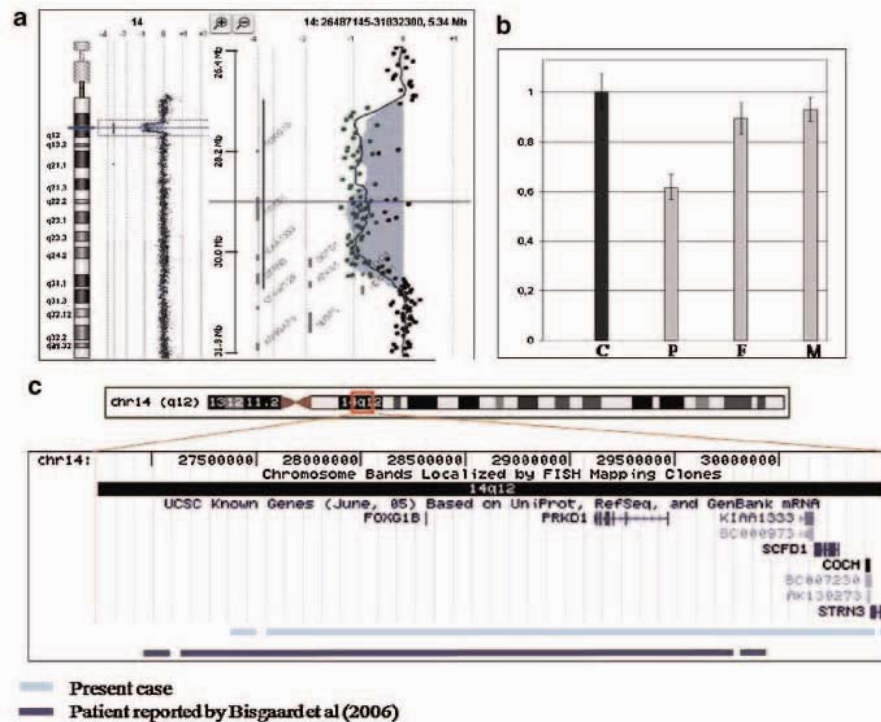


FIG. 2. Molecular data of the patient. (a) Array-CGH ratio profile. Chromosome 14 array-CGH ratio profile using DNA from the patient and a reference DNA from a normal female. On the left, the chromosome 14 ideogram. On the right, the log₂ ratio of the chromosome 14 probes plotted as a function of chromosomal position. Oligos with a value of zero represent equal fluorescence intensity ratio between sample and reference. Each dot represents a single probe (oligo) spotted on the array. Copy number loss shifts the ratio to the left (value of about $-2 \times$). (b) Real-time quantitative PCR validation experiment. *FOXG1B* ddCT ratios and standard deviations obtained for the patient (P), her parents (F, father, M, mother) and a control sample (C). The patient shows a ddCT ratio of about 0.6, indicating the presence of a single copy of *FOXG1B* (deletion), while the parents and the controls show ddCT ratios of about 1.0, indicating a double copy of the gene (normal). (c) Gene content of the deleted region (UCSC Genome Browser, <http://genome.ucsc.edu>). In the bottom part of the figure the deletion of our patient is compared with the deletion of the patient reported by Bisgaard. The deletion extension of this latter case has been defined by 105 K oligo array-CGH, performed in our laboratory. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Array-CGH analysis performed in both healthy parents showed a normal result indicating that the rearrangement is a *de novo* deletion (results not shown). To confirm array data, real-time quantitative PCR experiments were performed in the patient and her parents. The deletion was confirmed in the patient, while parents showed a normal result (Fig. 2b).

Array-CGH analysis was performed on DNA from patient described by Bisgaard et al. [2006], for a better definition of deletion extension and breakpoints. The deletion is about 3.12 Mb; with both breakpoints mapped in 14q12. The last present oligonucleotide is located in 26.90 Mb and the first deleted in 26.98 Mb. The distal breakpoint is located between 30.03 Mb, last oligonucleotide deleted, and 30.09 Mb, first oligonucleotide present (according to UCSC Genome Browser, <http://genome.ucsc.edu>, on Human May 2004 Assembly). Only two genes are included in the region: *FOXG1B* and *PRKD1*.

DISCUSSION

Given the clinical history and the neurological features of the patient, we first hypothesized a Rett spectrum disorder. The severe neurological impairment, dyspraxia, stereotypies, gastroesophageal reflux, constipation, scoliosis and cold extremities together with a normal perinatal period were quite suggestive of Rett syndrome. However, we were unable to demonstrate a mutation in one of the two main genes of Rett syndrome, *MECP2* and *CDKL5*.

Considering the presence of distinctive facial features, which are not present in Rett syndrome, we decided to perform the array-CGH analysis in the patient, and we found a pathogenic *de novo* deletion in 14q12. In this region no copy number variation is reported in normal individuals and it contains a very low amount of genes. In a 3 Mb region only five known genes are included. None of them is reported

in the databases of imprinted genes (<http://www.genem印rint.com> and <http://igc.otago.ac.nz>).

Although LCRs are not reported near the breakpoints (BioXRTdatabase: <http://projects.tcag.ca/bioxrt/>), the deletion identified in our patient is very similar to another recently reported by Bisgaard et al. [2006] (Table I). The described facial phenotype is very similar between the two patients: both present postnatal microcephaly, prominent metopic suture, epicanthic folds, bulbous nasal tip, tented upper lip, everted lower lip and large ears. Moreover, the two patients both have mental retardation and seizures. Two genes are included in the common deleted region: *FOXG1B* and *PRKD1*.

The *PRKD1* gene encodes for a cytosolic serine-threonine kinase that binds to the trans-Golgi network and regulates the fission of transport carriers specifically destined to the cell surface [Johannes et al., 1994]. To date a clinical phenotype has not been associated with mutations or deletions of this gene.

FOXG1B (forkhead box G1B) gene is a very interesting gene. This transcription factor is potentially a strong candidate gene for determining fore-brain size in vertebrates due to its role in the development of the telencephalon, where it promotes progenitor proliferation and suppresses premature neurogenesis [Bredenkamp et al., 2007].

Shoichet et al. [2005] reported a patient with an inversion in chromosome 14 disrupting at the breakpoint the *FOXG1B* gene. The patient exhibited severe cognitive disability associated with postnatal microcephaly, seizures, agenesis of the corpus callosum and frontal and parietal myelination defects. The patient was very severely affected being in the status of tetraplegia at the age of 7 years. Similar features are described also in two patients harboring larger deletions that involve *FOXG1B*, previously reported by Kannasaran et al. [2001]. It is plausible that neurological features result at least in part from *FOXG1B* haploinsufficiency.

On the telomeric end of the deletion, outside the deleted region shared with the patient reported by Bisgaard there are three additional genes: *SCFD1*, *COCH* and *STRN3*. *SCFD1* (Sec1 family domain containing 1) may be involved in vesicular transport between the endoplasmic reticulum and the Golgi and may interact functionally with Ras2 protein. To date a clinical phenotype has not been associated with mutations or deletions of this gene. *COCH* encodes for cochlin, the major non-collagen component of extracellular matrices of the inner ear. Dominant negative *COCH* mutations cause dominant progressive hearing loss and vestibular dysfunction (DFNA9) [Makishima et al.,

TABLE I. Comparison of the Clinical Features of the Present Case with the one Described by Bisgaard et al. [2006]

	Present case	Patient reported by Bisgaard et al. [2006]
Deletion breakpoints ^a	27.47 Mb 27.55 Mb 30.45 Mb 30.47 Mb	26.90 Mb 26.98 Mb 30.03 Mb 30.09 Mb
Age	7 years	11 months
Sex	F	F
Birth weight	2730 g (25 ^o cnt)	3200 g (<50 ^o cnt)
Birth length	48.5 cm (25 ^o cnt)	53 cm (<90 ^o cnt)
Birth OFC	32 cm (10–25 ^o cnt)	33 cm (25 ^o cnt)
Postnatal growth retardation	Present	N.A.
Psychomotor retardation/mental retardation	Severe	Present (Severity unknown)
Seizures	Present	Present
Brain MRI	Normal	Normal
Microcephaly	Yes (OFC 45.5 cm << 3 ^o cnt)	Yes
Prominent metopic suture	Present	Present
Ears	Large	Apparently large with prominent anhelices
Epicanthal folds	Bilateral	Bilateral
Nasal bridge	Depressed	Depressed
Nasal tip	Bulbous	Bulbous
Lips	Thick with tented upper lip and everted lower lip	Tented upper lip, everted lower lip
Prognathism	Present	N.A.
Hypotonia	Generalized	N.A.
Scoliosis	Mild	N.A.
Gastroesophageal reflux	Present	N.A.
Bruxism	Present	N.A.
Scialorrhea	Present	N.A.
Cold extremities	Present	N.A.
Postnatal regression	After 6 months	N.A.
Stereotypic movements	Constant of hands and tongue	N.A.

N.A.: not available.

^aOligonucleotide positions according to UCSC Genome Browser, <http://genome.ucsc.edu>, on Human May 2004 Assembly; numbers refers to last oligonucleotide present, first oligonucleotide deleted, last oligonucleotide deleted and first oligonucleotide present, respectively.

2005]. Our patient presents a normal auditory function and this is in accordance with the evidence that *COCH* haploinsufficiency does not cause DFNA9 [Makishima et al., 2005].

The distal breakpoint of the deletion interrupts the *STRN3* gene. *STRN3* or *SG2NA* is a supposedly nuclear protein, which shares with striatin identical protein-protein interaction domains and the same overall domain structure [Castets et al., 2000]. Proteins of the striatin family are principally expressed in neurons with a somato-dendritic localization and high concentration in dendritic spines. Recent investigations have demonstrated the involvement of proteins of the striatin family in numerous developmental and nondevelopmental signaling pathways as well as in vesicular trafficking [Gaillard et al., 2006]. This data could suggest a possible involvement also of *STRN3* gene in the neurological phenotype of our patient.

ACKNOWLEDGMENTS

This work was supported by Telethon grant GGP06170 and EuroRETT E-RARE Network to A.R. and grant from University of Siena (PAR 2006) to F.M. We gratefully acknowledge Anne Marie Bisgaard, Maria Kirchhoff and Monika Cohen for their collaboration in providing the DNA sample of their patient that has permitted a better definition of the deletion breakpoints.

REFERENCES

- Bisgaard AM, Kirchhoff M, Tumer Z, Jepsen B, Brondum-Nielsen K, Cohen M, Hamborg-Petersen B, Bryndorf T, Tommerup N, Skovby F. 2006. Additional chromosomal abnormalities in patients with a previously detected abnormal karyotype, mental retardation, and dysmorphic features. *Am J Med Genet Part A* 140A:2180–2187.
- Bredenkamp N, Seoighe C, Illing N. 2007. Comparative evolutionary analysis of the FoxG1 transcription factor from diverse vertebrates identifies conserved recognition sites for micro-RNA regulation. *Dev Genes Evol* 217:227–233.
- Castets F, Rakitina T, Gaillard S, Moqrich A, Mattei MG, Monneron A. 2000. Zinedin, *SG2NA*, and striatin are calmodulin-binding, WD repeat proteins principally expressed in the brain. *J Biol Chem* 275:19970–19977.
- Gaillard S, Bailly Y, Benoist M, Rakitina T, Kessler JP, Fronzaroli-Molinieres L, Dargent B, Castets F. 2006. Targeting of proteins of the striatin family to dendritic spines: role of the coiled-coil domain. *Traffic* 7:74–84.
- Johannes FJ, Prestle J, Eis S, Oberhagemann P, Pfizenmaier K. 1994. PKCu is a novel, atypical member of the protein kinase C family. *J Biol Chem* 269:6140–6148.
- Kannasaran D, O'Brien PC, Schuffenhauer S, Quarrell O, Lupski JR, Grammatico P, Ferguson-Smith MA, Cox DW. 2001. Defining the breakpoints of proximal chromosome 14q rearrangements in nine patients using flow-sorted chromosomes. *Am J Med Genet* 102:173–182.
- Livak K. 1997. ABI Prism 7700 Sequence Detection System, No. 2.
- Makishima T, Rodriguez CI, Robertson NG, Morton CC, Stewart CL, Griffith AJ. 2005. Targeted disruption of mouse *Coch* provides functional evidence that DFNA9 hearing loss is not a *COCH* haploinsufficiency disorder. *Hum Genet* 118:29–34.
- Pescucci C, Caselli R, Grosso S, Mencarelli M, Mari F, Farnetani M, Piccini B, Artuso R, Bruttini M, Priolo M, Zuffardi O, Gimelli S, Balestri P, Renieri A. 2007. 2q24-q31 deletion: Report of a case a review of the literature. *Eur J Med Genet* 50:21–32.
- Petek E, Plecko-Starting B, Windpassinger C, Egger H, Wagner K, Kroisel PM. 2003. Molecular characterisation of a 3.5 Mb interstitial 14q deletion in a child with several phenotypic anomalies. *J Med Genet* 40(4):e47.
- Shoichet SA, Kunde SA, Viertel P, Schell-Apacik C, von Voss H, Tommerup N, Ropers HH, Kalscheuer VM. 2005. Haploinsufficiency of novel *FOXP1B* variants in a patient with severe mental retardation, brain malformations and microcephaly. *Hum Genet* 117:536–544.

4.2 CNV in Autism and RETT syndrome

4.2.1 Genetic susceptibility factors in ASD patients

Practical hints for management of array-CGH output in autism spectrum disorders: susceptibility regions versus uncertain cofactors.

Filomena Tiziana Papa, Roberto Canitano, Maria Antonietta Mencarelli, Mucciolo Mafalda, Disciglio Vittoria, Annabella Marozza, Radice Lucia, Cinzia Castagnini, Laura Dosa, Marzia Pollazzon, Joussef Hayek, Alessandra Renieri, Francesca Mari.

Submitted to Am J Med Genet Part B

Practical hints for management of array-CGH output in autism spectrum disorders: susceptibility regions versus uncertain cofactors.

Filomena T. Papa¹, Roberto Canitano², Maria Antonietta Mencarelli¹, Mafalda Mucciolo¹, Vittoria Disciglio¹, Annabella Marozza¹, Lucia Radice², Cinzia Castagnini¹, Laura Dosa¹, Marzia Pollazzon¹, Joussef Hayek², Alessandra Renieri¹, Francesca Mari¹.

¹ Medical Genetics-Molecular Biology Department-University of Siena-Policlinico S. Maria alle Scotte-53100 Siena-Italy

² Child Neuropsychiatry-University Hospital of Siena-53100 Siena-Italy

BACKGROUND: Comparative genomic hybridization (CGH) has provided insights into the underlying genetic causes of autism spectrum disorders (ASD) by associating CNVs at 16p11.2, 16p13.1 and 15q13.3 with susceptibility to ASD in up to 1% of patients.

METHODS: Using 44K oligo array-CGH we analyzed 95 unrelated subjects with ASD, thoroughly investigated by both child neuropsychiatrists and clinical geneticists. In most cases parents were available for the analysis. All patients except 3 with Asperger syndrome, had mental retardation, 24% of them had congenital anomalies and 13% had epilepsy.

RESULTS: In fifteen patients (16%) CNV(s) were considered disease associated. These include two *de novo* deletions (del16p11.2; delXq12), five known susceptibility regions and eight new potential candidate regions (dup2q11.2, del20p12.1, del7q31.1, dup2p16.3, dup4p16.1, dup5q21.1, del9p13.1p13.2,

del17q25.3). In 20 patients (21%), including three patients with two rearrangements each, one inherited by the mother and one by the father, CNVs were classified "as uncertain cofactors". Overall, CNVs were more common among patients with congenital anomalies and less common in patients with epilepsy.

CONCLUSIONS: Our study confirmed that array-CGH analysis is able to detect the underlying genetic cause in about 16% of ASD patients, strongly indicating that it has become an essential diagnostic tool for ASD patient assessment. Moreover, this study underlines the difficulty in interpreting array-CGH data in an even larger number of patients (21%), and provides some practical recommendations for the management of such data and for the communication of findings to the family.

Keywords: array-CGH; CNVs; autism spectrum disorders; susceptibility region; uncertain cofactors

INTRODUCTION

Autism Spectrum Disorders (ASDs) present complex and heterogeneous aetiology with a strong evidence of genetic involvement. In addition to Autism, ASDs include Asperger syndrome and pervasive

developmental disorder not otherwise specified (PDD-NOS). The overall estimated prevalence of ASD is 1:160, while the prevalence of Autism is considered 1:500 with a 4:1 male to female ratio.[1]

The complexity of the disorder, resulting from the interaction of several genes and environmental factors, makes the identification of contributory genes extremely difficult. Recently, comparative genomic hybridization (CGH) technology has been used in research studies and in clinical assessment of ASDs in order to detect copy number variants (CNVs) throughout the whole genome.

The detection rate in different studies is variable from 9% to 44% and it is highly influenced by the resolution of the applied array platform, the clinical selection of patients (family history, associated anomalies) and the criteria used to define the CNVs as pathogenic.

One of the first applications of array-CGH analysis in ASDs was published in 2006. Array-CGH with 1 Mb resolution allowed the identification of 8 clinically relevant CNVs among a cohort of 29 patients with syndromic ASD (~28%).[2] Subsequently, a larger series of 427 unrelated ASD cases has been investigated by single nucleotide polymorphism (SNP) microarray. This analysis identified 277 CNVs in 44% of the patients. These were classified as pathogenic in 14% of cases (~7% *de novo* and ~7% overlapping/recurrent in 2 or more unrelated samples).[3] In the same year, Christian and colleagues, using BAC array, studied nearly 400 patients and showed an overall CNVs rate of 11.6%. [4] More recently, Qiao Y *et al* investigated 100 subjects with syndromic ASD using BAC array-CGH and identified 9 pathogenic CNVs (9%) and three rearrangements of unknown significance.[5] In the latter study, the pathogenic role of CNVs was determined on the basis of these criteria: CNVs not reported in the database of genomic variants or in the control population; X-linked CNVs present in males and maternally inherited; and CNVs already reported in ASD.

Overall, array-CGH analysis increased the diagnostic yield in ASDs and allowed the identification of new genetic

determinants. In addition to the well known recurrent rearrangements involving the 15q11q13, 16p13 and 22q13 regions, recurrent microdeletions and microduplications at 16p11.2 have been recently identified and are shown to confer susceptibility to ASDs in up to 1% of autistic patients.[3]

Beside these emerging data, the clinical interpretation of several newly discovered private CNVs is still uncertain. In the above reported literature, only a fraction of CNVs ranging from 9% to 27% have been classified as pathogenic. Our analysis focused on the identification of new pathogenic loci and on the practical management of those CNVs with uncertain clinical significance.

METHODS

Study cohort.

A cohort of 95 unrelated patients with ASD with unknown aetiology was selected for this study. Eighty were males and 15 were females (M:F; 5.3:1). The age range was 2 years and 8 months and 22 years and 2 months with a mean of 10 years and 10 months, and median of 10 years and 4 months. This cohort included 3 patients with Asperger syndrome, 48 with Autism and 44 with PDD-NOS. All patients belong to "Cell Lines and DNA Bank of Rett syndrome, X linked mental retardation and other genetic diseases" - Telethon Genetic Biobank Network, and were: Asperger = AU237, MR1297, MR1346; Autism = MR841, MR593, MR755, AU277, MR1081, AU243, MR886, MR1187, AU212, MR681, MR925, PW66, MR1203, AU204, AU303, MR1460, AU274, MR322, AU180, AU295, AU193, MR1475, MR868, MR815, MR940, MR1505, MR798, AU265, AU197, AU229, AU237, MR1181, MR746, MR460, MR1297, MR850, AU186, AU271, AU256, MR1269, MR708, RET196, MR1681, MR1346, MR1537; PDD-NOS = AU135, AU94, AU142, MR790, AU227, MR518, MR148, MR1748, AU183t, XXM1, MR521, MR1005, MR889, MR1708, MR1022,

MR326, MR1436 AS100, AU232, MR1601, MR1392, MR1108, MR1075, AS88, AU224, MR1647, MR1408, MR853, MR362, MR384, AU209, AU289, AU290, MR901, AU293, AU221, MR714, MR1204, AU157, MR1477, MR1645, MR1643, MR1191, MR1635.

The ASD diagnosis for all subjects was made by the same child neuropsychiatrist (RC) based on standardized Diagnostic and Statistical Manual 4th Edition (DSM-IV) criteria using Autism Diagnostic Interview-Revised (ADI-R) and/or Autism Diagnostic Observation Schedule Generic (ADOS-G) standards. All patients underwent a comprehensive evaluation by a clinical geneticist (FM or MAM) who excluded a diagnosis of a recognizable syndrome, and all patients, except those with microcephaly, tested negative for *FMRI* gene expansion.

Array-CGH analysis: Genomic DNA of the patients was isolated from an EDTA peripheral blood sample using the QIAamp DNA Blood Kit according to the manufacturer's protocol (Qiagen, www.qiagen.com). Genomic DNA of normal male and female controls was obtained from Promega. Ten micrograms of genomic DNA from the patient (test sample) and the control (reference sample) were sonicated. Test and reference DNA samples were subsequently purified using dedicated columns (DNA Clean and Concentrator, Zymo Research) and the appropriate DNA concentrations were determined by a DyNA Quant™ 200 Fluorometer (GE Healthcare).

Array-CGH analysis was performed using commercially available oligonucleotide microarrays containing about 43,000 60-mer probes (Human Genome CGH Microarray 44B Kit, Agilent Technologies, Santa Clara, California) as previously reported.[6] The

median functional resolution was 75-100 Kb. Copy number variations (CNVs) were considered significant if they were defined by three or more oligonucleotides, and were not present in the Database of Genomic Variants. Confirmation of results was performed by a second independent experiment. Segregation analysis of each rearrangement was performed in 22/35 (63%) using the same technique.

Statistical analysis. Differences in proportions were calculated using Fisher Exact Test with $p < 0.05$ set for significance.

RESULTS

CNV discovery.

Among the 95 patients, 35 patients exhibited at least one rearrangement (Tables 1 and 2). In these 35 patients, 49 specific rearrangements were identified: 22 deletions and 27 duplications, 45 in autosomes and 4 in the X-chromosome. The size of rearrangements ranged from 2Kb to 1.8Mb. The mean size of rearrangements was 326 Kb. Among the 49 rearrangements two were *de novo*, 13 were inherited by the mother and 13 by the father, while for 21 inheritance was unknown. Twenty-five patients had only one rearrangement and 10 had two or more rearrangements up to four in one case. Interestingly, three patients had two rearrangements, one inherited by the father and the second by the mother, suggesting digenic inheritance.

In fifteen patients (16%) CNV(s) were considered disease associated. These include two *de novo* deletions (del16p11.2; delXq12), five known susceptibility regions and eight new potential candidate regions (dup2q11.2, del20p12.1, del7q31.1, dup2p16.3, dup4p16.1, dup5q21.1, del9p13.1p13.2, del17q25.3). In 20 patients (21%) CNVs were classified "as uncertain cofactors".

Table 1 Susceptibility regions (Class I)

Sample ID	Gender	Chromosome	Rearrangement	Inheritance	Proximal breakpoint	Distal breakpoint	Size	Comments
Proved <i>de novo</i> (Class Ia)								
MR518	M	16p11.2	Loss	De novo	29 581 455	30 027 213	450 Kb	Susceptibility to ASD.[3]
AU142	M	Xq12	Loss	De novo	65 732 215	65 811 681	80 Kb	One gene: <i>EDA2R</i> , a susceptibility gene for androgenic alopecia.[7] The distal breakpoint is located 1.36 Mb from the <i>XLID</i> gene, <i>OPHN1</i> .
Known susceptibility regions (Class Ib)								
MR1537	M	16p13.11	Gain	maternal	14 852 061	15 867 585	1.1 Mb	Susceptibility to ASD.[8]
AU221	M	15q13.3	Gain	unknown	29 809 025	30 298 096	500 Kb	Susceptibility to ASD.[9]
XXM1	M	17q12	Gain	paternal	31 925 650	33 726 698	1.8 Mb	Susceptibility to cognitive impairment and behavioral abnormalities.[10]
MR1475 (*)	M	11p12	Loss	maternal	40 239 489	41 342 038	1.1 Mb	Recurrent and overlapping locus in ASD.[3]
MR746	M	15q11.2	Loss	unknown	20 335 887	20 636 537	300 Kb	Susceptibility to developmental delay, behavioral problems and dysmorphisms.[11]
New susceptibility regions (Class Ic)								
MR1346	M	2q11.2	Gain	maternal	99 991 724	100 088 972	97 Kb	One gene: <i>AFF3</i> associated with limb and brain abnormalities when deleted.[12]
MR681	M	20p12.1	Loss	maternal	14 772 372	15 216 002	444 Kb	One gene: <i>C20orf133</i> , disrupted in a patient with Niikawa-Kuroki syndrome.[13]
MR460 (*)	M	2p16.3	Gain	paternal	47 866 371	47 938 504	72 Kb	Two genes including <i>MSH6</i> , mutated in Lynch syndrome. The proximal breakpoint is located 3Mb far from the <i>neurexin 1</i> gene which has been associated with schizophrenia.[14]
MR1392 (*)	M	7q31.1	Loss	unknown	110 917 835	110 989 204	70 Kb	One gene: <i>IMMP2L</i> , a candidate gene for ASD.[15]
AU224	M	4p16.1	Gain	paternal	9 686 502	9 751 028	65 Kb	One gene: <i>WRDI</i> , upregulated in the dorsolateral prefrontal cortex in schizophrenia patients.[16]
AU197	M	5q21.1	Gain	paternal	100 219 716	100 401 892	182 Kb	Only one gene: <i>ST8SIA4</i> , involved in polysialylation of the neural cell adhesion molecule, a critical step in neural development.[17]
AU183	M	9p13.1p13.2	Loss	unknown	37 847 269	38 040 719	193 Kb	Three genes, including <i>SHB</i> , playing a role in neuronal cells[18]; and <i>WDR32 (DCAF10)</i> interacting with the <i>XLID</i> gene, <i>CUL4B</i> . [19]
AU271 (*)	M	17q25.3	Loss	unknown	78 536 419	78 653 660	117 Kb	Two genes, including <i>METRNL</i> , which codes a glial cell differentiation regulator-like protein. [20]

(*) These patients have additional CNV of unknown significance, see table 2.

Table 2 Uncertain cofactors (Class II)

Sample ID	Gender	Chromosome	Rearrangement	Inheritance	Proximal breakpoint	Distal breakpoint	Size	Gene content
Likely X-linked recessive inheritance (Class IIa)								
PW66	M	Xp22.31	Gain	maternal	7 565 292	8 226 181	660 Kb	This region includes 3 genes (VCX, PNPLA4 and VCX2) and is located in distal region involved in Xp22.31 microduplication.
MR940	M	Xq22.3	Gain	maternal	106 238 368	106 515 657	277 Kb	3 genes
MR1681	M	Xq26.2	Gain	unknown	130 398 834	130 788 239	390 Kb	2 genes
Likely digenic inheritance (Class IIb)								
AS100	M	1p22.1	Loss	maternal	87 179 292	87 273 767	94 Kb	No genes
		4q35.2	Loss	paternal	189 846 235	190 706 272	860 Kb	No genes
MR815	M	4q32.2	Gain	paternal	162 406 396	162 900 008	493 Kb	1 gene
		12p11.23	Gain	maternal	27 270 178	27 659 659	390 Kb	4 genes
RET196	F	9p21.3	Loss	maternal	21 431 614	21 537 336	105 Kb	2 genes
		1q23.1	Gain	paternal	111 421 787	111 470 536	50 Kb	1 gene
Likely cofactors for additional signs (Class IIc)								
MR1475 (*)	M	3q29	Gain	maternal	199 058 690	199 288 161	230 Kb	5 genes including RPL35A mutated in heterozygous state in Diamond-Blackfan Anemia.
Unknown significance (Class II d)								
MR853	M	1q43	Gain	paternal	234 698 142	234 814 728	116 Kb	3 genes including EDARADD gene associated with anhidrotic ectodermal dysplasia when mutated both in homozygous and heterozygous state.
MR1505	F	1p21.2	Loss	unknown	101 246 819	101 276 111	29 Kb	1 gene
MR1748	F	2p15	Gain	unknown	63 459 048	63 670 473	210 Kb	2 genes
		4p13	Gain	unknown	44 545 342	45 213 670	668 Kb	No genes
MR886	M	4p16.3p16.2	Gain	maternal	2 902 096	3 114 307	212 Kb	5 genes including HTT gene responsible for Huntington disease in heterozygous state (gain of function).
AU293	M	5p13.3	Loss	unknown	37 552 360	38 292 964	740 Kb	2 genes including GDNF gene associated with susceptibility to Hirschsprung disease.
AU303	M	8p23.1	Loss	unknown	9 933 149	10 030 338	97 Kb	1 gene
MR1022	M	9p24.1	Loss	maternal	6 318 511	6 321 081	2 Kb	1 gene
AU193	F	10q24.1	Gain	paternal	99 369 370	99 498 368	130 Kb	6 genes including ZFYVE27 gene responsible for spastic paraplegia when mutated in heterozygous state.
AU235	M	10q21.1	Gain	unknown	59 944 705	60 678 299	733 Kb	4 genes
MR384	M	14q21.1	Loss	paternal	40 088 478	40 380 622	292 Kb	No genes
MR798	M	15q15	Loss	paternal	41 676 219	41 830 276	154 Kb	6 genes including STRC gene associated with deafness when mutated in homozygous state.
MR1501	M	17q12	Loss	unknown	30 711 469	30 762 521	51 Kb	2 genes
		21q22.11	Gain	unknown	32 572 536	32 869 001	296 Kb	6 genes including MRAP gene associated with glucocorticoid deficiency when mutated in homozygous state.
		10q26.3	Gain	unknown	135 202 362	135 222 424	20 Kb	2 genes
AU277	M	21q21.3	Loss	paternal	28 834 275	28 948 424	114 Kb	1 gene
MR460 (*)	M	7q34	Loss	paternal	142 469 982	142 591 458	121 Kb	3 genes
MR1392 (*)	M	2q21.1	Gain	unknown	130 731 967	130 861 076	129 Kb	3 genes
		10q26.3	Loss	unknown	135 104 029	135 222 424	118 Kb	3 genes
		3q13.2	Loss	unknown	113 666 633	113 912 199	245 Kb	4 genes
AU271 (*)	M	17q21.31	Loss	unknown	41 544 224	41 706 870	163 Kb	Only one gene KIAA1267. This gene extends into commonly variable region around 41.6Mb and in part its sequence overlaps the distal region involved in 17q21.31 microdeletion syndrome. Note: no omim gene.
PW66 (**)		17p13.3	Gain	maternal	955 992	1 131 225	175 Kb	2 genes
		19q13.32	Gain	unknown	50 043 547	50 104 453	60 Kb	2 genes including APOE associated with hyperlipoproteinemia, Alzheimer disease-2, lipoprotein glomerulopathy, sea-blue histiocyte disease. Moreover APOE is associated with susceptibility to myocardial infarction.

(*) These patients have one additional pathogenic CNV, see table 1.

(**) These patients have one additional pathogenic CNV, see table 2 – Class IIa.

CNVs affecting susceptibility regions

Six patients showed a rearrangement in known susceptibility regions for ASD: 11p12, 15q11.2, 15q13.3, 16p11.2, 16p13.11, and 17q12 (Table 1).[3] [11] [9] [8] [10] The 16p11.2 rearrangement was proven to be *de novo*, while the remaining rearrangements were inherited or with unknown origin. The case with 17q12 CNV was previously reported in detail.[21]

While the rearrangements in 16p11.2, 16p13.1 and 17q12 fall within the range of a previously reported susceptibility region [3] [8] [10], the rearrangements in 15q13.3, 15q11.2 and 11p12 call for specific description. [3] [22] [23] [24] [25] [8] [26] [27] [11]

Szafranski et al. recently investigated the microduplication in 15q13.3 of about 500kb that includes the *CHRNA7* gene and overlaps the class 3 microduplication. They demonstrated that the mutation may exert a predisposition to neurodevelopmental and neuropsychiatric phenotypes in association with other genetic modifiers.[9] In our case, the *de novo* or inherited origin of the rearrangement could not be assessed: the *CHRNA7* duplication was excluded in mother's DNA, while father was not available for the analysis. The family history was intensely investigated for the presence of cognitive impairment or neuropsychiatric disorder, which were both ruled out.

The microdeletion 15q11.2 was recently described by Doornbos *et al*, who suggested a possible predisposing effect to developmental delay, behavioural problems and dysmorphic features, such as hypertelorism, palate and ear anomalies.[11] The rearrangement, partially overlapping PWS/AS critical region, is of about 300 kb and it includes four genes. Three of them seem to be important for neurological development. The *NIPAI* gene is highly expressed in neuronal tissue and encodes a magnesium membrane transporter. Gain of function

mutations in this gene have been associated with autosomal dominant spastic paraplegia 6.[28] The *CYFIP1* gene is a protein that interacts with FMRP (fragile X mental retardation protein) and GTPase RAC1, a protein involved in the development and maintenance of neuronal structures.[29] A third gene, *TUBGCP5*, which is highly expressed in the subthalamic nuclei of the brain, could be involved in the behavioral disorders, such as ADHD and obsessive-compulsive behavior.[30]

Another patient had a deletion in 11p12, a recurrent rearrangement in ASD patients in the series described by Marshall *et al*. [3] The deletion in our patient is of about 1 Mb and is larger than the deletion reported by Marshall *et al*. It disrupts only one gene, *NGL1*, which is highly expressed in the striatum and in cerebral cortex, and is a specific binding partner for *NTNG1*, a member of the netrin family of axon guidance molecules. Functional experiments using mouse thalamic neurons and chick embryos suggested that the interaction of the two genes is crucial for the growth of thalamocortical axons.[31]

Identification of novel susceptibility regions

Nine new susceptibility regions were identified on the basis of either the gene content or *de novo* inheritance (Table 1). One case had a *de novo* deletion of about 100 Kb in **Xq12**, including a single gene: *EDA2R*. This gene, also known as *XEDAR*, is the X-linked ectodysplasin receptor, a type III transmembrane protein with an extracellular N terminus and a cellular cytoplasmic C terminus.[32]

The remaining eight CNVs either included genes that, according to their expression and function or according to literature data, were good candidates for the neurological phenotype or fall at 1-3Mb distance from such genes (Table 1). The **7q31.1** deletion includes the *IMMP2L* gene, a candidate gene for ASD by association studies.[15] The **9p13.1p13.2**

deletion contains the *SHB* gene, which plays a role in neuronal cells, and the *WDR32* gene (also known as DCAF10), which interacts with *CUL4B*, an X-linked mental retardation gene.[19] In one male patient an inherited **20p12.1** deletion of nearly 450 Kb interrupting the *C20orf133* has been identified. In 2007 a *de novo* deletion of the same gene was identified in a patient with a clinical diagnosis of Niikawa-Kuroki syndrome.[13] Our patient does not show clinical features of this syndrome and the analysis of different cohorts of patients with Niikawa-Kuroki syndrome did not reveal any *C20orf133* mutation. Nevertheless, this gene remains a good candidate gene for the neurocognitive phenotype.

The duplication in **2p16.3** included two genes: *MSH6* and *FBXO11* (also known as *VITI*). *MSH6* is a component of the post-replicative DNA mismatch repair system and defects in *MSH6* are the cause of Lynch syndrome.[14] *FBXO11* is known to be down-regulated in melanocytes from patients with vitiligo, a skin disorder that results in depigmentation, while polymorphisms in this gene are associated with chronic otitis media with effusion and recurrent otitis media.[33] The duplication of *FBXO11* may be responsible for the presence in the patient of hyperchromic areas localized in the abdomen.[34] About 3Mb from the proximal breakpoint is the neurexin 1 gene (*NRXN1*), which has been implicated in a variety of conditions including autism, schizophrenia, and nicotine dependence.[14]

The small duplication on **4p16.1** included only the *WDR1* gene. This gene is included in a 200 kb region linked by haplotype analysis to bipolar affective disorder. *WDR1* encodes a brain expressed protein that provides protein-protein interactions thought to induce the disassembly of actin filaments. This function offers a plausible mechanism for neuronal dysfunction.[16] The **5q21.1** duplication included the *ST8SIA4* gene that promotes neurite outgrowth and sprouting

and has a role in development and neuronal plasticity.[17] The deletion in **17q25.3** included only two genes: *B3GNTL1* and *METRNL*. *B3GNTL1* codes for a glycosyltransferase, a subgroup of enzymes involved in a wide range of functions in all living organisms. *METRNL* is a small gene, with a guanine cytosine (GC)-rich 5' untranslated region resembling the features of genes involved in epigenetic control. The gene product is a protein akin to meteorin, a secreted protein expressed in undifferentiated neural progenitors and in the astroglial lineage. In the embryonal stages, meteorin is involved in the development of the nervous system, contributing to glial cell differentiation and axonal network formation.[20] The **2q11.2** duplication includes a single gene, *AFF3* (also known as *LAF4*), which is interrupted. This gene, which belongs to a family of four genes that includes *FMR2*, encodes a nuclear transcriptional activator that is preferentially expressed in lymphoid tissue, and that may function in lymphoid development and oncogenesis. Haploinsufficiency for *AFF3* has been recently associated with brain, limb and urogenital anomalies.[12]

Uncertain cofactors

A total of 34 CNVs of unknown significance were identified (Table 2). Twenty-eight were present in 20 patients, who had only this kind of rearrangements. Instead, four patients had the remaining 6 rearrangements of unknown significance in association with a pathogenic CNV (Table 1). All these CNVs were neither reported in the database of genomic variant nor previously identified in other controls/patients. They ranged in size from 2Kb to 940Kb and they had a gene content from 0/1 to 7 genes. The mean content of genes was 2.5

Correlation of clinical features with CNVs

Clinical data of all patients is provided in Table 3. Between the cohort of patients with pathogenic CNVs and the cohort without CNVs, no statistically

relevant differences were identified in gender, weight, head circumference and peculiar facial features. A statistically significant correlation was found between epilepsy and absence of CNVs, especially considering cases without CNVs versus cases with CNVs - all classes (p=0,002). A nearly statistically significant association was identified between congenital anomalies and the presence of

CNVs - class I only (p=0,09). CNVs - all classes are over-represented in patients with Autism in respect to PDD-NOS, both considering the solely group of patients with Autism (p=0,06) and that with Autism plus Asperger (p=0,07). In addition CNVs - all classes are under-represented (p=0,06) among patients with height >97^ocnt.

Table 3 Comparison of clinical features of patients with and without CNVs.

Phenotypes	All cases	Cases without CNVs ^(a)	Cases with CNVs - all classes ^(b)	Cases with CNVs - class I only ^(c)	P value ^(d)	P value ^(e)
Total cases	95	60	35	15		
Gender						
Male	80	50	30	15	0.22	0.09
Female	15	10	5	0		
ASDs						
Autism	48	27	21	7	0.06	0.22
PDD-NOS	44	31	13	7	0.07	0.21
Asperger	3	2	1	0	-	-
ID						
Yes (IQ<70)	92	58	34	15	-	-
No (IQ≥70)	3	2	1	0		
Weight <3 rd centile	4	3	1	0	0.38	0.51
Height <3 rd centile	4	3	1	0	0.38	0.51
Weight >97 th centile	17	11	6	3	0.22	0.28
Height >97 th centile	13	11	2	3	0.06	0.28
Microcephaly	5	2	3	0	0.20	0.64
Macrocephaly	20	14	6	3	0.16	0.26
Epilepsy	12	11	1	0	0.02	0.07
Peculiar facial features	59	38	21	9	0.16	0.22
Congenital anomalies	23	13	10	6	0.14	0.09

Note a, b, c: number of patients.

Note d: P value comparing cases without CNVs versus cases with CNVs (all classes).

Note e: P value comparing cases without CNVs versus cases with CNVs (class I).

DISCUSSION

We report the investigation of CNVs in a cohort of 95 patients with ASD. We could confidently assign pathogenicity in 15 cases, which corresponds to about 16% of the cohort. This is higher than the 9% and 11.6% reported by Qiao Y *et al* 2009 and Christian *et al* 2008, respectively, owing to the higher resolution of oligo-array platform in respect to BAC platform. [5] [4] On the basis of this, the beneficial use of higher resolution array is clear. Our detection rate, considering only pathogenic CNVs, is similar to that reported in Marshall 2008, 14%, who used SNP array. [3] We prioritized CNV pathogenicity using the following criteria. Firstly, we considered the *de novo* inheritance of CNV (class Ia of table 1). Although it is possible to detect a *de novo* event of a benign variant, in the clinical practice we suggest that all *de novo* events be considered as pathogenic, keeping in mind that using this as the only criterion, a percentage of false positive cases will be included. As a second criterion, we used the literature data on known susceptibility CNVs (class Ib). We suggest clearly stating in genetic counseling reports, that these CNVs, although in most cases inherited from healthy parents, are a cofactor of the disease. As a third criterion we used literature data on association studies and knowledge on gene function/expression (class Ic). In this case, we propose that it should be indicated in the medical report that the region is a “likely” cofactor in the disease. Overall, in our opinion CNVs of class Ia, Ib and Ic should be considered as pathogenic.

As an example of class Ic, we discuss more extensively the case of the 7q31 deletion. The involvement of the 7q31 deletion is supported by three pieces of evidence. Firstly, this region includes the *IMMP2L* gene, designated autism susceptibility locus 1 (AUTS1). Secondly, a disruption of *IMMP2L* has been reported in a patient with Tourette syndrome. Lastly, *IMMP2L* intron 3 contains a gene

that is highly expressed in fetal brain (*LRRN3*).[15]

In addition to the 15 cases with pathogenic CNVs, we identified CNVs of unknown significance (class II) in 20 cases, which correspond to about 21% of the cohort. This group, which is even larger than class I group in our cohort, represents a challenge in genetic counseling. Previous studies reported this category at a variable rate from 3% in Qiao *et al* to 30% in Marshall *et al*. [35] [3] In these cases a deep investigation of gene content together with 1Mb of the surrounding region is essential. We suggest that these CNVs should be considered to be “uncertain cofactors” and to clearly state in genetic counseling reports that “an involvement of this CNV in the disease is not completely ruled out”.

CNVs in class II can be prioritized using the following combination of criteria. Firstly, we considered those rearrangements on the X chromosome identified in a male patient and maternally inherited, since they could potentially imply an X-linked recessive defect (class IIa of table 2). Secondly, we considered the concurrent inheritance of two rearrangements, one from the mother and one from the father, potentially to imply digenic inheritance (class IIb).

As a third criterion, we examined literature data pertaining to genes that were potentially implicated in secondary, but not necessarily psychiatric neurological, phenotypic signs. These data implied that rearrangements in these genes were, at least partially, correlated with the phenotype (class IIc). One example of this category is the maternally inherited 3q29 microduplication that included the *RPL35A* gene, whose loss of function mutations in heterozygous state cause Diamond-Blackfan Anemia. In this condition macrocytic anemia, short stature, microcephaly and mental retardation are present. Interestingly, our patient (AU271) has mental retardation associated with microcytic anemia, tall stature, and

macrocephaly, these latter characteristics being present also in the mother. Therefore, one can be hypothesized that the *RLP35A* gene duplication is responsible for the reciprocal phenotype of Diamond-Blackfan Anemia present both in the patient and in the mother, while the susceptibility to autism of the patient was well explained by the concurrent 11p12 microdeletion (table 1-class Ib) . All the remaining cases that do not fit the three previous categories are classified as class IId.

Overall, in order to obtain a correct interpretation of array CGH output in Autistic Spectrum Disorders we suggest not only accurate literature revision and gene content analysis but also accurate parents investigation both from molecular and clinical point of view. This process makes array-CGH analysis among one of the more complex of genetic investigation. In explaining class IIa output during the genetic counseling a recurrence risk up to 50% of male sibling have to be considered. In class IIb the consultant has to explain that the recurrence risk may rise up to the 25% independently to the gender. In class IIc specific clinical signs may be allocated to specific rearrangements potentially leading the accurate prediction of fetus phenotype. In class IId output it could be clearly stated that the recurrence risk is not predictable.

Overall, in this cohort CNVs were identified in 35 patients out of 95, corresponding to a detection rate of 37%. Duplications and deletions had the same frequency even considering only pathogenic CNVs (class I). These data are in contrast with the CNVs identified in syndromic ID in which deletions are over-represented.[36] In syndromic ID the rearrangement is usually *de novo*, and it is considered the major determinant of the disease. The reciprocal duplication is usually associated with a milder phenotype that may escape CNVs investigation. On the contrary, ASDs are considered multifactorial diseases in which the CNV may represent one among several

etiological factors. Therefore, deletions and duplications are expected to occur at the same rate.

Our most important finding coming from the correlation of clinical features with CNVs was that positive output of array-CGH analysis is more expected in Autism and Asperger than in PDD-NOS pinpoint that CNV may represent a stronger cofactor in the first two disorders. While the association with CNVs and congenital anomalies was expected, the inverse relationship with epilepsy is a quite unexpected finding. This finding deserves to be confirmed in other larger cohort of patients.

In conclusion, our study confirmed that array-CGH analysis is able to detect the underlying genetic susceptibility factors in a consistent number of ASD patients, strongly indicating that it has become an essential diagnostic tool for assessing ASD patients. Moreover, this study has provided additional data on the genetic causes of ASD, and has proposed to take into account the possible interaction between genes located in different chromosomal regions when attempting to identify new disease genes. Furthermore, our findings underscore a challenge in genetic counselling in interpreting array-CGH data in a consistent number of patients, and we propose some practical recommendations on the management of such data and on the communication of these findings to the family.

Acknowledgements

This work was supported by Ricerca Finalizzata 2007 Ministero della Salute to A.R. and by PAR 2006 University of Siena to FM. We acknowledge "Cell Lines and DNA Bank of Rett syndrome, X linked mental retardation and other genetic diseases" (Medical Genetics-Siena)-Telethon Genetic Biobank Network (Project No. GTB07001C)

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in Journal of Medical Genetics and any other BMJGL products to exploit all subsidiary rights, as set out in our licence.

References

- 1.Fombonne E. Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry* 2005;66 Suppl 10:3-8.
- 2.Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimaki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. Strong association of de novo copy number mutations with autism. *Science* 2007;316:445-9.
- 3.Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapuram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficocioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82:477-88.
- 4.Christian SL, Brune CW, Sudi J, Kumar RA, Liu S, Karamohamed S, Badner JA, Matsui S, Conroy J, McQuaid D, Gergel J, Hatchwell E, Gilliam TC, Gershon ES, Nowak NJ, Dobyns WB, Cook EH, Jr. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry* 2008;63:1111-7.
- 5.Qiao Y, Riendeau N, Koochek M, Liu X, Harvard C, Hildebrand MJ, Holden JJ, Rajcan-Separovic E, Lewis ME. Phenomic determinants of genomic variation in autism spectrum disorders. *J Med Genet* 2009;46:680-8.
- 6.Pescucci C, Caselli R, Grosso S, Mencarelli MA, Mari F, Farnetani MA, Piccini B, Artuso R, Bruttini M, Priolo M, Zuffardi O, Gimelli S, Balestri P, Renieri A. 2q24-q31 deletion: report of a case and review of the literature. *Eur J Med Genet* 2007;50:21-32.
- 7.Prodi DA, Pirastu N, Maninchedda G, Sassu A, Picciau A, Palmas MA, Mossa A, Persico I, Adamo M, Angius A, Pirastu M. EDA2R is associated with androgenetic alopecia. *J Invest Dermatol* 2008;128:2268-70.
- 8.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, Wiczorek G, Ropers HH. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* 2007;28:674-82.
- 9.Szafranski P, Schaaf CP, Person RE, Gibson IB, Xia Z, Mahadevan S, Wiszniewska J, Bacino CA, Lalani S, Potocki L, Kang SH, Patel A, Cheung SW, Probst FJ, Graham BH, Shinawi M, Beaudet AL, Stankiewicz P. Structures and molecular mechanisms for common 15q13.3 microduplications involving CHRNA7: benign or pathological? *Hum Mutat* 2010;31:840-50.
- 10.Nagamani SC, Erez A, Shen J, Li C, Roeder E, Cox S, Karaviti L, Pearson M, Kang SH, Sahoo T, Lalani SR, Stankiewicz P, Sutton VR, Cheung SW. Clinical spectrum associated with recurrent genomic rearrangements in chromosome 17q12. *Eur J Hum Genet* 2010;18:278-84.
- 11.Doornbos M, Sikkema-Raddatz B, Ruijvenkamp CAL, Dijkhuizen T, Bijlsma EK, Gijsbers ACJ, Hilhorst-Hofstee Y, Hordijk R, Verbruggen KT, Kerstjens-Frederikse WS, van Essen T, Kok K, van Silfhout AT, Breuning M, van Ravenswaaij-Arts CMA. Nine patients with a microdeletion 15q11.2 between breakpoints 1 and 2 of the Prader-Willi critical region, possibly associated with behavioural disturbances. *European Journal of Medical Genetics* 2009;52:108-115.
- 12.Steichen-Gersdorf E, Gassner I, Superti-Furga A, Ullmann R, Stricker S, Klopocki E, Mundlos S. Triangular tibia with fibular aplasia associated with a microdeletion on 2q11.2 encompassing LAF4. *Clin Genet* 2008;74:560-5.
- 13.Maas NM, Van de Putte T, Melotte C, Francis A, Schrandt-Stumpel CT, Sanlaville

- D, Genevieve D, Lyonnet S, Dimitrov B, Devriendt K, Fryns JP, Vermeesch JR. The C20orf133 gene is disrupted in a patient with Kabuki syndrome. *J Med Genet* 2007;44:562-9.
14. Ching MS, Shen Y, Tan WH, Jeste SS, Morrow EM, Chen X, Mukaddes NM, Yoo SY, Hanson E, Hundley R, Austin C, Becker RE, Berry GT, Driscoll K, Engle EC, Friedman S, Gusella JF, Hisama FM, Irons MB, Lafiosca T, LeClair E, Miller DT, Neessen M, Picker JD, Rappaport L, Rooney CM, Sarco DP, Stoler JM, Walsh CA, Wolff RR, Zhang T, Nasir RH, Wu BL. Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. *Am J Med Genet B Neuropsychiatr Genet* 2010;153B:937-47.
15. Maestrini E, Pagnamenta AT, Lamb JA, Bacchelli E, Sykes NH, Sousa I, Toma C, Barnby G, Butler H, Winchester L, Scerri TS, Minopoli F, Reichert J, Cai G, Buxbaum JD, Korvatska O, Schellenberg GD, Dawson G, de Bildt A, Minderaa RB, Mulder EJ, Morris AP, Bailey AJ, Monaco AP. High-density SNP association study and copy number variation analysis of the AUTS1 and AUTS5 loci implicate the IMMP2L-DOCK4 gene region in autism susceptibility. *Mol Psychiatry* 2009;15:954-68.
16. Le Hellard S, Lee AJ, Underwood S, Thomson PA, Morris SW, Torrance HS, Anderson SM, Adams RR, Navarro P, Christoforou A, Houlihan LM, Detera-Wadleigh S, Owen MJ, Asherson P, Muir WJ, Blackwood DH, Wray NR, Porteous DJ, Evans KL. Haplotype analysis and a novel allele-sharing method refines a chromosome 4p locus linked to bipolar affective disorder. *Biol Psychiatry* 2007;61:797-805.
17. Angata K, Suzuki M, Fukuda M. ST8Sia II and ST8Sia IV polysialyltransferases exhibit marked differences in utilizing various acceptors containing oligosialic acid and short polysialic acid. The basis for cooperative polysialylation by two enzymes. *J Biol Chem* 2002;277:36808-17.
18. Karlsson T, Kullander K, Welsh M. The Src homology 2 domain protein Shb transmits basic fibroblast growth factor- and nerve growth factor-dependent differentiation signals in PC12 cells. *Cell Growth Differ* 1998;9:757-66.
19. Angers S, Li T, Yi X, MacCoss MJ, Moon RT, Zheng N. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* 2006;443:590-3.
20. Nishino J, Yamashita K, Hashiguchi H, Fujii H, Shimazaki T, Hamada H. Meteorin: a secreted protein that regulates glial cell differentiation and promotes axonal extension. *Embo J* 2004;23:1998-2008.
21. Mencarelli MA, Katzaki E, Papa FT, Sampieri K, Caselli R, Uliana V, Pollazzon M, Canitano R, Mostardini R, Grosso S, Longo I, Ariani F, Meloni I, Hayek J, Balestri P, Mari F, Renieri A. Private inherited microdeletion/microduplications: implications in clinical practice. *Eur J Med Genet* 2008;51:409-16.
22. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;358:667-75.
23. Behjati F, Shafaghati Y, Firouzabadi SG, Kahrizi K, Bagherizadeh I, Najmabadi H, Bint S, Ogilvie C. M-banding characterization of a 16p11.2p13.1 tandem duplication in a child with autism, neurodevelopmental delay and dysmorphism. *Eur J Med Genet* 2008;51:608-14.
24. van Bon BW, Mefford HC, Menten B, Koolen DA, Sharp AJ, Nillesen WM, Innis JW, de Ravel TJ, Mercer CL, Fichera M, Stewart H, Connell LE, Ounap K, Lachlan K, Castle B, Van der Aa N, van Ravenswaaij C, Nobrega MA, Serra-Juhe C, Simonic I, de Leeuw N, Pfundt R, Bongers EM, Baker C, Finnemore P, Huang S, Maloney VK, Crolla JA, van Kalmthout M, Elia M, Vandeweyer G, Fryns JP, Janssens S, Foulds N, Reitano S, Smith K, Parkel S, Loeys B, Woods CG, Oostra A, Speleman F, Pereira AC, Kurg A, Willatt L, Knight SJ, Vermeesch JR, Romano C, Barber JC, Mortier G, Perez-Jurado LA, Kooy F, Brunner HG, Eichler EE, Kleefstra T,

- de Vries BB. Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from non-pathogenic to a severe outcome. *J Med Genet* 2009;46:511-23.
25. Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH, van Haeringen A, Franssen van de Putte DE, Anderlid BM, Lundin J, Lapunzina P, Perez Jurado LA, Delle Chiaie B, Loeys B, Menten B, Oostra A, Verhelst H, Amor DJ, Bruno DL, van Essen AJ, Hordijk R, Sikkema-Raddatz B, Verbruggen KT, Jongmans MC, Pfundt R, Reeser HM, Breuning MH, Ruivenkamp CA. Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur J Med Genet* 2009;52:77-87.
26. Hannes FD, Sharp AJ, Mefford HC, de Ravel T, Ruivenkamp CA, Breuning MH, Fryns JP, Devriendt K, Van Buggenhout G, Vogels A, Stewart H, Hennekam RC, Cooper GM, Regan R, Knight SJ, Eichler EE, Vermeesch JR. Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *J Med Genet* 2009;46:223-32.
27. Li F, Shen Y, Kohler U, Sharkey FH, Menon D, Coulleaux L, Malan V, Rio M, McMullan DJ, Cox H, Fagan KA, Gaunt L, Metcalfe K, Heinrich U, Hislop G, Maye U, Sutcliffe M, Wu BL, Thiel BD, Mulchandani S, Conlin LK, Spinner NB, Murphy KM, Batista DA. Interstitial microduplication of Xp22.31: Causative of intellectual disability or benign copy number variant? *Eur J Med Genet* 2010;53:93-9.
28. Rainier S, Chai JH, Tokarz D, Nicholls RD, Fink JK. NIPA1 gene mutations cause autosomal dominant hereditary spastic paraplegia (SPG6). *Am J Hum Genet* 2003;73:967-71.
29. Schenck A, Bardoni B, Moro A, Bagni C, Mandel JL. A highly conserved protein family interacting with the fragile X mental retardation protein (FMRP) and displaying selective interactions with FMRP-related proteins FXR1P and FXR2P. *Proc Natl Acad Sci U S A* 2001;98:8844-9.
30. Chai JH, Locke DP, Grealley JM, Knoll JH, Ohta T, Dunai J, Yavor A, Eichler EE, Nicholls RD. Identification of four highly conserved genes between breakpoint hotspots BP1 and BP2 of the Prader-Willi/Angelman syndromes deletion region that have undergone evolutionary transposition mediated by flanking duplicons. *Am J Hum Genet* 2003;73:898-925.
31. Lin JC, Ho WH, Gurney A, Rosenthal A. The netrin-G1 ligand NGL-1 promotes the outgrowth of thalamocortical axons. *Nat Neurosci* 2003;6:1270-6.
32. Yan M, Wang LC, Hymowitz SG, Schilbach S, Lee J, Goddard A, de Vos AM, Gao WQ, Dixit VM. Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* 2000;290:523-7.
33. Guan CP, Lin FQ, Hong WS, Fu LF, Zhou MN, Xu W, Xu AE. [Expression of InnVit/FBXO11 in vitiligo and its role in tyrosinase export from endoplasmic reticulum]. *Zhonghua Yi Xue Za Zhi* 2010;90:1126-30.
34. Segade F, Daly KA, Allred D, Hicks PJ, Cox M, Brown M, Hardisty-Hughes RE, Brown SD, Rich SS, Bowden DW. Association of the FBXO11 gene with chronic otitis media with effusion and recurrent otitis media: the Minnesota COME/ROM Family Study. *Arch Otolaryngol Head Neck Surg* 2006;132:729-33.
35. Qiao Y, Harvard C, Tyson C, Liu X, Fawcett C, Pavlidis P, Holden JJ, Lewis ME, Rajcan-Separovic E. Outcome of array CGH analysis for 255 subjects with intellectual disability and search for candidate genes using bioinformatics. *Hum Genet* 2010;128:179-94.
36. Wagenstaller J, Spranger S, Lorenz-Depiereux B, Kazmierczak B, Nathrath M, Wahl D, Heye B, Glaser D, Liebscher V, Meitinger T, Strom TM. Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation. *Am J Hum Genet* 2007;81:768-79.

4.2.2 Genetic modifier factor

**Investigation of modifier genes within copy number variations in
Rett syndrome.**

Rosangela Artuso*, Papa Filomena Tiziana*, Mucciolo Mafalda, Dag H. Yasui, Keith W. Dunaway, Vittoria Disciglio, Maria Antonietta Mencarelli, Marzia Pollazzon, Francesca Ariani, Francesca Mari, Janine M. LaSalle, Alessandra Renieri.

Submitted to Molecular Autism

Investigation of modifier genes within copy number variations in Rett syndrome.

Rosangela Artuso^{1*}, Filomena T Papa^{1*}, Mafalda Mucciolo¹, Dag H Yasui², Keith W Dunaway², Vittoria Disciglio¹, Maria Antonietta Mencarelli¹, Marzia Pollazzon¹, Francesca Ariani¹, Francesca Mari¹, Janine M LaSalle², Alessandra Renieri^{1§}

¹ Medical Genetics, University of Siena Italy

² Medical Microbiology and Immunology, Genome Center, School of Medicine, University of California, Davis, CA

*These authors contributed equally to this work

§Corresponding author

Abstract

Background Mutations in the X-linked *MECP2* gene are responsible for two different phenotypes in females, classical and Zappella variant Rett syndrome. In classical Rett, girls show somatic hypoevolutism, are unable to speak, possess severely impaired motor function and have acute mental retardation. Zappella variant patients recover the ability to speak in either simple or complex often ecolalic phrases, regain purposeful hand function, have moderate mental retardation and generally show normal head circumference, weight and height.

Methods We explored the hypothesis that Copy Number Variants (CNVs) leading to over/underexpression of genes may modulate the phenotype by comparing by array-CGH two discordant pairs of Rett sisters and four additional discordant pairs of unrelated Rett girls matched by mutation type. By ChIP-chip analysis we also identify hypothetical MeCP2 targets included in the identified CNVs.

Results Our study failed to identify one major common modifier gene/region, suggesting genetic modifiers may be complex and variable between cases. However, we demonstrated that several common CNVs are enriched for MeCP2 bound promoters and could influence the severity of the phenotype in complex ways.

On 1p36.13 *CROCC*, an hypothetical MeCP2 target, results duplicated in a Zappella variant and deleted in a classic Rett patient. It encodes for a major structural component of the ciliary rootlet and ciliary motility is required for correct brain development and function. On 1q31.3 *CFHR1* and *CFHR3*, deleted in a Zappella variant and duplicated in two classic Rett patients, are involved in the regulation of complement whose proteins are important for CNS synapse elimination. The duplication on 10q11.22, present in two Zappella variant patients, includes *GPRIN2*, involved in the regulation of neurite outgrowth and *PPYR1*, a key regulator of energy homeostasis and directly implicated in the regulation of food intake.

Conclusions Finding genetic modifiers of the Rett phenotype by CNVs analyses is by nature complex because MeCP2 pathways are influenced by both genetic and epigenetic variables. However this study highlights important hints to further investigate by expression and/or statistical analysis in a higher number of patients.

Background

Rett syndrome (RTT, OMIM#312750) is an X-linked neurodevelopmental disorder predominantly affecting females. In the classic form, after a period of normal development (6-18 months), patients show growth retardation and regression of speech and purposeful hand movements, with appearance of stereotyped hand movements, microcephaly, autism, seizures, and somatic hypoevolutism [1] [2]. RTT syndrome has been the object of extensive investigations, revealing a wide spectrum of clinical phenotypes including: the Zappella variant (Z-RTT), the early onset seizure variant, the congenital variant, the 'forme fruste', and the late regression variant [3]. Z-RTT, firstly described by M. Zappella in 1992, represents the most common RTT variant and it is characterized by the recovery of the ability to speak in single words or third person phrases and by an improvement of purposeful hand movements [4] [5]. Z-RTT patients also show milder intellectual disabilities (up to IQ of 50) and often normal head circumference, weight and height respect to classic RTT [5].

De novo mutations in the *MECP2* gene (Xq28) account for the majority of girls with classic RTT (95%) and for about half of cases with Z-RTT. Only a few familial cases have been reported. Some cases have been explained by skewing of X-inactivation towards the wild type allele in an asymptomatic carrier [6] [7]. In others cases, germline mosaicism has been a possible explanation [8] [9] [10].

RTT symptoms can vary depending on the pattern of X-chromosome inactivation, the type of *MECP2* mutation, and environmental factors [5] [11] [12]. However, all these factors are not sufficient to explain RTT clinical variability suggesting that other elements may be involved. In fact there are cases of RTT sisters with identical *MECP2* mutation, balanced X-inactivation, similar

environmental stimulation and discordant phenotype (one classic and one Z-RTT sister) [6] [8].

Copy Number Variations (CNVs) are segments of DNA ranging from kilobases to multiple megabases in length that are present at a variable number of copies compared with the reference genome sequence. It has been demonstrated that CNVs are associated with quantitative gene expression levels that in some cases are likely to have causative, functional effects [13]. CNVs have been reported to be associated with human diseases such as neurological and autoimmune disorders and cancer [14] [15] [16]. CNVs, to a larger extent than Single Nucleotide Polymorphisms (SNPs), represent an important source of variability in both phenotypically normal subjects and individuals with diseases [17] [18]. It is therefore reasonable to hypothesize that CNVs can modulate the phenotypic expression of RTT syndrome.

In order to test this hypothesis, we analyzed by array-CGH two couples of RTT sisters and four additional couples of unrelated RTT girls matched by mutation type showing discordant phenotype (classic and Z-RTT). By ChIP-chip analysis we also identify hypothetical *MeCP2* targets included in the identified CNVs.

Methods

Patients

From the Italian RTT database and biobank, we recruited two rare familial cases with two RTT sisters with discordant phenotype: one classic (#897 and #138) and one Z-RTT (#896 and #139) (www.biobank.unisi.it) [19]. These cases were screened by both DHPLC and MLPA techniques to identify *MECP2* mutations. The first couple bears a *MECP2* large deletion involving exon 3 and exon 4, while the second one has a late truncating mutation: c.1157del32. Clinical descriptions of these patients have been

reported in previous manuscripts [6] [8]. Furthermore, we selected four more pairs (#565/601, #185/119, #421/109, #402/368) of unrelated RTT patients showing discordant phenotype (classic and Z-RTT) coupled by *MECP2* mutation type (c.1163del26, p.R306C, c.1159del44, p.R133C). All cases contained in the bank have been clinically evaluated by the Medical Genetics Unit of Siena. Patients were classified in classic and RTT variant according to the international criteria [20].

Genomic DNA isolation

Genomic DNA of the patients was isolated from an EDTA peripheral blood sample using the QIAamp DNA Blood Kit according to the manufacturer's protocol (Qiagen, www.qiagen.com). Genomic DNA of normal male and female controls was obtained from Promega. Ten micrograms of genomic DNA from the patient (test sample) and the control (reference sample) were sonicated. Test and reference DNA samples were subsequently purified using dedicated columns (DNA Clean and Concentrator, Zymo Research) and the appropriate DNA concentrations were determined by a DyNA Quant™ 200 Fluorometer (GE Healthcare).

Array Comparative Genomic Hybridization

Array CGH analysis was performed using commercially available oligonucleotide microarrays containing about 99,000 60-mer probes with an estimated average resolution of about 65 Kb. Physical positions of the probes correspond to the UCSC genome browser - NCBI build 36/hg18, March 2006 (<http://genome.ucsc.edu>).

DNA labeling was performed according to the Agilent Genomic DNA Labeling Kit Plus protocol (Oligonucleotide Array-Based CGH for Genomic DNA Analysis 2.0v). Genomic DNA (3.5 µg) from patients with classical Rett syndrome was mixed with Cy5-dNTP while genomic DNA from patient with Zappella variant was mixed with Cy3-

dNTP, as previously reported [21]. The array was disassembled and washed according to the manufacturer protocol with wash buffers supplied with the Agilent 105A kit. The slides were dried and scanned using an Agilent G2565BA DNA microarray scanner.

Array-CGH image and data analysis

Image analysis was performed using the CGH Analytics software v. 5.0.14 with aberration algorithms Z-Score and ADM-1 provided by the CGH-Analytics software. The software automatically determines the fluorescence intensities of the spots for both fluorochromes performing background subtraction and data normalization, and compiles the data into a spreadsheet that links the fluorescent signal of every oligo on the array to the oligo name, its position on the array and its position in the genome. The linear order of the oligos is reconstituted in the ratio plots consistent with an ideogram. The ratio plot is arbitrarily assigned such that gains and losses in DNA copy number at a particular locus are observed as a deviation of the ratio plot from a modal value of 1.0.

Identification of MeCP2 bound promoters

ChIP-chip analysis was performed as described previously [22] using a modified protocol from Oberley et al. [23]. Briefly, chromatin from three separate SH-SY5Y neuroblastoma cultures differentiated by 48h treatment with phorbol 12-myristate 13-acetate (PMA) was crosslinked by addition of formaldehyde to 1% by volume. Crosslinked cells were lysed by incubation in SDS containing buffer and donec homogenization and then sonicated to fragment chromatin. MeCP2 bound chromatin was isolated by immunoprecipitation with a custom anti-MeCP2 antibody. As a control, chromatin was immunoprecipitated in parallel with anti-RNA polymerase II antibody (Covance, Berkeley, CA). DNA was isolated from immunoprecipitated

chromatin by proteinase K digestion and phenol/chloroform extraction prior to linear amplification by ligation mediated PCR (LM-PCR).

Amplicons from ChIP samples were subsequently labeled with Cy5 according to Nimblegen protocol (Roche Nimblegen, Madison, WI) and hybridized along with Cy3 labeled sonicated total SH-SY5Y genomic DNA to a commercial genome wide promoter microarray produced by Nimblegen according to the manufacturer's instructions. In this 1.5 kb promoter array, tiled oligonucleotide probes extend 1.3 kb upstream and 0.2 kb downstream of the transcriptional start sites of 24,275 human transcripts. Arrays were washed, dried and scanned using a GenePix 4000B scanner (Molecular Devices, Sunnyvale, CA) to obtain raw fluorescent intensity data. NimbleScan software (Nimblegen) was used to assemble log₂ signal intensity for each tiled oligonucleotide for further analysis. Statistical analysis of promoter ChIP-chip data indicated that 2600-4300 promoters were bound by MeCP2 with 1524 promoters common to two replicate hybridizations. Promoters were ranked according to MeCP2 binding levels for the two arrays.

Analyses on phenotypically discordant RTT pairs resulted in 28 CNV's that gave a list of 55 genes which could potentially modify RTT phenotype. A Perl script was used to search for these 55 genes within two experimentally replicated promoter hit files that ranked MeCP2 binding based on ChIP-chip analysis. Aliases of genes were acquired from genome.ucsc.edu and genelist.org.

Results

Results of CNVs analysis in phenotypically discordant RTT pairs are shown in Tables 1 and 2. Overall, we identified 29 CNVs, 28 of them being known polymorphic regions and one being an apparently private rearrangement duplicated in only one Z-RTT patient. This private CNV localized to 3q13.12,

extended for about 99 Kb and contained two coding genes: *GUCA1C* expressed in retina and *MORC1* expressed in testis.

Some polymorphic CNVs were found in more than one pair of RTT pairs. Among the 28 CNVs, we considered 14 of them as "unlikely modifier" since results were apparently not associated with phenotypic severity (Table 2). These include regions containing olfactory receptors and class II HLA molecules that are not expected to directly correlate with the phenotypic variability related to classic/Z-RTT phenotype.

To determine if the CNVs found in phenotypically discordant RTT pairs contained possible MeCP2 target genes, we compared promoter rankings of MeCP2 binding using promoter-wide ChIP-chip analysis [22]. The ranking from total number of genes from 1 to 24,134 is shown for two replicate MeCP2-ChIP microarrays (MeCP2 B and MeCP2 C promoter hits rank, Tables 1 and 2). Genes with promoters in the top 10% of MeCP2 promoter hits for at least one replicate are shaded in blue. Among CNVs classified as "likely modifiers", ChIP-chip analysis identified potential MeCP2 target genes within the 1p36.13 (*CROCC* gene whose duplication was found in the Z-RTT # 896 and deletion in the classic form #402) and the 2p25.2 (*TSSC1* gene whose deletion was found in the Z-RTT #896) region. Among CNVs classified as "unlikely modifiers", ChIP-chip analysis identified potential MeCP2 target genes on 14q11 (*OR4Q3* and *OR4Q1*, deleted in a classic patient #138 and duplicated in another classic patient #421) and on 16p11.2 (*NFATC2IP* and *SPNS1*, duplicated in both a classic #897 and a Z-RTT patient #368).

Discussion

In order to test the hypothesis that CNVs may modulate RTT phenotype, we analyzed by array-CGH two couples of RTT sisters and four additional couples of unrelated RTT girls matched by *MECP2* mutation type showing discordant phenotype: classic and Z-RTT. Our study

failed to identify one major common modifier gene/region, suggesting genetic modifiers may be complex and variable between cases (Tables 1 and 2). In total we found 29 CNVs that were divided into two groups: “likely modifiers” and “unlikely modifiers” (Tables 1 and 2).

Among the first group, the rearrangement on 1p36.13 includes *CROCC* (ciliary rootlet coiled-coil), that represents an interesting potential modifier gene. This gene is duplicated in the Z-RTT patient # 896 and deleted in the classical patient # 402, suggesting that change in its expression may modulate RTT outcome. Moreover, according to ChIP-chip analysis, *CROCC* could be a potential MeCP2 target gene (Table 1). *CROCC* encodes for a major structural component (Rootletin) of the ciliary rootlet, a cytoskeletal-like structure in ciliated cells which originates from the basal body at the proximal end of a cilium and extends proximally toward the cell nucleus. In nonciliated cells, a miniature ciliary rootlet is located at the centrosome and does not project a fibrous network into the cytoplasm. Rootletin is expressed in retina, brain, trachea and kidney [24]. Cilia generate specialized structures that perform critical functions of several broad types: sensation, development, fluid movement, sperm motility, and cell signaling. Their functional significance in tissues is reflected in the severity and diversity of pathologies caused by defects in cilia. These include anosmia, retinitis pigmentosa and retinal degeneration, polycystic kidney disease, diabetes, neural tube defects and neural patterning defects, chronic sinusitis and bronchiectasis, obesity, heterotaxias, polydactyly, and infertility [25] [26] [27]. Defects in cilia are therefore underlying causes of several diseases with pleiotropic symptoms [28]. Several pleiotropic disorders (Bardet-Biedl syndrome, Alstrom syndrome, Meckel-Gruber syndrome and Joubert syndrome) caused by disruption of the function of cilia present mental retardation or other cognitive defects as part of their

phenotypic spectrum [29]. The presence of cilia in different types of neurons supports the notion that dysfunction in specific neuronal populations might explain, at least in part, such defects. If *MeCP2* acts as a positive regulator of *CROCC*, it can be hypothesized that higher protein level due to three copies of such gene may counteract the *MeCP2* functional reduction, while lower protein level due to single gene copy may worsen the phenotype.

The *CFHR* gene family members (*CFHR1* and *CFHR3*) located on 1q31.3 are duplicated in classic girls (#185 and #402) and deleted in Z-RTT (#368), suggesting that the phenotype may benefit of reduced expression of these proteins involved in complement regulation [30]. The complement system is a tightly controlled part of the host innate immune defense. Imbalances in this control contribute to tissue injury and can result in autoimmune diseases. In particular, *CFHR1* and *CFHR3* was previously associated with hemolytic uremic syndrome (HUS) and age related macular degeneration (AMD) [31] [32] [33]. It is well known that the immune system participate in the development and functioning of the CNS and an immune etiology for RTT and autism has been recently hypothesized [34]. Interestingly, complement proteins have been demonstrated to be fundamental for CNS synapse elimination [35]. Morphological studies in postmortem brain samples from RTT individuals described a characteristic neuropathology which included decreased dendritic arborization, fewer dendritic spines, and increased packing density [36]. It is therefore possible that *CFHR* genes could be involved in the modelling of synaptic connections and that they could influence RTT severity.

The duplication on 10q11.22, present in two Z-RTT patients (#139 and #109), includes two interesting candidate modifier genes: *GPRIN2* and *PPYR1*. The *GPRIN2* gene, highly expressed in the cerebellum, interacts with activated

members of the Gi subfamily of G protein α subunits and acts together with *GPRIN1* to regulate neurite outgrowth [37]. *PPYRI*, also named as neuropeptide Y receptor or pancreatic polypeptide 1, is a key regulator of energy homeostasis and directly involved in the regulation of food intake. Previous studies have reinforced the potential influence of *PPYRI* on body weight in humans [38]. Moreover, it has been demonstrated that *PPYRI* knockout mice display lower body weight and reduced white adipose tissue [39]. Thus a higher level of *PPYRI* expression due to duplication may correlate with the higher body weight characterizing Z-RTT patients respect to classic RTT [5]. In contrast, a recent study demonstrated that 10q11.22 gain is associated with lower body mass index value in the Chinese population. However this CNV is bigger respect to the one reported here and includes two additional genes [40].

The 2p25.2 region, deleted in the Z-RTT patient #896, contains a potential MeCP2 target gene, *TSSCI* (Table 1). Very little is known about this gene. It shares sequence homology with RbAp48 and CAF1, suggesting that it could play a role in gene silencing [41].

The 1q42.12 region, duplicated in one Z-RTT patient (#896), includes *ENAH*. This gene was identified as a mammalian homolog of *Drosophila* Ena and initially named Mena (Mammalian enabled) [42]. It localizes to cell-substrate adhesion sites and sites of dynamic actin assembly and disassembly. It is a member of the Ena/VASP family that also includes VASP and EVL in vertebrates. Work carried out

in *Drosophila*, *C. elegans* and mice showed that these proteins participate in axonal outgrowth, dendritic morphology, synapse formation and also function downstream of attractive and repulsive axon guidance pathways [43]. Previous evidence shows that knocking out the three murine ENA/VASP proteins results in a blockade of axon fiber tract formation in the cortex in vivo, and that failure in neurite initiation is the underlying cause [44] [45]. *ENAH* therefore represents an interesting candidate gene as modifier in RTT. Further investigations are necessary in order to test whether the duplication of *ENAH* gene in Z-RTT #896 effectively corresponds to increase mRNA levels in brain and whether this mechanism is confined to one couple of discordant girls or is a common mechanism in Z-RTT possibly throughout SNP modulation.

Conclusions

One limitation of our study is that the number of patients is too low to perform a statistically significant analysis of CNVs in classic and Z-RTT and this is principally due to the difficulty in recruiting Z-RTT cases. Another limitation is that mRNA expression analysis has not been performed. However, this analysis would not be conclusive because it could be executed only in blood or other accessible human tissues. Regions/genes found rearranged could rather represent interesting hints for further studies in animal models or in new cellular models such as human induced pluripotent stem (iPS).

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by Ricerca Finalizzata 2007 Ministero della Salute to A.R. and by PAR 2006 University of Siena to FM. We acknowledge “Cell Lines and DNA Bank of Rett syndrome, X linked mental retardation and other genetic diseases” (Medical Genetics-Siena) - Telethon Genetic Biobank Network (Project No. GTB07001C).

Polymorphic CNVs	Breakpoints (Bp)	Gene content	MeCP2_B promoter hits rank	MeCP2_C promoter hits rank	897 C/ 896 Z	138 C/ 139 Z	565 C/ 601 Z	185 C/ 119 Z	421 C/ 109 Z	402 C/ 368 Z
1p36.13 {426 kb}	16,698,906- 17,124,554	ESPNP	-	-	Dup Z					Del C
		MSTP9	-	-						
		CROCC	658	16300						
1q31.3 {55kb}	195,011,344- 195,065,867	CFHR1	23144	7651				Dup C		Amp C/ Del Z
		CFHR3	20253	6994						
1q42.12 {139kb}	223,731,557- 223,870,819	ENAH	18604	13553	Dup Z					
2p25.2 {400 kb}	3,060,975- 3,460,506	TSSC1	941	3174	Del Z					
		TTC15	20740	21641						
2q37.3 {141 kb}	242,514,593- 242,655,973	-								Del Z
3q13.12 {99 kb}	110,116,098- 110,397,433	GUCA1C	19293	6167				Dup Z		
		MORC1	18317	20394						
		C3orf66	-	-						
5p15.33 {85 kb}	763,944- 848,744	ZDHC11	4349	13284	Dup Z					
6q27 {210 kb}	168,114,265- 168,324,002	MLLT4	-	-						Dup Z
		C6orf54	2389	6671						
		KIF25	8778	3159						
		FRMD1	15800	10530						
7p21.3 {89 kb}	11,720,901- 11,809,763	THSD7A	8520	8160						Del Z
8q21.3 {87kb}	87,136,222- 87,222,795	PSKH2	19413	9491						Dup C
		ATP6VOD2	22858	4087						
10q11.22 {172kb}	46,396,163- 46,568,496	GPRIN2	-	-		Dup Z				Dup Z
		PPYR1	16722	9812						
14q32.33 {125kb}	105,708,209- 105,833,372	SLK	-	-						Del Z
		COL17A1	-	-						
15q14 {49kb}	32,523,241- 32,572,315	-							Del Z	
16p11 {200kb}	34,399,543- 34,539,890	-	-	-						Dup Z
22q13.2 {43kb}	22,681,995- 22,712,221	GSTT1	9984	14237						Dup Z

Table 1 Likely modifier regions. Genes with promoters in the top 10% of MeCP2 promoter hits for at least one replicate are shaded in blue.

Polymorphic CNVs	Breakpoints (Bp)	Gene content	MeCP2_B promoter hits rank	MeCP2_C promoter hits rank	897 C/ 896 Z	138 C/ 139 Z	565 C/ 601 Z	185 C/ 119 Z	421 C/ 109 Z	402 C/ 368 Z
1q44{58kb}	246,794,522-246,852,126	OR2T34	22567	14244	Dup Z		**Del C	Del C	Del Z	
		OR2T10	22348	8566						
2p11{494kb}	89,401,838-89,895,566	*IGKV2D	-	-		Del Z				
		*IGKV1D	-	-						
		*IGKV3D	-	-						
3q26{134kb}	163,997,228-164,101,776	-				Del C	Del Z			Del C/ Del Z
3q29{36 kb}	196,905,767-196,942,158	MUC20	-	-			Dup C	Dup C/ Dup Z	Dup C	Dup C
4q13.2{108kb}	69,057,735-69,165,814	UGT2B17	21719	18336		Dup C			Del Z	Del C
6p21.32{65kb}	29,939,288-30004,636	HCG4P6	-	-	Del C					Del Z
6p21.33{77kb}	32,595,402-32,672,983	HLA-DRB5	-	-		Dup Z		**Amp C		Amp C/ Amp Z
		HLA-DRB1	-	-						
8p11.23{143kb}	39,356,595-39,499,752	ADAM5P	14070	2916	Dup C	Dup C	Amp Z	Amp C/ Amp Z		Amp C/ Amp Z
10q11.22{84kb}	47,017,598-47,161,232	-	-	-		Dup C	Dup Z			Dup Z
14q11{860kb}	18,624,383-19,484,013	OR11H13P	-	-		Del C			Dup C	
		OR4Q3	23539	330						
		OR4M1	24054	4686						
		OR4N2	23383	7030						
		OR4K2	21957	3567						
		OR4K5	21814	7944						
		OR4K1	23684	162						
15q11.2{727}	18,810,004-19,537,035	-			Del C		Del C	Del Z		Del C
16p11.2{220kb}	28,732,295-28,952,218	ATXN2L	-	-	Dup C					Dup Z
		TUFM	7097	3848						
		SH2B1	12675	12680						
		ATP2A1	15182	23566						
		RABEP2	10298	16794						
		CD19	18469	15604						
		NFATC2IP	23326	1627						
		SPNS1	1393	20317						
LAT	15145	14271								
17q21.31{163kb}	41,544,224-41,706,870	KIAA1267	-	-			Amp C/ Dup Z	Dup Z		Dup Z
22q13.2{43kb}	41,237,731-41,287,060	SERHL RRP7A SERHL2	-	-					Dup Z	

*16 isoforms

** smaller rearrangement in these patients

Table 2 Unlikely modifier regions. Genes with promoters in the top 10% of MeCP2 promoter hits for at least one replicate are shaded in blue.

References

1. Trevathan E, Adams MJ: **The epidemiology and public health significance of Rett syndrome.** *J Child Neurol* 1988, 3 Suppl:S17-20.
2. Lasalle JM, Yasui DH: **Evolving role of MeCP2 in Rett syndrome and autism.** *Epigenomics* 2009, 1(1):119-130.
3. Mencarelli MA, Spanhol-Rosseto A, Artuso R, Rondinella D, De Filippis R, Bahi-Buisson N, Nectoux J, Rubinsztajn R, Bienvenu T, Moncla A *et al*: **Novel FOXP1 mutations associated with the congenital variant of Rett syndrome.** *J Med Genet* 2010, 47(1):49-53.
4. Zappella M: **The Rett girls with preserved speech.** *Brain Dev* 1992, 14(2):98-101.
5. Renieri A, Mari F, Mencarelli MA, Scala E, Ariani F, Longo I, Meloni I, Cevenini G, Pini G, Hayek G *et al*: **Diagnostic criteria for the Zappella variant of Rett syndrome (the preserved speech variant).** *Brain Dev* 2009, 31(3):208-216.
6. Zappella M, Meloni I, Longo I, Hayek G, Renieri A: **Preserved speech variants of the Rett syndrome: molecular and clinical analysis.** *Am J Med Genet* 2001, 104(1):14-22.
7. Hardwick SA, Reuter K, Williamson SL, Vasudevan V, Donald J, Slater K, Bennetts B, Bebbington A, Leonard H, Williams SR *et al*: **Delineation of large deletions of the MECP2 gene in Rett syndrome patients, including a familial case with a male proband.** *Eur J Hum Genet* 2007, 15(12):1218-1229.
8. Scala E, Longo I, Ottimo F, Speciale C, Sampieri K, Katzaki E, Artuso R, Mencarelli MA, D'Ambrogio T, Vonella G *et al*: **MECP2 deletions and genotype-phenotype correlation in Rett syndrome.** *Am J Med Genet A* 2007, 143A(23):2775-2784.
9. Evans JC, Archer HL, Whatley SD, Clarke A: **Germline mosaicism for a MECP2 mutation in a man with two Rett daughters.** *Clin Genet* 2006, 70(4):336-338.
10. Mari F, Caselli R, Russo S, Cogliati F, Ariani F, Longo I, Bruttini M, Meloni I, Pescucci C, Schurfeld K *et al*: **Germline mosaicism in Rett syndrome identified by prenatal diagnosis.** *Clin Genet* 2005, 67(3):258-260.
11. Archer H, Evans J, Leonard H, Colvin L, Ravine D, Christodoulou J, Williamson S, Charman T, Bailey ME, Sampson J *et al*: **Correlation between clinical severity in patients with Rett syndrome with a p.R168X or p.T158M MECP2 mutation, and the direction and degree of skewing of X-chromosome inactivation.** *J Med Genet* 2007, 44(2):148-152.
12. Lonetti G, Angelucci A, Morando L, Boggio EM, Giustetto M, Pizzorusso T: **Early environmental enrichment moderates the behavioral and synaptic phenotype of MeCP2 null mice.** *Biol Psychiatry* 2010, 67(7):657-665.
13. Henriksen CN, Chaignat E, Reymond A: **Copy number variants, diseases and gene expression.** *Hum Mol Genet* 2009, 18(R1):R1-8.
14. Kalman B, Vitale E: **Structural chromosomal variations in neurological diseases.** *Neurologist* 2009, 15(5):245-253.
15. Niederer HA, Willcocks LC, Rayner TF, Yang W, Lau YL, Williams TN, Scott JA, Urban BC, Peshu N, Dunstan SJ *et al*: **Copy number, linkage disequilibrium and disease association in the FCGR locus.** *Hum Mol Genet* 2010, 19(16):3282-3294.
16. Ahn JW, Mann K, Walsh S, Shehab M, Hoang S, Docherty Z, Mohammed S, Mackie Ogilvie C: **Validation and implementation of array comparative genomic hybridisation as a first line test in place of postnatal karyotyping for genome imbalance.** *Mol Cytogenet* 2010, 3:9.
17. 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(9):949-951.
18. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W *et al*: **Global variation in copy number in the human genome.** *Nature* 2006, 444(7118):444-454.
19. Sampieri K, Meloni I, Scala E, Ariani F, Caselli R, Pescucci C, Longo I, Artuso R, Bruttini M, Mencarelli MA *et al*: **Italian Rett database and biobank.** *Hum Mutat* 2007, 28(4):329-335.

20. Hagberg B, Berg M, Steffenburg U: **Three decades of sociomedical experiences from West Swedish Rett females 4-60 years of age.** *Brain Dev* 2001, 23 Suppl 1:S28-31.
21. Sampieri K, Amenduni M, Papa FT, Katzaki E, Mencarelli MA, Marozza A, Epistolato MC, Toti P, Lazzi S, Bruttini M *et al*: **Array comparative genomic hybridization in retinoma and retinoblastoma tissues.** *Cancer Sci* 2009, 100(3):465-471.
22. Yasui DH, Peddada S, Bieda MC, Vallero RO, Hogart A, Nagarajan RP, Thatcher KN, Farnham PJ, Lasalle JM: **Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes.** *Proc Natl Acad Sci U S A* 2007, 104(49):19416-19421.
23. Oberley MJ, Tsao J, Yau P, Farnham PJ: **High-throughput screening of chromatin immunoprecipitates using CpG-island microarrays.** *Methods Enzymol* 2004, 376:315-334.
24. Yang J, Liu X, Yue G, Adamian M, Bulgakov O, Li T: **Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet.** *J Cell Biol* 2002, 159(3):431-440.
25. Kurtoglu S, Atabek ME, Kumandas S, Keskin M: **Diabetes mellitus type 1: association with Rett syndrome.** *Pediatr Int* 2005, 47(1):90-91.
26. Segawa M: **Early motor disturbances in Rett syndrome and its pathophysiological importance.** *Brain Dev* 2005, 27 Suppl 1:S54-S58.
27. Adegbola AA, Gonzales ML, Chess A, LaSalle JM, Cox GF: **A novel hypomorphic MECP2 point mutation is associated with a neuropsychiatric phenotype.** *Hum Genet* 2009, 124(6):615-623.
28. McClintock TS, Glasser CE, Bose SC, Bergman DA: **Tissue expression patterns identify mouse cilia genes.** *Physiol Genomics* 2008, 32(2):198-206.
29. Badano JL, Mitsuma N, Beales PL, Katsanis N: **The ciliopathies: an emerging class of human genetic disorders.** *Annu Rev Genomics Hum Genet* 2006, 7:125-148.
30. Male DA, Ormsby RJ, Ranganathan S, Giannakis E, Gordon DL: **Complement factor H: sequence analysis of 221 kb of human genomic DNA containing the entire fH, fHR-1 and fHR-3 genes.** *Mol Immunol* 2000, 37(1-2):41-52.
31. Hughes AE, Orr N, Esfandiary H, Diaz-Torres M, Goodship T, Chakravarthy U: **A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration.** *Nat Genet* 2006, 38(10):1173-1177.
32. Raychaudhuri S, Ripke S, Li M, Neale BM, Fagerness J, Reynolds R, Sobrin L, Swaroop A, Abecasis G, Seddon JM *et al*: **Associations of CFHR1-CFHR3 deletion and a CFH SNP to age-related macular degeneration are not independent.** *Nat Genet* 2010, 42(7):553-555; author reply 555-556.
33. Jozsi M, Licht C, Strobel S, Zipfel SL, Richter H, Heinen S, Zipfel PF, Skerka C: **Factor H autoantibodies in atypical hemolytic uremic syndrome correlate with CFHR1/CFHR3 deficiency.** *Blood* 2008, 111(3):1512-1514.
34. Derecki NC, Privman E, Kipnis J: **Rett syndrome and other autism spectrum disorders--brain diseases of immune malfunction?** *Mol Psychiatry* 2010, 15(4):355-363.
35. Stevens B, Allen NJ, Vazquez LE, Howell GR, Christopherson KS, Nouri N, Micheva KD, Mehalow AK, Huberman AD, Stafford B *et al*: **The classical complement cascade mediates CNS synapse elimination.** *Cell* 2007, 131(6):1164-1178.
36. Armstrong DD: **Neuropathology of Rett syndrome.** *J Child Neurol* 2005, 20(9):747-753.
37. Iida N, Kozasa T: **Identification and biochemical analysis of GRIN1 and GRIN2.** *Methods Enzymol* 2004, 390:475-483.
38. Kamiji MM, Inui A: **Neuropeptide y receptor selective ligands in the treatment of obesity.** *Endocr Rev* 2007, 28(6):664-684.
39. Sainsbury A, Schwarzer C, Couzens M, Fetissov S, Furtinger S, Jenkins A, Cox HM, Sperk G, Hokfelt T, Herzog H: **Important role of hypothalamic Y2 receptors in body weight regulation revealed in conditional knockout mice.** *Proc Natl Acad Sci U S A* 2002, 99(13):8938-8943.
40. Sha BY, Yang TL, Zhao LJ, Chen XD, Guo Y, Chen Y, Pan F, Zhang ZX, Dong SS, Xu XH *et al*: **Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population.** *J Hum Genet* 2009, 54(4):199-202.

41. Hu RJ, Lee MP, Connors TD, Johnson LA, Burn TC, Su K, Landes GM, Feinberg AP: **A 2.5-Mb transcript map of a tumor-suppressing subchromosomal transferable fragment from 11p15.5, and isolation and sequence analysis of three novel genes.** *Genomics* 1997, 46(1):9-17.
42. Gertler FB, Niebuhr K, Reinhard M, Wehland J, Soriano P: **Mena, a relative of VASP and Drosophila Enabled, is implicated in the control of microfilament dynamics.** *Cell* 1996, 87(2):227-239.
43. Lanier LM, Gates MA, Witke W, Menzies AS, Wehman AM, Macklis JD, Kwiatkowski D, Soriano P, Gertler FB: **Mena is required for neurulation and commissure formation.** *Neuron* 1999, 22(2):313-325.
44. Kwiatkowski AV, Rubinson DA, Dent EW, Edward van Veen J, Leslie JD, Zhang J, Mebane LM, Philippar U, Pinheiro EM, Burds AA *et al*: **Ena/VASP Is Required for neuritogenesis in the developing cortex.** *Neuron* 2007, 56(3):441-455.
45. Dent EW, Kwiatkowski AV, Mebane LM, Philippar U, Barzik M, Rubinson DA, Gupton S, Van Veen JE, Furman C, Zhang J *et al*: **Filopodia are required for cortical neurite initiation.** *Nat Cell Biol* 2007, 9(12):1347-1359.

4.3 Human Copy Number Variations Database

index - Windows Internet Explorer
 http://www.gvarianti.hondelinux.net/gvarianti/index.php

DATABASE of HUMAN CNVs based on build 36 (hg18)

A tool to compare CNVs identified in MR patients
 hosted by IRCCS OASI Maria SS, Troina

Total CNVs :	13944
N° of patients :	2664
Causative CNVs:	1177

Involved Centers		
Lab Code	Affiliazione	Contatti
Bergamo	Laboratorio Genetica Molecolare - UCCD LGM degli Ospedali Riuniti di Bergamo	Maria Iacone miaacone@ospedaleiruniti.bergamo.it
Fatebenefratelli	Ospedale San Pietro Fatebenefratelli - Centro di Ricerca Scientifica - Unità di Genetica Medica Roma	Viola Alesi viola83a@hotmail.com
Firenze	SCD Diagnostico Genetico - AOU Careggi, Firenze	Francoeca Tomcolli tomcolli@ao-careggi.toscana.it pesuccic@ao-careggi.toscana.it
Genova	S.C. Laboratorio di Genetica Umana E.O. Ospedali Galliera Genova	Michela Malacarne michela.malacarne@galliera.it
Lecco	Laboratorio di Citogenetica, IRCCS Eugenio Medea Lecco	Clara Bonaglio clara.bonaglio@hp.inf.it
Milano	Laboratorio di Citogenetica Medica e Genetica Molecolare, IRCCS, Istituto Auxologico Italiano, Milano.	Daniela Giardino giardino@auxologico.it Lucia Ballarati l.ballarati@auxologico.it
Monza	Laboratorio di Genetica Medica Dipartimento di Neuroscienze e Tecnologie Biomediche Università degli Studi di Milano-Bicocca Via Cadore 48, 20052 Monza, Italia Laboratorio di Genetica Medica Azienda Ospedaliera San Gerardo Via Perugines 33, 20052 Monza, Italia	Sirena Redaelli sirena.redaelli@unimib.it
Padova	Laboratorio Genetica Molecolare Servizio di Genetica Clinica ed Epidemiologica Padova	Monica Forzan monica.forzan@unipd.it Chiara Rigon chiara.rigon@unipd.it
Pavia	Biologia Generale e Genetica Medica, Università di Pavia	Orsotto Zuffardi zuffardi@pavia.it Roberto Ciccone ciccone@unipv.it Annalisa Vetro annalisa.vetro@unipv.it
Pisa	U.O. Citogenetica e Genetica Molecolare AOU Pisana Osp. S. Chiara Pisa	Paolo Simi Angelo Valetto a.valetto@ao-pisa.toscana.it Veronica Bertini v.bertini@ao-pisa.toscana.it
Roma	IRCCS-C.S.S. San Giovanni Rotondo and C.S.S.-Mendel Institute, Rome, Italy	Antonio Novelli anovelli@csc-mendel.it Laura Bernardini l.bernardini@csc-mendel.it
San_G_Rotondo	Servizio di Genetica Medica, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (FG)	Massimo Carella Massimo Carella m.carella@operapadrepio.it
Siena	Genetica Medica, Università di Siena	Alessandra Renieri Alessandra Renieri <renieri@unisi.it> Maria Antonietta Mencarelli mar.mencarelli@ao-siena.toscana.it
Tor_Vergata	U.O.C. Laboratorio di Genetica Medica - Fondazione Policlinico Tor Vergata, Roma	Diana Postorivo dianapostorivo@gmail.com
Torino	SCdU Genetica Medica, AOU San Giovanni Battista di Torino, Via Santena 19, 10126 Torino	Nicola Migone nicola.migone@unito.it
Trieste	S.C. Laboratorio di Genetica Medica Dipartimento di Medicina Molecolare e di Laboratorio IRCCS Istituto per l'Infanzia "Burlo Garofino" Trieste	Yanna Pecile pecile@burlo.trieste.it
Troina	Laboratorio di Diagnosi Genetica e U.O. di Pediatria e Genetica Medica, IRCCS Oasi M. SS, Troina	Marco Fichera mfichera@oasi.en.it
Varese	Laboratorio di Genetica Medica Ospedale di Circolo Fondazione Macchi Viale Borri 21100 Varese	Rosario Casalone r.casalone@ospedale.varese.it
Vicenza	U.O.S. Genetica e Biologia Molecolare Servizio di Immunematologia Trapiantologia e Genetica Umana Azienda ULSS N.6 "VICENZA"	Anna Montaldi genetica@ulssvicenza.it

Chromosome: Start: Stop:

Fig. 3 Homepage of Human Copy Number Variations Database

Molecular karyotype	Inheritance	Proximal breakpoints	Distal breakpoints	Disease genes
dup(1)(p22.3p22.3){0.1 Mb}	M	87 179 292	87 273 826	
dup(1)(p12p12){0.2 Mb}	P	118 294 990	118 303 853	
dup(1)(q43q43){0.3 Mb}	P	234 698 142	234 814 787	EDARADD
dup(2)(p16.3p16.3){0.3 Mb}	P	47 866 371	47 938 504	MSH6
dup(2)(q13q13){1.33 Mb}	P, M	110 199 004	110 337 631	NPHP1*
dup(2)(q14.2q14.2){0.6 Mb}	P	119 843 354	120 283 921	
del(2)(q32.2q32.2){0.13 Mb}	M	189 573 876	189 633 678	COL3A1, COL5A2
dup(2)(q36.3q36.3){1.3 Mb}	P	229 638 936	230 533 004	
dup(3)(p22.1p22.1){0.5 Mb}	M	41 729 441	41 940 304	
dup(3)(p21.1p14.3){0.7 Mb}	Brother	53 736 895	54 395 857	
dup(3)(p14.2p14.2){0.8 Mb}	M	59 931 959	60 267 321	
dup(4)(p16.2p16.3){0.23 Mb}	M	2 902 096	3 114 307	HTT
dup(4)(p16.1p16.1){0.06 Mb}	P	9 686 502	9 751 987	
del(4)(q13.3q13.3){0.16 Mb}	P	74 105 031	74 245 908	
dup(4)(q13.3q13.3){1.2 Mb}	P	75 458 705	76 190 527	
dup(4)(q32.2q32.2){1.1 Mb}	P	162 406 396	162 900 008	
dup(4)(q34.1q34.1){1.1 Mb}	M	174 912 401	175 910 342	HPGD
del(4)(q35.1q35.2){1.2 Mb}	P	186 466 884	187 446 875	TLR3, CYP4V2, KLKB1, F11
del(4)(q35.2q35.2){1.4 Mb}	P	189 846 235	191 133 668	
dup(5)(p13.2p13.2){0.4 Mb}	M	37 327 849	37 552 419	
dup(5)(q21.1q21.1){0.2 Mb}	P	100 219 716	100 401 892	
del(6)(p24.3p24.3){0.2 Mb}	M	9 953 753	10 085 727	
del(7)(q31.33q31.33){0.1 Mb}	M	124 394 469	124 482 617	
dup(8)(q12.1q12.1){1.0 Mb}	P	56 590 647	57 085 155	
del(9)(p24.1p24.1){0.14 Mb}	M	6 318 511	6 321 081	
dup(9)(p23p23){0.60 Mb}	M	11 447 340	11 687 635	
dup(9)(q22.32q22.32){1.0 Mb}	M	97 470 882	98 104 136	HSD17B3
dup(10)(q21.3q21.3){0.5 Mb}	P	69 661 546	70 076 154	
del(11)(p14.2p14.3){0.8 Mb}	P	26 245 417	26 577 744	
dup(11)(p11.2p11.2){0.08 Mb}	P	44 108 216	44 184 944	EXT2**
del(11)(q22.3q22.3){0.12 Mb}	M	107 761 839	107 807 671	
dup(12)(p12.1p12.1){0.22 Mb}	Sister	22 369 323	22 586 257	
dup(12)(p11.22p11.23){0.6 Mb}	M	27 270 178	27 659 659	
dup(12)(p11.22p11.22){0.4 Mb}	M	29 489 556	29 844 564	
dup(12)(q21.31q21.31){0.8 Mb}	P	81 274 846	81 814 232	
dup(12)(q21.31q21.32){1.6 Mb}	P	84 474 053	86 008 465	
dup(12)(q24.11q24.11){0.1 Mb}	M	108 061 580	108 123 789	
dup(13)(q13.2q13.3){0.8 Mb}	P	34 517 499	35 022 728	
dup(13)(q21.33q22.1){0.2 Mb}	M	72 153 664	72 228 260	

dup(13)(q31.1q31.1){0.3 Mb}	P	80 241 336	80 460 454	
del(13)(q31.3q32.1){0.5 Mb}	P	93 853 568	93 890 089	
del(14)(q21.1){0.4 Mb}	P	40 088 478	40 380 681	
dup(15)(q24.1q24.1){0.3 Mb}	M	72 398 300	72 587 604	SEMA7A
dup(15)(q26.3q26.3){2.9 Mb}	P	96 430 271	99 137 984	IGF1R, MEF2A
del(16)(p13.2p13.3){0.6 Mb}	M	6 367 765	6 368 498	
del(16)(p12.2p12.1){0.3 Mb}	P	21 507 188	21 745 052	OTOA
dup(16)(q23.2q23.2){0.55 Mb}	M	79 075 510	79 361 411	
dup(16)(q23.3q23.3){1.63 Mb}	M	80 773 415	82 193 537	
dup(16)(q24.3q24.3) {0.08 Mb}	P	88 376 785	88 436 861	FANCA
del(17)(q12q12){0.2 Mb}	P	30 711 469	30 762 580	
dup(17)(q12q12){1.8 Mb}	P	31 925 650	33 726 757	ACACA, HNF1B
dup(19)(q13.42q13.42){0.16 Mb}	P	59 446 274	59 537 113	
del(20)(p12.1p12.1){0.44 Mb}	M	14 772 372	15 216 002	
del(21)(q22.3q22.3){0.12 Mb}	P	44 701 761	44 825 966	

P: inherited from father

M: inherited from mother

*nephronoptosis region

** Potocki-Shaffer syndrome region

*** Autism susceptibility region

Table 3a. Autosomal inherited private rearrangements found in 330 patients. The base pair positions are according to UCSC Genome Browser, <http://genome.ucsc.edu>, on Human May 2006 Assembly.

Molecular karyotype	Patient sex	Inheritance	Proximal breakpoints	Distal breakpoints
dup(X)(p22.31p22.31){0.8 Mb}	F	P	8 458 107	9 042 705
dup(X)(p11.4p11.4){0.2 Mb}	F	P	38 376 483	38 432 836
dup(X)(q22.3q22.3){0.43 Mb}	M	M	106 238 368	106 515 657
del(X)(q25q25){3.8 Mb}	M	M	124 267 011	127 627 617
dup(X)(q26.2q26.2){0.6 Mb}	F	P	130 526 525	130 788 298
dup(X)(q27.2q27.2){0.21 Mb}	M	M	140 617 007	140 824 390
dup(X)(q28q28){0.4 Mb}	F	P	148 880 115	149 137 402

Table 3b. Inherited private rearrangements on the X chromosome found in 330 patients.

Molecular karyotype	Disease genes	number	Frequency %
dup(2)(q13q13){1.33 Mb}	NPHP1	3	2.14%
del(2)(q32.2q32.2){0.13 Mb}	COL3A1, COL5A2	1	0.71%
dup(3)(p22.1p22.1){0.5 Mb}		1	0.71%
dup(4)(p16.1p16.1){0.06 Mb}		2	1.42%
dup(9)(p23p23){0.60 Mb}		1	0.71%
dup(13)(q13.2q13.3){0.8 Mb}		2	1.42%

Table 4a. Autosomal private rearrangements found in 140 controls.

Molecular karyotype	Disease genes	number	Frequency %
dup(X)(p22.31p22.31){0.8 Mb}	KAL1	1	0.71%

Table 4b. Inherited private rearrangements on the X chromosome found in 140 controls.

5. DISCUSSION, CONCLUSIONS and FUTURE PERSPECTIVES

5.1 DISCUSSION

Microarray-based comparative genomic hybridization has revolutionized clinical cytogenetic, as it provides a relatively quick method to scan the genome for gains and losses of chromosomal material with significantly higher resolution and greater clinical yield than was previously possible. Numerous different aCGH platforms have emerged and have been used successfully in the diagnostic setting. In the past few years, these new methodologies have led to the identification of novel genomic disorders in patients with developmental delay/intellectual disability and/or multiple congenital anomalies as well as to the discovery that each individual carries inherited copy number variations whose contributions to genetic variation, and complex disease are not well understood. Although aCGH is currently being used as an adjunct test for to standard karyotype analysis, it may to become the genetic test of choice, especially in cases of idiopathic ID/MCA. [4]

As such y, in this work this innovative technique was employed to study a group of MCA/ID patients. In all cases where the aCGH analysis indicated that the rearrangements arisen "*de novo*", we performed standard karyotype analysis of both parents in order to exclude the presence of a balanced rearrangement favouring the unbalanced rearrangement in the child. In all cases but two the parents' karyotype resulted normal and during genetic counselling a low recurrence risk was given to the family.

Among 696 patients, we found 426 (61%) negative while in 88 patients (13%) the analysis is still ongoing. We identified an inherited or *de novo* rearrangement in 165 cases (24%) while in 17 cases (2%) we detected novel *de novo* deletions not reported in the literature (Fig. 4).

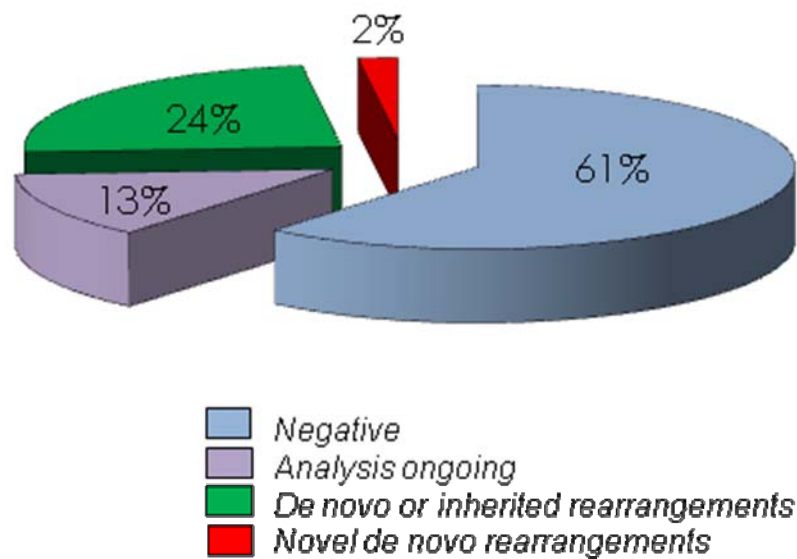


Fig. 4 | Number and type of rearrangements identified by oligo aCGH in a cohort of 696 MCA/ID patients.

Among 165 positive cases 10% were novel *de novo* rearrangements, 15% were rearrangements of known syndromes, 40% were non polymorphic rearrangements inherited from mother, 35% were non polymorphic rearrangements inherited from father.

The strategy for the identification of the candidate gene responsible for the clinical features is to analyze the genic content of rearrangements. In the patient with a *de novo* 3 Mb interstitial deletion of chromosome 14q12 this approach allowed us the identification of a new gene: *FOXG1B*. Later *FOXG1B* was demonstrated to be a gene responsible for the most severe form of RTT, the congenital variant. [97] The patient showed dysmorphic features and a Rett-like clinical course, including normal perinatal period, postnatal microcephaly, seizures, and severe intellectual disability. The deleted region was gene poor, indeed it contained only five genes. Among them, *FOXG1* turned out to be a very interesting gene because it encodes a brain-specific transcriptional repressor.

A first analysis of this gene with a combination of both DHPLC and real-time quantitative PCR in a cohort of 53 *MECP2/CDKL5* mutation-negative RTT patients, allowed the characterization of a different *de novo FOXG1* truncating mutation in two RTT congenital variant patients. Soon after, *FOXG1* mutations have been identified in several patients, independently classified as congenital Rett variants from France, Spain, Latvia and Canada. [98]

The search of additional patients having overlapping 14q12 deletions in the existing databases, (ECARUCA and DECIPHER) which collect cases with chromosomal aberrations, allowed the identification, in the DECIPHER database, of two new patients and the further characterization of the 14q12 microdeletion syndrome. [99]

Nevertheless, when attempting to carry out a genotype-phenotype correlation after the identification of a certain submicroscopic anomaly, the classic approach of disease mapping based on the sole genic content may not be sufficient.

Autism spectrum disorders are a group of common neurodevelopmental conditions characterised by impairments in communication, social interaction, and behaviour. At least one of every 166 children is likely to be affected by this spectrum of disorders, with a male: female (M:F) ratio of 4:1, and it is the most heritable of all complex neuropsychiatric conditions. Since ASDs are considered complex genetic disorders, resulting from the interaction of several genes and environmental factors, the lumping together of all cases of ASD, with no subgrouping based on phenotypic characteristics, makes the identification of contributory genes extremely difficult. The fact that autism is known to be associated with several distinct medical/genetic disorders further highlights its genetic heterogeneity. A variety of different approaches to identify genes for ASDs have been undertaken, including cytogenetic assessment for chromosome

abnormalities, genome scan linkage studies, and association studies. The overlap of results among different study methods is limited, largely attributed to the significant clinical and genetic heterogeneity of ASDs based on variable behavioural indices, variations among study populations, and limitation of methods for detecting ASD susceptibility genes of mild to moderate effect. [100]

We selected 95 unrelated patients with ASD. In addition to the well known recurrent rearrangements involving the 15q11q13, 16p13 and 22q13 regions, recurrent microdeletions and microduplications at 16p11.2 have been recently identified and have been shown to confer susceptibility to ASDs in up to 1% of autistic patients. In fifteen patients (16%) CNVs were considered as disease associated, while in 20 patients (21%) CNVs were classified as uncertain cofactor. Our findings have important clinical and research implications. These results emphasise that aCGH analysis is able to detect genetic cause in ASD patients, stimulating the need for further investigation of CNVs in ASD. These findings underscore a challenge in genetic counselling in interpreting aCGH data and give practical advices about the management of such data and the communication process to the family.

The presence of a CNVs in a coding region usually correlates with changes in the abundance of corresponding transcripts. Absence or excess of the protein product of a dosage sensitive gene may influence cell differentiation or migration and tissue formation early during development. Excess of a protein produced as a result of gene amplification may also lead to protein misfolding in an age-dependent manner and overload intracellular pathways of protein transport, proteosome degradation, or recycling commonly seen in late onset neurodegenerative disorders. Intracellular aggregation of proteins and inclusions

may affect protein-protein interactions, enhance free radical formation, cause mitochondrial damage and dysfunction, and initiate apoptosis

In addition, genomic rearrangements may also be associated with molecular mechanisms other than affecting transcript levels to influence gene dosage and expression. Such complex mechanisms include gene interruption, gene fusion, unmasking a recessive allele or silenced gene, and interruption of regulatory gene-gene and chromosomal interactions. The investigations of the roles of CNVs during evolution and their biologic significance in health and diseases are rapidly progressing. Even before the completion of the Human Genome Project, the pathogenic significance of gene dosage was realized in several disorders of the central and peripheral nervous system. The highlighted examples demonstrate how gene dosage effects may influence cell function and development of common disorders often characterized by heterogeneous genetic etiology. [101] To test the hypothesis that CNVs can modulate the phenotypic expression of RTT syndrome we analyzed by aCGH two couples of RTT sisters and four additional couples of unrelated RTT girls matched by mutation type showing discordant phenotype (classic and Z-RTT). Results of CNVs analysis of phenotypically discordant RTT pairs are shown in Tables 1 and 2. Overall, we indentified 28 CNVs, 27 of them being known polymorphic regions and one being an apparently private rearrangement duplicated in only one Z-TT patient. This private CNV localized to 3q13.12, extended for about 99 Kb and contained three genes expressed in retina (*GUCA1C*), testis (*MORC1*) and one putative uncharacterized protein C3orf66.

Some polymorphic CNVs were found in more than one pair of phenotypically discordant RTT pairs. Among the 29 CNVs we considered 14 of them as "unlikely modifier regions" since results were apparently not associated with phenotypic severity (Table 2). For instance, the CNV was observed in both a classic and a Z-RTT member of different pairs. These include regions containing

olfactory receptors and class II HLA molecules that are not expected to directly correlate with the phenotypic variability related to classic/Z-RTT phenotype. The remaining 15 CNVs were considered potential modifiers (Table 1). For 3 of them results were consistent in two couples. Analyses on phenotypically discordant RTT pairs resulted in 28 CNV's that gave a list of 55 genes which could potentially modify RTT phenotype. A Perl script was used to search for these 55 genes within two experimentally replicated promoter hit files that ranked MeCP2 binding based on ChIP-chip analysis. Aliases of genes were acquired from genome.ucsc.edu and genelist.org. However, twenty genes were not found on the promoter hit list primarily because most these genes were discovered after the array was designed. The ranking from total number of genes from 1 to 24,134 is shown for two replicate MeCP2-ChIP microarrays (MeCP2 B and MeCP2 C promoter hits rank, Tables 1 and 2). Genes with promoters in the top 10% of MeCP2 promoter hits for at least one replicate are shaded in blue.

In conclusion, our study failed to identify one major common modifier gene/region suggesting that genetic modifiers may be complex and different case by case. However, we identified several genes within loci that may have a role in modulating the diseases in single cases. Environmental and epigenetic differences may further modify RTT syndrome phenotype and this study suggests that investigating genetic modifiers of the RTT phenotype by CNV analyses is by nature complex because MeCP2 pathways and gene targets are influenced by both genetic and epigenetic variables.

Microarray-based comparative genomic hybridisation enables the accurate detection of submicroscopic CNVs and has been increasingly used to investigate patients with intellectual disability or congenital abnormalities in some clinical settings. In aCGH analysis the CNVs have been identified using the aberration

algorithms Z-Score and ADM-1 provided by the CGH-Analytics software. The presence/absence of CNVs depends both on the experimental procedure and the aberration algorithm used to this purpose. The intrinsic variability of these procedures does inevitably lead to false negative and false positive calls. In order to reduce the false positive results, we disregarded all the calls containing less than three contiguous probes for experiments of excellent quality. We adopted some criteria to discern between benign and pathogenic CNVs. Firstly, we considered benign, and thus disregarded, all the CNVs reported in the last release of DGVs. [102] Since this catalogue contains many benign CNVs detected by BAC arrays, which in general are overestimated in size, [68] we compared our CNVs only with those reported in the database detected by oligo/SNP arrays. Then, we considered as clinically relevant all the CNVs arisen *de novo*, the imbalances known to be pathogenic and the CNVs encompassing genes whose up/downregulation is known to cause a disease.

Until August 2009 we selected of a cohort of 330 patients with MCA/ID in which we identified 61 CNVs, ranging from 0.06 to 2.9 Mb, that were not described in the DGVs (Tab. 3a and 3b). Among these 14 of which are regions that include disease genes, 44 are gains and 17 are losses. In all cases the same rearrangement was inherited from a healthy parent. Fifty nine of the rearrangements were private while 2 were found in more patients. In particular a del17q12 was found in 3 patients and a del9p24.1 was found in 2 patients. We collected 140 subjects from the first year of Medicine and Surgery faculty at University of Siena and from IRCCS OASI Maria SS. Troina and Biologia Generale e Genetica Medica, University of Pavia to detect the frequency of selected CNVs in normal population. In the 140 control subjects analyzed we identified only 7 out of 61 rearrangements previously described (11.4%) (Tab. 4a and 4b). In particular a dup2q13 of 1.33 Mb was identified in 2.14% of control subjects; a dup4p16.1 of 0.06 Mb and a dup13q13 of 0.8 Mb were identified in

1.4% of control subjects; a del2q32.2 of 0.3 Mb, a dup3p22.1 of 0.5 Mb and a dup9p23 of 0.6 Mb were identified in 0.7% of control subjects. Moreover, none of the 2 rearrangements identified in more than one patient (del17q12q12 and del9p24.1) was detected until known in the normal population. Each of the 7 rearrangements found in the control population has the same extension of those previously detected in the 330 patients. By analyzing 140 control subjects we found that 13.6% of the duplications and 6.2% of the deletions detected were recurrent in the control population. The dup2q13 seems to be the most frequent rearrangement detected so far. Three of the 7 recurrent rearrangements are regions that include known disease genes. One of these 3 rearrangements is included in region known as a susceptibility locus in autism (dupXp22.31).

The project of "Genomic Structural Variation Studies in Mentally Retarded and Normal Individuals in Italy" was born from collaboration with other two genetic units to characterize Italian CNVs not reported in DGVs. Currently, after 2 years, the total of Italian genetic units that decided to share aCGH data increased to 19. (Fig.3) In HCNV database about 14,000 CNVs are collected, so far.

The HCNV database resulted to be an helpful tool. In fact increasing the CNVs detected by aCGH, the clinician will have more data to explain not polymorphic but inherited or *de novo* rearrangements overlapping with other patients of database. Our knowledge is still limited, you can not accurately predict the clinical syndrome based on genetic backgrounds, you can only predict a possible variation.

5.2 CONCLUSIONS AND FUTURE PERSPECTIVES

aCGH can detect cytogenetic anomalies at a resolution unachievable by conventional techniques. Imbalances as small as the size of the used clones will identify new syndromes and will elucidate the molecular basis of clinically recognized syndromes. Given the great achievements in the past few years, aCGH will change the diagnostic approach to many congenital and acquired genetic diseases such as intellectual disability, birth defects, and cancer. However, diagnostic guidelines should be established though in order to enable the reliable and accurate detection of chromosome imbalances by aCGH in the cytogenetic laboratory. [103]

Our study allowed the characterization of several chromosomal imbalances in patients with complex phenotype, confirming the power of the aCGH method to clarify the molecular basis of these difficult cases. Through the employment of this innovative approach, several families finally received a definitive diagnosis and a correct recurrence risk.

For the future we plan to continue the consultation of the literature to find new disease candidate genes and new emerging low penetrance syndromes. We plan to reconsider the rearrangements found in our cohort and to provide to our patients a correct clinical follow up. We will continue to re-analyze our cohort paying attention to the CNVs regions and taking into consideration the possibility of modifier genes.

6. REFERENCES

REFERENCES

1. Rosenberg, C., et al., *Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents*. J Med Genet, 2006. **43**(2): p. 180-6.
2. Faas, B.H., et al., *Detection of cryptic subtelomeric imbalances in fetuses with ultrasound abnormalities*. Eur J Med Genet, 2008. **51**(6): p. 511-9.
3. Trask, B.J., *Fluorescence in situ hybridization: applications in cytogenetics and gene mapping*. Trends Genet, 1991. **7**(5): p. 149-54.
4. Edelmann, L. and K. Hirschhorn, *Clinical utility of array CGH for the detection of chromosomal imbalances associated with mental retardation and multiple congenital anomalies*. Ann N Y Acad Sci, 2009. **1151**: p. 157-66.
5. Solinas-Toldo, S., et al., *Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances*. Genes Chromosomes Cancer, 1997. **20**(4): p. 399-407.
6. Pinkel, D., et al., *High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays*. Nat Genet, 1998. **20**(2): p. 207-11.
7. Cai, W.W., et al., *Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays*. Nat Biotechnol, 2002. **20**(4): p. 393-6.
8. Ylstra, B., et al., *BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH)*. Nucleic Acids Res, 2006. **34**(2): p. 445-50.
9. Ishkanian, A.S., et al., *A tiling resolution DNA microarray with complete coverage of the human genome*. Nat Genet, 2004. **36**(3): p. 299-303.
10. Shaikh, T.H., *Oligonucleotide arrays for high-resolution analysis of copy number alteration in mental retardation/multiple congenital anomalies*. Genet Med, 2007. **9**(9): p. 617-25.
11. Bejjani, B.A., et al., *Use of targeted array-based CGH for the clinical diagnosis of chromosomal imbalance: is less more?* Am J Med Genet A, 2005. **134**(3): p. 259-67.
12. Bejjani, B.A. and L.G. Shaffer, *Application of array-based comparative genomic hybridization to clinical diagnostics*. J Mol Diagn, 2006. **8**(5): p. 528-33.
13. Shaffer, L.G., *Risk estimates for uniparental disomy following prenatal detection of a nonhomologous Robertsonian translocation*. Prenat Diagn, 2006. **26**(4): p. 303-7.
14. Le Caignec, C., et al., *Detection of genomic imbalances by array based comparative genomic hybridisation in fetuses with multiple malformations*. J Med Genet, 2005. **42**(2): p. 121-8.
15. Sahoo, T., et al., *Microarray based comparative genomic hybridization testing in deletion bearing patients with Angelman syndrome: genotype-phenotype correlations*. J Med Genet, 2006. **43**(6): p. 512-6.
16. Kitsiou-Tzeli, S., et al., *Prenatal diagnosis of a de novo partial trisomy 10p12.1-12.2 pter originating from an unbalanced translocation onto 15qter and confirmed with array CGH*. Prenat Diagn, 2008. **28**(8): p. 770-2.
17. Veltman, J.A. and B.B. de Vries, *Diagnostic genome profiling: unbiased whole genome or targeted analysis?* J Mol Diagn, 2006. **8**(5): p. 534-7; discussion 537-9.
18. Toruner, G.A., et al., *An oligonucleotide based array-CGH system for detection of genome wide copy number changes including subtelomeric regions for genetic evaluation of mental retardation*. Am J Med Genet A, 2007. **143A**(8): p. 824-9.

19. Baldwin, E.L., et al., *Enhanced detection of clinically relevant genomic imbalances using a targeted plus whole genome oligonucleotide microarray*. Genet Med, 2008. **10**(6): p. 415-29.
20. Veltman, J.A. and B.B. de Vries, *Whole-genome array comparative genome hybridization: the preferred diagnostic choice in postnatal clinical cytogenetics*. J Mol Diagn, 2007. **9**(2): p. 277.
21. Selzer, R.R., et al., *Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH*. Genes Chromosomes Cancer, 2005. **44**(3): p. 305-19.
22. Urban, A.E., et al., *High-resolution mapping of DNA copy alterations in human chromosome 22 using high-density tiling oligonucleotide arrays*. Proc Natl Acad Sci U S A, 2006. **103**(12): p. 4534-9.
23. Wong, L.J., et al., *Utility of oligonucleotide array-based comparative genomic hybridization for detection of target gene deletions*. Clin Chem, 2008. **54**(7): p. 1141-8.
24. Ahn, J.W., et al., *Validation and implementation of array comparative genomic hybridisation as a first line test in place of postnatal karyotyping for genome imbalance*. Mol Cytogenet, 2010. **3**: p. 9.
25. Shaffer, L.G. and T.H. Bui, *Molecular cytogenetic and rapid aneuploidy detection methods in prenatal diagnosis*. Am J Med Genet C Semin Med Genet, 2007. **145C**(1): p. 87-98.
26. Shevell, M.I., et al., *Array comparative genomic hybridization in global developmental delay*. Am J Med Genet B Neuropsychiatr Genet, 2008. **147B**(7): p. 1101-8.
27. Vermeesch, J.R., et al., *Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis*. J Histochem Cytochem, 2005. **53**(3): p. 413-22.
28. Cytrynbaum, C.S., et al., *Advances in overgrowth syndromes: clinical classification to molecular delineation in Sotos syndrome and Beckwith-Wiedemann syndrome*. Curr Opin Pediatr, 2005. **17**(6): p. 740-6.
29. Gropman, A.L., et al., *New developments in Smith-Magenis syndrome (del 17p11.2)*. Curr Opin Neurol, 2007. **20**(2): p. 125-34.
30. Ballif, B.C., et al., *Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication*. Mol Cytogenet, 2008. **1**: p. 8.
31. Ballif, B.C., et al., *Discovery of a previously unrecognized microdeletion syndrome of 16p11.2-p12.2*. Nat Genet, 2007. **39**(9): p. 1071-3.
32. Ballif, B.C., et al., *Identification of a previously unrecognized microdeletion syndrome of 16q11.2q12.2*. Clin Genet, 2008. **74**(5): p. 469-75.
33. Koolen, D.A., et al., *A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism*. Nat Genet, 2006. **38**(9): p. 999-1001.
34. Sharp, A.J., et al., *Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome*. Nat Genet, 2006. **38**(9): p. 1038-42.
35. Sharp, A.J., et al., *Characterization of a recurrent 15q24 microdeletion syndrome*. Hum Mol Genet, 2007. **16**(5): p. 567-72.
36. Vissers, L.E., et al., *Mutations in a new member of the chromodomain gene family cause CHARGE syndrome*. Nat Genet, 2004. **36**(9): p. 955-7.
37. Bejjani, B.A. and L.G. Shaffer, *Clinical utility of contemporary molecular cytogenetics*. Annu Rev Genomics Hum Genet, 2008. **9**: p. 71-86.
38. Neill, N.J., et al., *Comparative analysis of copy number detection by whole-genome BAC and oligonucleotide array CGH*. Mol Cytogenet, 2010. **3**: p. 11.

39. Krepischi-Santos, A.C., et al., *Whole-genome array-CGH screening in undiagnosed syndromic patients: old syndromes revisited and new alterations*. *Cytogenet Genome Res*, 2006. **115**(3-4): p. 254-61.
40. Firth, H.V., et al., *DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources*. *Am J Hum Genet*, 2009. **84**(4): p. 524-33.
41. Gu, W., F. Zhang, and J.R. Lupski, *Mechanisms for human genomic rearrangements*. *Pathogenetics*, 2008. **1**(1): p. 4.
42. Turner, D.J., et al., *Germline rates of de novo meiotic deletions and duplications causing several genomic disorders*. *Nat Genet*, 2008. **40**(1): p. 90-5.
43. Berg, J.S., et al., *Speech delay and autism spectrum behaviors are frequently associated with duplication of the 7q11.23 Williams-Beuren syndrome region*. *Genet Med*, 2007. **9**(7): p. 427-41.
44. Potocki, L., et al., *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*, 2007. **80**(4): p. 633-49.
45. Yobb, T.M., et al., *Microduplication and triplication of 22q11.2: a highly variable syndrome*. *Am J Hum Genet*, 2005. **76**(5): p. 865-76.
46. Brunetti-Pierri, N., et al., *Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities*. *Nat Genet*, 2008. **40**(12): p. 1466-71.
47. Van der Aa, N., et al., *Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome*. *Eur J Med Genet*, 2009. **52**(2-3): p. 94-100.
48. Hased, S.J., et al., *A new genomic duplication syndrome complementary to the velocardiofacial (22q11 deletion) syndrome*. *Clin Genet*, 2004. **65**(5): p. 400-4.
49. Potocki, L., et al., *Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion*. *Nat Genet*, 2000. **24**(1): p. 84-7.
50. de Vries, B.B., et al., *Diagnostic genome profiling in mental retardation*. *Am J Hum Genet*, 2005. **77**(4): p. 606-16.
51. Engels, H., et al., *A novel microdeletion syndrome involving 5q14.3-q15: clinical and molecular cytogenetic characterization of three patients*. *Eur J Hum Genet*, 2009. **17**(12): p. 1592-9.
52. Ogilvie, C.M., et al., *A novel deletion in proximal 22q associated with cardiac septal defects and microcephaly: a case report*. *Mol Cytogenet*, 2009. **2**: p. 9.
53. Pescucci, C., et al., *2q24-q31 deletion: report of a case and review of the literature*. *Eur J Med Genet*, 2007. **50**(1): p. 21-32.
54. Mencarelli, M.A., et al., *Clinical and molecular characterization of a patient with a 2q31.2-32.3 deletion identified by array-CGH*. *Am J Med Genet A*, 2007. **143A**(8): p. 858-65.
55. Vissers, L.E., et al., *Identification of disease genes by whole genome CGH arrays*. *Hum Mol Genet*, 2005. **14 Spec No. 2**: p. R215-23.
56. Bielanska, M., S.L. Tan, and A. Ao, *Chromosomal mosaicism throughout human preimplantation development in vitro: incidence, type, and relevance to embryo outcome*. *Hum Reprod*, 2002. **17**(2): p. 413-9.
57. Hassold, T., *Mosaic trisomies in human spontaneous abortions*. *Hum Genet*, 1982. **61**(1): p. 31-5.
58. Ledbetter, D.H., et al., *Cytogenetic results from the U.S. Collaborative Study on CVS*. *Prenat Diagn*, 1992. **12**(5): p. 317-45.
59. Wang, B.B., C.H. Rubin, and J. Williams, 3rd, *Mosaicism in chorionic villus sampling: an analysis of incidence and chromosomes involved in 2612 consecutive cases*. *Prenat Diagn*, 1993. **13**(3): p. 179-90.

60. Ballif, B.C., et al., *Detection of low-level mosaicism by array CGH in routine diagnostic specimens*. Am J Med Genet A, 2006. **140**(24): p. 2757-67.
61. Cheung, S.W., et al., *Microarray-based CGH detects chromosomal mosaicism not revealed by conventional cytogenetics*. Am J Med Genet A, 2007. **143A**(15): p. 1679-86.
62. Sebat, J., et al., *Strong association of de novo copy number mutations with autism*. Science, 2007. **316**(5823): p. 445-9.
63. Marshall, C.R., et al., *Structural variation of chromosomes in autism spectrum disorder*. Am J Hum Genet, 2008. **82**(2): p. 477-88.
64. Walsh, T., et al., *Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia*. Science, 2008. **320**(5875): p. 539-43.
65. Stefansson, H., et al., *Large recurrent microdeletions associated with schizophrenia*. Nature, 2008. **455**(7210): p. 232-6.
66. Wagenstaller, J., et al., *Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation*. Am J Hum Genet, 2007. **81**(4): p. 768-79.
67. Friedman, J.M., et al., *Oligonucleotide microarray analysis of genomic imbalance in children with mental retardation*. Am J Hum Genet, 2006. **79**(3): p. 500-13.
68. Redon, R., et al., *Global variation in copy number in the human genome*. Nature, 2006. **444**(7118): p. 444-54.
69. Pinto, D., et al., *Copy-number variation in control population cohorts*. Hum Mol Genet, 2007. **16 Spec No. 2**: p. R168-73.
70. Levy, S., et al., *The diploid genome sequence of an individual human*. PLoS Biol, 2007. **5**(10): p. e254.
71. Perry, G.H., et al., *Copy number variation and evolution in humans and chimpanzees*. Genome Res, 2008. **18**(11): p. 1698-710.
72. Korb, J.O., et al., *Paired-end mapping reveals extensive structural variation in the human genome*. Science, 2007. **318**(5849): p. 420-6.
73. Koolen, D.A., et al., *Genomic microarrays in mental retardation: a practical workflow for diagnostic applications*. Hum Mutat, 2009. **30**(3): p. 283-92.
74. Lee, C., A.J. Iafrate, and A.R. Brothman, *Copy number variations and clinical cytogenetic diagnosis of constitutional disorders*. Nat Genet, 2007. **39**(7 Suppl): p. S48-54.
75. McMullan, D.J., et al., *Molecular karyotyping of patients with unexplained mental retardation by SNP arrays: a multicenter study*. Hum Mutat, 2009. **30**(7): p. 1082-92.
76. McCarroll, S.A., et al., *Integrated detection and population-genetic analysis of SNPs and copy number variation*. Nat Genet, 2008. **40**(10): p. 1166-74.
77. Cheung, J., et al., *Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence*. Genome Biol, 2003. **4**(4): p. R25.
78. Hehir-Kwa, J.Y., et al., *Accurate distinction of pathogenic from benign CNVs in mental retardation*. PLoS Comput Biol, 2010. **6**(4): p. e1000752.
79. Lupski, J.R., *Genomic rearrangements and sporadic disease*. Nat Genet, 2007. **39**(7 Suppl): p. S43-7.
80. Lupski, J.R., *New mutations and intellectual function*. Nat Genet, 2010. **42**(12): p. 1036-8.
81. Henshall, J.M., V.A. Whan, and B.J. Norris, *Reconstructing CNV genotypes using segregation analysis: combining pedigree information with CNV assay*. Genet Sel Evol, 2010. **42**: p. 34.
82. Beckmann, J.S., X. Estivill, and S.E. Antonarakis, *Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability*. Nat Rev Genet, 2007. **8**(8): p. 639-46.

83. Chakrabarti, S. and E. Fombonne, *Pervasive developmental disorders in preschool children: confirmation of high prevalence*. Am J Psychiatry, 2005. **162**(6): p. 1133-41.
84. Christian, S.L., et al., *Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder*. Biol Psychiatry, 2008. **63**(12): p. 1111-7.
85. Bailey, A., et al., *Autism as a strongly genetic disorder: evidence from a British twin study*. Psychol Med, 1995. **25**(1): p. 63-77.
86. Hurley, R.S., et al., *The broad autism phenotype questionnaire*. J Autism Dev Disord, 2007. **37**(9): p. 1679-90.
87. Constantino, J.N. and R.D. Todd, *Intergenerational transmission of subthreshold autistic traits in the general population*. Biol Psychiatry, 2005. **57**(6): p. 655-60.
88. Veenstra-VanderWeele, J. and E.H. Cook, Jr., *Molecular genetics of autism spectrum disorder*. Mol Psychiatry, 2004. **9**(9): p. 819-32.
89. Abrahams, B.S. and D.H. Geschwind, *Advances in autism genetics: on the threshold of a new neurobiology*. Nat Rev Genet, 2008. **9**(5): p. 341-55.
90. Ghahramani Seno, M.M., et al., *Gene and miRNA expression profiles in autism spectrum disorders*. Brain Res, 2010.
91. Wetmore, D.Z. and C.C. Garner, *Emerging pharmacotherapies for neurodevelopmental disorders*. J Dev Behav Pediatr, 2010. **31**(7): p. 564-81.
92. Mencarelli, M.A., et al., *Novel FOXP1 mutations associated with the congenital variant of Rett syndrome*. J Med Genet, 2010. **47**(1): p. 49-53.
93. Gropman, A.L. and M.L. Batshaw, *Epigenetics, copy number variation, and other molecular mechanisms underlying neurodevelopmental disabilities: new insights and diagnostic approaches*. J Dev Behav Pediatr, 2010. **31**(7): p. 582-91.
94. Sambrook, J. and M.J. Gething, *Protein structure. Chaperones, paperones*. Nature, 1989. **342**(6247): p. 224-5.
95. Ariani, F., et al., *Real-time quantitative PCR as a routine method for screening large rearrangements in Rett syndrome: Report of one case of MECP2 deletion and one case of MECP2 duplication*. Hum Mutat, 2004. **24**(2): p. 172-7.
96. Koolen, D.A., et al., *Screening for subtelomeric rearrangements in 210 patients with unexplained mental retardation using multiplex ligation dependent probe amplification (MLPA)*. J Med Genet, 2004. **41**(12): p. 892-9.
97. Ariani, F., et al., *FOXP1 is responsible for the congenital variant of Rett syndrome*. Am J Hum Genet, 2008. **83**(1): p. 89-93.
98. Jacob, F.D., et al., *Atypical Rett syndrome with selective FOXP1 deletion detected by comparative genomic hybridization: case report and review of literature*. Eur J Hum Genet, 2009. **17**(12): p. 1577-81.
99. Mencarelli, M.A., et al., *14q12 Microdeletion syndrome and congenital variant of Rett syndrome*. Eur J Med Genet, 2009. **52**(2-3): p. 148-52.
100. Qiao, Y., et al., *Phenomic determinants of genomic variation in autism spectrum disorders*. J Med Genet, 2009. **46**(10): p. 680-8.
101. Kalman, B. and E. Vitale, *Structural chromosomal variations in neurological diseases*. Neurologist, 2009. **15**(5): p. 245-53.
102. Iafrate, A.J., et al., *Detection of large-scale variation in the human genome*. Nat Genet, 2004. **36**(9): p. 949-51.
103. Shaffer, L.G. and B.A. Bejjani, *A cytogeneticist's perspective on genomic microarrays*. Hum Reprod Update, 2004. **10**(3): p. 221-6.

Curriculum Vitae

Personal data

Name: Filomena Tiziana
Surname: Papa
Date of birth: Feb 02, 1977
Place of birth: Foggia
Nationality: Italian

EDUCATION AND TRAINING

1996: School-leaving certificate

Scientific High School: “Guglielmo Marconi” of Foggia.

2005: Degree in Biological Sciences

in the field of Molecular Biology, achieved on November 15 with marks 106/110. Degree thesis in Molecular biology about "proteomic analysis of oocytes of *Danio Rerio*"
Supervisor: Prof. Luca Bini

2005: Professional Qualification for Biologist

2006: Winner of four year PhD in “Oncology and Genetics” - October 1th, 2006

2006-2010: Scholarship for a 4-year Doctoral in Oncology and Genetics

– Medical Genetics Course (coordinator prof. A. Renieri), University of Siena.

From April 2010: International Exchange Scholar – Department of Neurlogy, University of Kentucky, Lexington-USA

TEACHING RESPONSIBILITIES

Component of examen committee for:

Integrated course in Biology	1st year	Degree course in Medical School	from 2007-2009
------------------------------	----------	---------------------------------	----------------

Integrated course in Genetics	2nd year	Degree course in Midwifery	from 2007-2009
-------------------------------	----------	----------------------------	----------------

Lessons

Tutorial teaching activities in Medical Genetics :

Genetics	2nd year	Degree course in Medical School	from 2007-2009
----------	----------	---------------------------------	----------------

PUBLICATIONS

A 2.6 Mb deletion of 6q24.3-25.1 in a patient with growth failure, cardiac septal defect, thin upperlip and asymmetric dysmorphic ears.

Eur J Med Genet. 2007 Jul-Aug;50(4):315-21.

Caselli R, Mencarelli MA, **Papa FT**, Uliana V, Schiavone S, Strambi M, Pescucci C, Ariani F, Rossi V, Longo I, Meloni I, Renieri A, Mari F.

Clinical and molecular characterization of Italian patients affected by Cohen syndrome.

J Hum Genet 2007, 52(12):1011-7.

Katzaki E, Pescucci C, Uliana V, **Papa FT**, Ariani F, Meloni I, Priolo M, Selicorni A, Milani D, Fischetto R, Celle ME, Grasso R, Dallapiccola B, Brancati F, Bordignon M, Tenconi R, Federico A, Mari F, Renieri A, Longo I.

Expanding the phenotype of 22q11 deletion syndrome: the MURCS association.

Clin Dysmorphol 2008, Jan;17(1):13-17

Uliana V, Giordano N, Caselli R, **Papa FT**, Ariani F, Marcocci E, Gianetti E, Martini G, Papakostas P, Rollo F, Meloni I, Mari F, Priolo M, Renieri A, Nuti R.

Delineation of 7q36.1-36.2 deletion syndrome: long QT, renal hypoplasia and mental retardation. Am J Med Genet A. 2008 May 1;146A(9):1195-9.

Caselli R, Mencarelli MA, **Papa F**, Ariani F, Longo I, Meloni I, Vonella G, Hayek G, Renieri R, Mari F.

A 3 Mb deletion in 14q12 causes severe mental retardation, mild facial dysmorphisms and Rett-like features.

Am J Med Genetic Part A 2008 Aug 1;146A(15):1994-8

Papa FT, Mencarelli MA, Caselli R, Katzaki E, Sampietri K, Meloni I, Ariani F, Longo I, Maggio A, Balestri P, Grosso S, Farnetani MA, Berardi R, Mari F, Renieri A.

Private inherited microdeletion/microduplications: implications in clinical practice.

Eur J Med Genet. 2008 Sep-Oct;51(5):409-16.

Mencarelli MA, Katzaki E, **Papa FT**, Sampieri K, Caselli R, Uliana V, Pollazzon M, Canitano R, Mostardini R, Grosso S, Longo I, Ariani F, Meloni I, Hayek J, Balestri P, Mari F, Renieri A.

Array comparative genomic hybridization in retinoma and retinoblastoma tissues.

Cancer Sci. 2009 Mar;100(3):465-71.

Sampieri K, Amenduni M, **Papa FT**, Katzaki E, Mencarelli MA, Marozza A, Epistolato MC, Toti P, Lazzi S, Bruttini M, De Filippis R, De Francesco S, Longo I, Meloni I, Mari F, Acquaviva A, Hadjistilianou T, Renieri A, Ariani F.

14q12 Microdeletion syndrome and congenital variant of Rett syndrome.

Eur J Med Genet. 2009 Mar-Jun;52(2-3):148-52.

Mencarelli MA, Kleefstra T, Katzaki E, **Papa FT**, Cohen M, Pfundt R, Ariani F, Meloni I, Mari F, Renieri A.

A 9.3 Mb microdeletion of 3q27.3q29 associated with psychomotor and growth delay, tricuspid valve dysplasia and bifid thumb.

Eur J Med Genet. 2009 Mar-Jun;52(2-3):131-3.

Pollazzon M, Grosso S, **Papa FT**, Katzaki E, Marozza A, Mencarelli MA, Uliana V, Balestri P, Mari F, Renieri A.

Is HSD17B1 a new sex reversal gene in human?
Mol Cell Endocrinol. 2009 Dec 10;313(1-2):70.
Katzaki E, **Papa FT**, Mucciolo M, Uliana V, Renieri A.

3.2 Mb microdeletion in chromosome 7 bands q22.2-q22.3 associated with overgrowth and delayed bone age.

Eur J Med Genet. 2010 May-Jun;53(3):168-70. Epub 2010 Feb 26.

Uliana V, Grosso S, Cioni M, Ariani F, **Papa FT**, Tamburello S, Rossi E, Katzaki E, Mucciolo M, Marozza A, Pollazzon M, Mencarelli MA, Mari F, Balestri P, Renieri A.

Syndromic mental retardation with thrombocytopenia due to 21q22.11q22.12 deletion: Report of three patients.

Am J Med Genet A. 2010 Jul;152A(7):1711-7.

Katzaki E, Morin G, Pollazzon M, **Papa FT**, Buoni S, Hayek J, Andrieux J, Lecerf L, Popovici C, Receveur A, Mathieu-Dramard M, Renieri A, Mari F, Philip N.

Association between primary open-angle glaucoma (POAG) and WDR36 sequence variance in Italian families affected by POAG.

Br J Ophthalmol. 2010 Sep 2. [Epub ahead of print]

Frezzotti P, Pescucci C, **Papa FT**, Iester M, Mittica V, Motolese I, Peruzzi S, Artuso R, Longo I, Mencarelli MA, Mittica P, Motolese E, Renieri A.

Practical hints for management of array-CGH output in autism spectrum disorders: susceptibility regions versus uncertain cofactors.

Am J Med Genet Part B 2010 submitted

Papa FT, Canitano R, Mencarelli MA, Mucciolo M, Disciglio V, Marozza A, Radice L, Castagnini C, Dosa L, Pollazzon M, Hayek J, Renieri A, Mari F.

BOOK CHAPTER

Collaboration for a book chapter:

Mental Retardation. Chapter 7: X-Linked Mental Retardation: a Diagnostic, Clinical and Molecular update.

Caselli R, **Papa FT**, Ariani F, Meloni I, Renieri A. Mariani Foundation Paediatric Neurology Editor. 2007

