



UNIVERSITÀ
DI SIENA
1240

Ph.D in Genetics, Oncology and Clinical
Medicine

**New insights into the pathogenic
mechanisms associated with CNVs:
duplication of 17p13.3, mirror effect in
16p11.2 and recessive phenotype in
22q11.22.**

Mafalda Mucciolo

Supervisor: Prof. Alessandra Renieri

Thesis suitable for the title of “Doctor Europaeus”

Doctoral School in Medical Genetics

Academic year 2011-2012

XXIV cycle

Ph.D dissertation board

Prof. Rosanna Abbate

University of Florence, Florence, Italy

Prof. Sritharan Kadiramanathan

University of London, London, UK

Prof. Stylianos Antonarakis

University of Geneva, Geneva, Switzerland

Prof. Alessandra Renieri

University of Siena, Siena, Italy

Ph.D thesis reviewers

Prof. Thomas Liehr

Institute of Human Genetics, University of Jena, Germany

Prof. Béla Melegh

Department of Medical Genetics, University of Pécs, Hungary



Institut für Humangenetik

Institutsdirektor

Prof. Dr. med. C. Hübner

E-Mail:

Christian.Huebner@mti.uni-jena.de

Ansprechpartner

Molekulare Zytogenetik

PD Dr. rer. nat./med. habil. T. Liehr

Tel.: 03641 9-35533

Fax: 03641 9-35582

Labor: 03641 9-35538

E-Mail:

Thomas.Liehr@mti.uni-jena.de

Besuchsadresse:
Kollegiengasse 10
07743 Jena

Postadresse:
Postfach
07740 Jena

Jena, 11/15/2012

About Ph.D. thesis of Mafalda Mucciolo:

"New insights into the pathogenic mechanisms associated with CNVs: duplication of 17p13.3, mirror effect in 16p11.2 and recessive phenotype in 22q11.22."

The work presented by Mafalda Mucciolo was performed at the Doctoral School in Biomedicine and Immunological sciences (Siena, Italy) in the laboratory of Prof. Dr. Alessandra Renieri.

The question she worked on was the following:

Two types of genomic disorders can be distinguished: syndromic forms where the phenotypic features are largely invariant and fully penetrant, and those where the same genomic rearrangement associates with a variant clinical outcomes. For the latter cases two ideas/theories shall be tested:

- a) some microdeletion syndromes could go together with activation of otherwise "recessive" mutations of genes present only in one copy after deletion of the other allele
- b) CNVs can be responsible of complex disorders in association with multiple high-penetrant alleles of low frequency.

These two theories were tested in cases 22q11.2 microdeletion, the 16p11.2 microdeletion/duplication and the 10q11.22 deletion/duplication.

As far as I understand no final conclusion could be drawn to answer the question finally. However, in two cases with 22q11.2 microdeletion M. Mucciolo found a mutation in the monosomic region, i.e. in the chromosome 22 without deletion, which is in support of above mentioned idea a). Idea b) is more supported by the recently suggested 'two-hit-model' of CNV – which is also discussed by M. Mucciolo. So overall, both mechanisms could contribute here.

Overall, big parts of her results were already published or are preparation for publishing (cumulative PhD thesis) – one even is a coauthorship in Nature.

Concerning her Ph.D.-thesis she put the data together in a form which meets international criteria and discussed her results thoroughly. I strongly recommend that her Ph.D. thesis is accepted.

Sincerely

Genetische Beratung
Fr. OÄ Dr. med. I. Schreyer
Hr. Prof. Dr. med. C. Hübner
Tel.: 03641 9-34924
Fax: 03641 9-34925

Molekulargenetische Diagnostik
Hr. OA Dr. med. I. Kurth
Tel.: 03641 9-34877
Fax: 03641 9-35502
Labor: 03641 9-35542

Zytogenetik
Fr. Dr. A. Weise
Fr. Dr. K. Mrasek
Tel.: 03641 9-35530
Fax: 03641 9-35518
Labor: 03641 9-35586

Molekulare Zytogenetik
Hr. PD Dr. T. Liehr
Tel.: 03641 9-35533
Fax: 03641 9-35582
Labor: 03641 9-35538

Tumorgenetik
Fr. Dr. A. Glaser
Tel.: 03641 9-35534
Fax: 03641 9-35518
Labor: 03641 9-35512

Funktionelle Genetik
Hr. Prof. Dr. med. C. Hübner
Hr. OA Dr. med. I. Kurth
Tel.: 03641 9-34877
Fax: 03641 9-35518
Labor: 03641 9-35511

Core-Unit Chipapplikation
Hr. Prof. Dr. F. v. Eggeling
Tel.: 03641 9-35526
Fax: 03641 9-35518
Labor: 03641 9-35528

Molekulargenetik
Hr. Prof. Dr. A. Baniahmad
Tel.: 03641 9-35524
Fax: 03641 9-34706

Anthropologie
Fr. PD Dr. K. Kromeyer-Hauschild
Tel.: 03641 9-34617
Fax: 03641 9-34618



UNIVERSITA' DEGLI STUDI DI SIENA

DIPARTIMENTO DI BIOTECNOLOGIE

Sezione di Genetica: Policlinico Le Scotte

Prof. Alessandra Renieri tel 0577 233303 FAX 0577 233325 e-mail

Siena, 26-11-2012

To:
Prof. Thomas Liehr
Institute of Human Genetics
University of Jena, Germany

Dear prof. Liehr,

I have really appreciated your comments on my work, and I would like to thank you very much for reviewing my Thesis.

I look forward to continue my research on genomic disorders in order to contribute to better define the molecular mechanisms underlying the phenotypic variability.

Best regards,

Mafalda Mucciolo



TO WHOM IT MAY CONCERN

24 November, 2012

Review of the PhD thesis: “New insights into the pathogenic mechanisms associated with CNVs: duplication of 17p13.3, mirror effect in 16p11.2 and recessive phenotype in 22q11.22.”
by Mafalda Mucciolo

The doctoral thesis written by Mafalda Mucciolo comprises 100 pages including a list of references. The sections of the thesis are proportional, and follow the usual structure of the doctoral theses.

The thesis is a carefully assembled work considering both content and format. The studies presented in the thesis represent new approach with highly sophisticated methods and to the study of genomic disorders; hereby, their significance is for both medical and scientifically outstanding. The logical structure of the dissertation is easy to follow. The clinical descriptions are precise and detailed. The figures and tables are appropriate and correctly reflects the information discussed through the text. The discussion and conclusion part are also well written, the conclusions are sound.

The thesis is focusing on four key research topics:

1. Reciprocal duplication in Miller-Dieker syndrome
2. Microdeletion and microduplication in 16p11.2
3. Microdeletion and microduplication in 10q11.22
4. Microdeletion unmasking recessive phenotype

Sudy of all of them provided valuable new data on genomic rearrangements in the background of genomic disorders which remained mostly hidden in the past due to the limited resolution of conventional cytogenetic techniques. The research activity demonstrated in the thesis contributed to better understanding and further delineation of the features associated with novel microduplication syndromes as well.

Comments and questions:

1. I would suggest to insert a list of abbreviations used in the thesis, it is almost conventional, and helps the reader.
2. There are many disorders, clinical conditions mentioned in the text. The use of MIM numbers (if they are available) can be helpful.
3. The microdeletion and microduplication in 10q11.22 is not discussed in the introduction. The importance of this issue is described firstly in the chapter „Aims and outlines of the study”.
4. Rephrase the sentence: „Moreover we reported two unrelated girls carrying a duplications of the Miller-Dieker region at 17p13.3.”
5. A typing error is on the page 30: „The overall phenotype of these two cases is complicated by the presence of a second copy number variation ,and some phenotypic features of our patients can be attributed either to 9p deletion or 10q deletion.”
6. An unnecessary space occurs in the nomenclature in a sentence on page 35: „MLPA analysis confirmed the presence of a duplication of the area containing the *RPH3AL* probe on chromosome 17p13.3 in both patients, a deletion of the *PAOX* probe on chromosome 10 q26.3 in Patient 1, and a deletion of the *DMRT1* probe on chromosome 9 p24.3 in Patient 2 (data not shown)”.
7. A typing error is on the top of page 60: „In a first analysis performed by array-CGH in our cohort of patients, we identified 12 individuals sharing an overlapping CNVs in 10q11.22 (3 deletions and 9 duplications).”
8. In the „Materials and Methods” can be read that custom available oligonucleotide arrays were used for analyses. However, by identification of CNVs in 10q11.22 the selected and previously analysed 292 patients were negative for deletions and duplications in 10q11.22 (by array-CGH 44K). We can read on the page 60 that the 44K slides have only one probe located in the 10q11.22 region.
The question arises, why not another array was used in the examination with specific probes representing this region better.

The thesis is based on research work published in the scientific literature which is the evidence of the successful presentation. The articles meet the PhD requirements in their number and level. The topic chosen by the candidate and her supervisor represent a new wave, utilizes new generation techniques, and very exciting field of the postgenomic studies. Apart from comments arisen by the present reviewer, the substantial work and high ranked publications of Mafalda Mucciolo and her coworkers presented in this doctoral thesis fulfills the requirements of a doctoral thesis and is suitable for achieving the title of “Doctor Philosophiae”.

Béla Melegh, M.D., Ph.D., D.Sc.
professor of Medical Genetics, Pediatrics,
and Laboratory Genetics

Siena, 26-11-2012

To:
Prof. Béla Melegh
Department of Medical Genetics
University of Pécs, Hungary

Dear Prof. Melegh,

I really appreciated your careful revision of my PhD thesis. I thank you very much for your overall comments.

I am herein including a detailed response to your questions:

1. *I would suggest to insert a list of abbreviations used in the thesis, it is almost conventional, and helps the reader.*

I added a list of all the abbreviations used in the thesis.

2. *There are many disorders, clinical conditions mentioned in the text. The use of MIM numbers (if they are available) can be helpful.*

Whenever possible I associated each disorder to the correspondent MIM number.

3. *The microdeletion and microduplication in 10q11.22 is not discussed in the introduction. The importance of this issue is described firstly in the chapter „Aims and outlines of the study”.*

A new paragraph about microdeletions and microduplications in 10q11.22 has been added in the introduction.

4. *Rephrase the sentence: „Moreover we reported two unrelated girls carrying a duplications of the Miller-Dieker region at 17p13.3.”*

I changed the sentence in: „Moreover we reported two unrelated girls carrying a duplication of the Miller-Dieker region at 17p13.3.”

5. *A typing error is on the page 30: „The overall phenotype of these two cases is complicated by the presence of a second copy number variation ,and some*

phenotypic features of our patients can be attributed either to 9p deletion or 10q deletion.”

I corrected the error on page 30: „The overall phenotype of these two cases is complicated by the presence of a second copy number variation, and some phenotypic features of our patients can be attributed either to 9p deletion or 10q deletion.”

6. *An unnecessary space occurs in the nomenclature in a sentence on page 35: „MLPA analysis confirmed the presence of a duplication of the area containing the RPH3AL probe on chromosome 17p13.3 in both patients, a deletion of the PAOX probe on chromosome 10 q26.3 in Patient 1, and a deletion of the DMRT1 probe on chromosome 9 p24.3 in Patient 2 (data not shown)”.*

I removed the spaces in the sentence on page 35: „MLPA analysis confirmed the presence of a duplication of the area containing the RPH3AL probe on chromosome 17p13.3 in both patients, a deletion of the PAOX probe on chromosome 10q26.3 in Patient 1, and a deletion of the DMRT1 probe on chromosome 9p24.3 in Patient 2 (data not shown)”.

7. *A typing error is on the top of page 60: „In a first analysis performed by array-CGH in our cohort of patients, we identified 12 individuals sharing an overlapping CNVs in 10q11.22 (3 deletions and 9 duplications).”*

I rephrased the sentence on page 60: „In a first analysis performed by array-CGH in our cohort of patients, we identified 12 individuals sharing a overlapping CNV in 10q11.22 (3 deletions and 9 duplications).”

8. *In the „Materials and Methods” can be read that custom available oligonucleotide arrays were used for analyses. However, by identification of CNVs in 10q11.22 the selected and previously analysed 292 patients were negative for deletions and duplications in 10q11.22 (by array-CGH 44K). We can read on the page 60 that the 44K slides have only one probe located in the 10q11.22 region. The question arises, why not another array was used in the examination with specific probes representing this region better.*

Considering the labour effort required in producing a custom array with respect to that necessary to design specific MLPA probes for this region, we decided to start our screening using a MLPA assay. However, taking into account the positive results achieved until now, a custom array could be a more adequate solution in order to better define the exact breakpoints of CNVs occurring in 10q11.22 region.

Thank you again for your suggestions.

Best regards,

Mafalda Mucciolo

INDEX

Acknowledgements

List of abbreviations.....	1
1. INTRODUCTION	
1.1. Historical overview.....	4
1.2. Array – CGH Methodologies.....	8
1.3. Clinical utility of array-CGH.....	9
1.3.1. Discovering new syndromes.....	10
1.3.2. Expanding the phenotypic spectrum of known syndrome.....	10
1.3.3. Identifying the reciprocal products of known abnormalities.....	11
1.3.4. Identifying the genomic lesions in known conditions.....	11
1.3.5. Increasing the frequency of copy number variations across the genome.....	12
1.4. Copy number variation (CNVs).....	13
1.5. Copy number variation and phenotypic variability.....	14
1.6. Reciprocal duplication of the Miller-Dieker region.....	15
1.7. Microdeletion and microduplication in 16p11.2.....	16
1.8. Microdeletion and microduplication in 10q11.22.....	16
1.9. Phenotype variability in 22q11.2 deletion syndrome.....	17
2. AIM AND OUTLINE OF THE STUDY.....	19

3. MATERIALS AND METHODS	
3.1. Patients collection.....	23
3.2. Array-based CGH.....	23
3.2.1. Samples preparation.....	23
3.2.2. Human oligonucleotides array.....	23
3.3 Real-time quantitative PCR	25
3.4 Multiplex Ligation-dependent Probe Amplification (MLPA).....	25
4. RESULTS	
4.1. Reciprocal duplication of known deletion syndrome.....	28
4.2. Microdeletion and microduplication in 16p11.2.....	46
4.3. Microdeletion and microduplication in 10q11.22.....	56
4.4. Microdeletion unmasking recessive phenotype.....	70
5. DISCUSSION.....	84
6. CONCLUSIONS and FUTURE PERSPECTIVES.....	91
7. REFERENCES.....	94
<i>Curriculum vitae</i>	101

Acknowledgements

Four years ago I started a long and unknown route named "Doctoral school" that to my eyes appeared so hard and tortuous. Today, at the end of the trip, if I look back I see a really steep path...but now I am on the top!

I would like to thank for this Prof. Alessandra Renieri, who believed in me and gave me the possibility to increase my scientific knowledge.

A great thanks to all the senior researchers, whose experience always led my work.

A special thanks to Viria, companion of adventures and misadventures of the "wonderful world of array-CGH" and a thanks also to Sonia, Elisa and Enea, new members of the array-group.

Thanks to all my colleagues for their constantly support but most of all for their friendship.

Thanks to Prof. Antonarakis, who gave me the possibility to attend his lab and to all his co-workers who made really exciting my experience in Geneva.

A special thanks to Francesco who always repeated me "you can do it".

And a big thanks to my family whose love always encouraged me to follow my dreams.

List of abbreviations

22q11.2DS = deletion syndrome

aCGH = array based comparative genomic hybridization

ASDs = autism-spectrum disorders

BACs = bacterial artificial chromosomes

BMI = body mass index

BSS = Bernard-Soulier syndrome

CCRs = complex chromosome rearrangements

CGH = comparative genomic hybridization

CMT1A = Charcot–Marie–Tooth syndrome type 1A

CNVs = copy number variants

DD = developmental delay

FISH = Fluorescent in situ hybridization

FoSTeS = fork stalling template switching

HNPP = hereditary neuropathy with liability to pressure palsies

ID = intellectual deficit

ILS = isolated lissencephaly

LCRs = between low copy repeats

MCA = multiple congenital anomalies

MDS = Miller–Dieker syndrome

m-FISH = multicolour FISH

MLPA = Multiplex Ligation-dependant Probe Amplification

MURCS = Mullerian Renal Cervico-thoracic Somite anomalies

NAHR = non-allelic homologous recombination

NHEJ = non-homologous end joining

NMD = nonsense-mediated mRNA decay

NSID = Non Syndromic Intellectual Disability

OFC = occipitofrontal circumference

RPA = Relative Peak Area

SID = Syndromic Intellectual Disability

SKY = spectral karyotyping

SMCs = supernumerary marker chromosomes

UCRs = ultra conserved regions

UTR = untranslated region

VCFS = velo cardio facial syndrome

1. INTRODUCTION

1) INTRODUCTION

1.1 Historical overview

The past 50 years have seen an explosion of methodological advances in molecular cytogenetic technology. These cytogenetic techniques added colour to the black and white world of conventional banding. Cytogenetic analysis of Giemsa-stained metaphase chromosomes (Fig.1a), identifies balanced and unbalanced structural and numerical chromosomal abnormalities (Shinawi 2008). However, even high resolution karyotypes (Yunis 1976) are unable to detect many known microdeletion syndromes, which range from 3-5 Mb in size, and cannot detect smaller aberrations. 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. Fluorescent in situ hybridization (FISH), which is the annealing of fluorescently labelled locus-specific probes to their complementary sequences in the genome, allowed for the detection of specific microdeletion syndromes (Trask 1991) (Fig.1b¹-b²). FISH technique can be used to map loci on specific chromosomes, detect both structural chromosomal rearrangements and numerical chromosomal abnormalities, and reveal cryptic abnormalities such as small deletions. FISH analysis is, however a time-consuming, targeted method that requires prior knowledge of the chromosomal region of interest and therefore interrogates one or more candidate chromosomal loci at a time. Therefore 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.

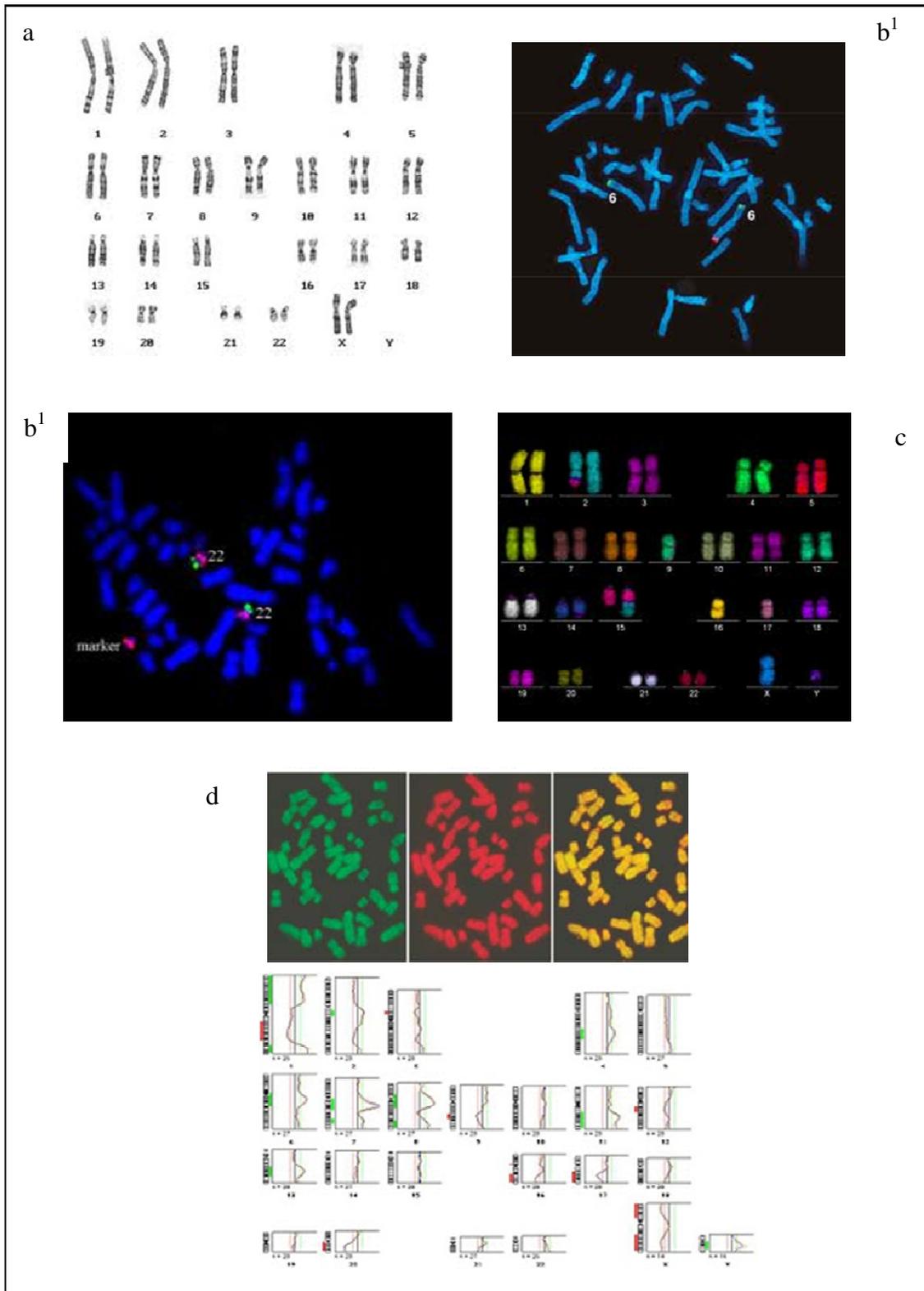


Fig.1 a. G banded karyotype. b¹-b². Fluorescence in situ hybridization (FISH) of metaphase human chromosomes. c. multicolour FISH (m-FISH). d. Comparative Genomic Hybridisation experiment.

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 (Fig.1c). CGH was developed initially as a molecular tool in tumor cytogenetics (Kallioniemi 1992). It detects genomic imbalances and determines the map position of gains and losses of chromosomes or chromosomal sub-regions on normal reference metaphase preparations using a small amount of DNA. In this technique, patient and reference whole-genome DNA are differentially labelled and co-hybridized to normal metaphase spread on glass slides. Unbalanced chromosomal rearrangements at a resolution of ~3-10 Mb across the whole genome can be detected by differential hybridization signals (Kirchhoff 1999) (Fig.1d). This method 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 (Yunis 1976). 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 (Fig.2). However, the substitution of the metaphase chromosomes with target DNAs robotically spotted immobilized onto glass microscope slides using split metal pins or glass capillaries has significantly enhanced the resolution and simplified the analysis procedure (Shinawi 2008).

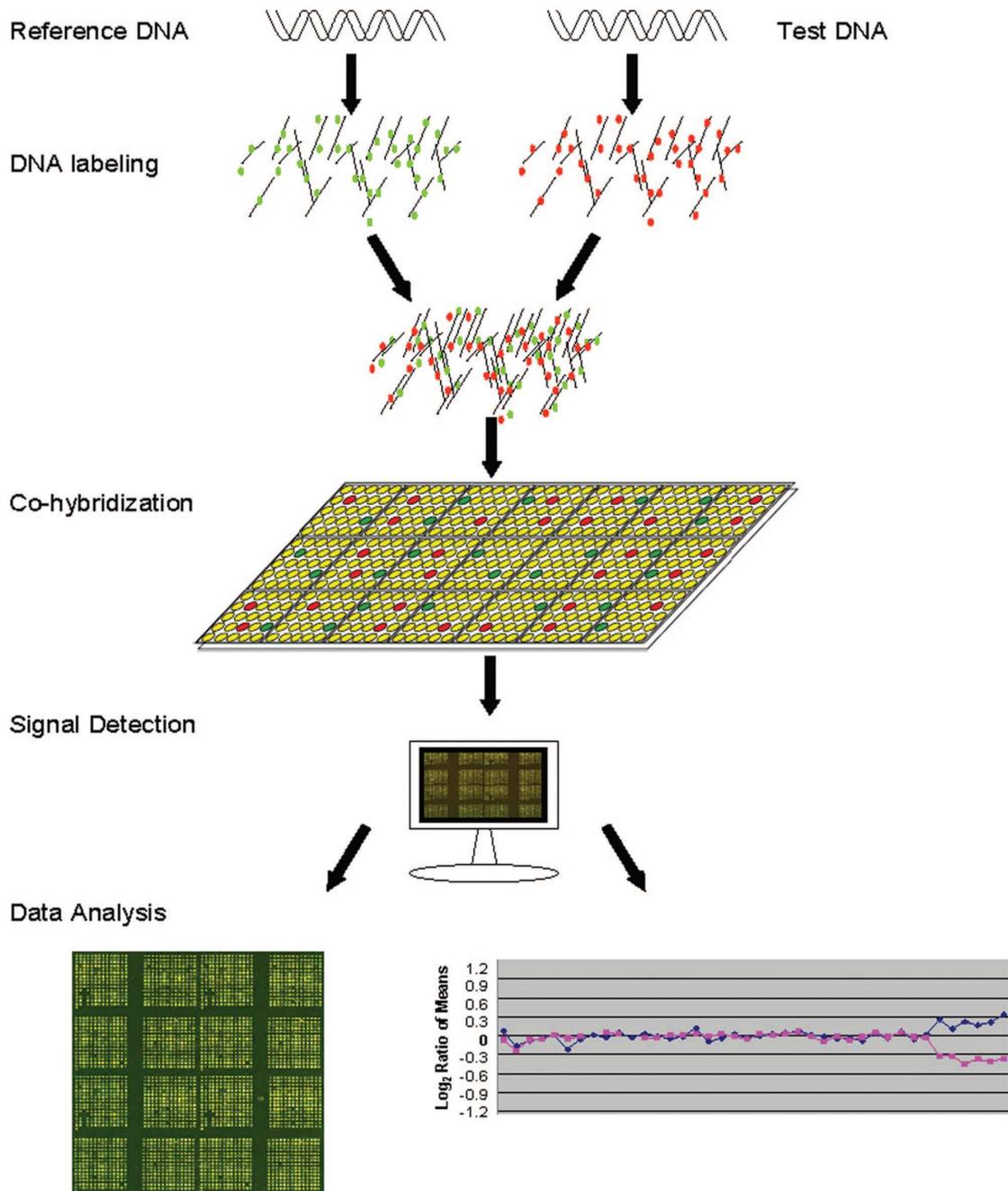


Fig.2 Schematic representation of an array-CGH experiment. Test and reference DNA are differentially labelled, co-precipitated and hybridised to an array. After wash procedures, the slides are analysed through a scanner and fluorescence intensities of each probe are determined. After imaging processing 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. In this representation, copy number loss shift the ratio to the left and copy number gains shift the ratio to the right.

The higher resolution and throughput with possibilities for automation, robustness, simplicity, high reproducibility and precise mapping of aberrations are the most significant advantages of aCGH over cytogenetic methods. In addition, there is no need for cell culture, making the turn around time shorter than in cytogenetic methods. As with other clinical diagnostic methods, there are limitations in aCGH technology. aCGH is not able to identify balanced rearrangements such as translocations and inversions and low-level mosaicism for unbalanced numeric or structural rearrangements.

1.2 Array – CGH Methodologies

In aCGH, equal amounts of labelled genomic DNA from a test and a reference sample are co-hybridized to an array containing the DNA targets. Genomic DNA of the patient and control are differentially labelled with Cyanine 3 (Cy3) and Cyanine 5 (Cy5). The slides are scanned into image files using a microarray scanner. The spot intensities are measured and the image files are quantified using feature extraction software, and text file outputs from the quantitative analyses are imported into software programs for copy number analysis (Fig.2) (Cheung 2005, Lu 2007). The resulting ratio of the fluorescence intensities is proportional to the ratio of the copy numbers of DNA sequences in the test and reference genomes.

Two major types of array targets are currently being utilized. Initially, bacterial artificial chromosomes (BACs) were the array target of choice (Pinkel 1998). However, now 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 (Ylstra 2006). 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 (Ishkanian 2004). Oligonucleotide arrays offer a flexible format with the potential for very high

resolution and customization. Several different platforms are available for oligonucleotide arrays that range from 25- to 85mers in length, 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 (Shaikh 2007). Both BAC and oligonucleotide arrays have been used successfully to detect copy number changes in patients with intellectual deficit (ID), multiple congenital anomalies (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 sub-telomeres and those responsible for known microdeletion/microduplication syndromes, but does not have probes that span the whole genome (Bejjani 2005, Bejjani 2006, Shaffer 2006). A whole genome or tiling path array offers full genome coverage with different resolution. The resolution of array CGH is defined by two main factors: 1) the size of the nucleic acid targets and 2) the density of coverage over the genome; the smaller the size of the nucleic acid targets and the more contiguous the targets on the native chromosome, the higher the resolution of the array.

1.3 Clinical utility of array-CGH

The considerable gap in resolution conventional cytogenetic techniques (5-10 Mb pairs) and molecular biology techniques (base pairs) has been bridged by aCGH, which allows the detection of genomic imbalances associated with phenotype of unknown genetic aetiology. This new technology has driven a technical convergence between molecular diagnostics and clinical cytogenetics, questioned our understanding of the complexity of the human genome and revolutionized the practice of medical genetics. The use of aCGH in research and diagnostics has resulted in the identification of many new syndromes, expanded our knowledge about the phenotypic spectrum of existing conditions, identified the reciprocal products of known abnormalities, elucidated the genomic lesions in known conditions, and ascertained the unexpected frequency of copy number variations across the genome.

1.3.1 Discovering new syndromes

Deletion and duplication syndromes represent recurrent chromosomal abnormalities that are associated with distinct phenotypes. These microdeletions/microduplications often occur between low copy repeats (LCRs) and are commonly because of non-allelic homologous recombination (NAHR) events (Lupski 1998). The detection of a de novo genomic imbalance in a single patient does not prove pathogenicity. Only the identification of similar genomic imbalances with a recognizable phenotype can help clarify the role of these genomic changes in causing the specific clinical features and will ultimately define a genetic syndrome. Therefore, 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 aetiology, to a "genotype-first" approach, in which a collection of individuals with similar copy-number imbalances can be examined for common clinical features (Neill 2010).

1.3.2 Expanding the phenotypic spectrum of known syndrome.

“Known syndrome” are defined as syndromes exhibiting a spectrum of signs and symptoms sufficient to encourage the clinician to proceed with a specific test in order to confirm the clinical diagnosis. The ascertainment through whole-genome screening of syndromic patients by array-CGH leads to the recognition of a broader spectrum of features for already described syndromes ranging from severe phenotype to a normal phenotype (van Bon 2009). 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.

1.3.3 Identifying the reciprocal products of known abnormalities

Many the well-known microdeletion syndromes are mediated by segmental duplications sequences (Lupski 1998). The clinical phenotypes associated with the reciprocal microduplications of the same genomic regions are, however, less well characterized probably because, in general, individuals with duplications tend to have a milder phenotype than those with the complementary deletions and this milder phenotype may not lead to clinical investigation (Van der Aa 2009; Hased 2004; Potocki 2000). The introduction of aCGH in clinical practice has 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 intellectual deficit/developmental delay (ID/DD), the frequency of microduplications at the common microdeletion syndrome loci will likely increase (Bejjani and Shaffer 2008).

1.3.4 Identifying the genomic lesions in known conditions

The high resolution afforded by array CGH has been used to define candidate regions for putative genes responsible for human genetic diseases. A good example is the discovery of a candidate gene for CHARGE syndrome (MIM#214800), a pleiotropic disorder comprising of coloboma, heart defects, choanal atresia, retarded growth and development, genital and/or urinary abnormalities, ear anomalies and deafness. Vissers and colleagues (Vissers 2004) hybridized cell lines from two individuals with CHARGE syndrome onto a genome-wide array with a 1Mb resolution. The authors narrowed a candidate region for CHARGE syndrome on 8q12 based on data from two individuals, one with a ~5 Mb deletion and another with a more complex rearrangement comprising two deletions that overlapped that of the first deletion subject. These results allowed the authors to focus on only nine genes in the region and detect heterozygous mutations in the gene *CHD7*, which was eventually shown to be the gene for CHARGE syndrome. The high resolution of that array was crucial in refining the critical region for this disease and in reducing the number of candidate genes to be investigated further.

1.3.5 Increasing the frequency of copy number variations across the genome

Array CGH has the ability to detect submicroscopic gains and losses of the genome at very high resolution and is performed with the goal of identifying pathogenic chromosomal aberrations or copy number variants (CNVs) that are directly responsible for the observed clinical phenotype. However, CNVs have been described in the literature that are present in phenotypically normal individuals and in some cases occur at a high frequency in the general population (Iafate 2004; Sebat 2004; Sharp 2005; Redon 2006; McCarroll 2007). Some of these aberrations are apparently benign CNVs and are usually inherited from a parent (Lee 2007). If identical alterations are found either in one of the unaffected parents, or in independent normal controls, they most probably have no direct phenotypic consequences; however, low penetrance and variable expressivity of the phenotype may complicate the analysis and genetic counseling. Currently, the publicly available CNV databases assist in making decisions about the clinical significance of imbalances detected by microarrays. Examples of such databases are the Database of Genomic Variants (<http://projects.tcag.ca/variation>). When determined as de novo in origin genomic imbalances are considered more likely pathological (Tyson 2005). This can be further supported if the implicated region contains gene(s) with functions compatible with the abnormal clinical findings or previously described patients with a similar genomic imbalance and a similar phenotype. The de novo occurrence of copy number alteration is, however, not an absolute evidence of its pathogenicity and caution must be exercised for possible non paternity. Moreover genetic modifiers or thresholds involving other copy-number alterations could play a role in the manifestation of clinical features, or other independent mutations elsewhere in the genome may obfuscate the interpretation of such data.

1.4 Copy number variations (CNVs).

CNVs can either be inherited or caused by de novo mutations of different size. They range from 1 kb to several Mb in size and, therefore, with increasing resolution of aCGH platforms more variations will be detected. These structural variants show variable copy number when compared to a reference genome and include both deletions and duplications of genomic loci (Feuk 2006). They have been reported to encompass as much as 12% of the genome (Redon 2006) and today several molecular mechanisms are known to be responsible for the occurrence of CNVs within the genome (Gu 2008). The major mechanisms underlying the former is non-allelic homologous recombination for recurrent rearrangements, and non-homologous end joining (NHEJ) for non-recurrent rearrangements. NAHR can use either region-specific low-copy-repeat (LCRs or segmental duplications) or sometimes repetitive sequences (e.g. Alu or LINE) as homologous recombination substrates, yielding recurrent events with clustered breakpoints (Lee 2007). When LCRs are located on the same chromosome and in direct orientation, NAHR results in deletion and/or duplication. Inversions result when LCRs on the same chromosome are in opposite orientation; whilst NAHR between LCRs located in different chromosomes result in translocation (Colnaghi 2011). However, a number of disease-associated rearrangements are not explained readily by either the NAHR or NHEJ recombinational mechanisms. Lee et al, proposed a new DNA replication-based mechanism termed FoSTeS to parsimoniously explain the generation of these complex rearrangements in the human genome. According to the FoSTeS model, during DNA replication, the active replication fork can stall and switch templates using complementary template microhomology to anneal and prime DNA replication (Lee 2007). The rearrangements generated by FoSTeS can be diverse in scale, from genomic duplications affecting megabases of the human genome to small deletions involving a single gene or only one exon. These different sized rearrangements implicate FoSTeS in CNVs of all sizes and in the evolution of both human genomes and genes (Zhang 2009).

1.5 Copy number variation and phenotypic variability.

It is now known that any individual carries ~1000 CNV ranging from 443 bp to 1.28 Mb (Conrad et al 2010). This can lead to either too many or too few dosage sensitive genes, which might result in phenotypic variability, complex behavioural traits and disease susceptibility. Interestingly, CNVs have not only been associated with disease, but also with genome evolution and adaptive traits. The *AMY1* gene, which encodes a protein that catalyses the first step in digestion of dietary starch and glycogen, constitutes an interesting example. It has been found that the copy number of this gene is three times higher in humans compared to chimpanzees, suggesting that humans were favoured in the gene dosage due to a concomitant increase of starch consumption (Perry et al, 2007). However, it still remains the problem to understand if CNV means disease and how these structural changes and gene dosage alterations contribute to phenotypic outcomes. Actually we know that CNVs affecting specific genes or chromosomal regions, can lead to susceptibility and predisposition to certain diseases such as HIV, lupus, nephritis, pancreatitis and psoriasis among many other phenotypes (Canales 2011). However, it has been shown that individuals carrying the same rearrangement, for instance within an affected family, show differences in the manifestation of the investigated phenotype.

There are several explanations for variable expressivity and clinical heterogeneity in genomic disorders. First, the breakpoints of the events may not be identical. Atypical deletions and duplications involving contiguous dosage-sensitive genes within the region often explained the observed clinical variability in many genomic disorders. Further studies demonstrated that the variability can be due to the presence of an additional large deletion or duplication in the proband that resulted in a sensitized genetic background and consequently a more pronounced phenotype (Girirajan 2010). However the commonly proposed functional impact of a CNV has been the haploinsufficiency or dosage sensitivity for one or more genes within the genomic region, or the possibility that a recessive gene resides within the rearranged region.

1.6 Reciprocal duplication of the Miller-Dieker region.

The short arm of chromosome 17 is particularly prone to submicroscopical rearrangements due to a high density of low copy repeats. Thus, the proximal 17p region harbours regions with microdeletion and reciprocal microduplication syndromes, each caused by non-allelic homologous recombination: CMT1A (Charcot–Marie–Tooth syndrome type 1A) (MIM#118220), due to a duplication at 17p11.2; HNPP (hereditary neuropathy with liability to pressure palsies) (MIM#162500), due to a reciprocal deletion, Smith–Magenis syndrome (MIM#182290), caused by a deletion at 17p11.2; and the relatively recently described Potocki–Lupski syndrome (MIM#610883), due to a reciprocal duplication at 17p11.2 (Stankiewicz 2003; Potocki 2000). Deletions in the more distal region 17p13.3, including the *PAFAH1B1* gene (encoding LIS1), result in the brain malformation lissencephaly, with reduced gyration of the cerebral surface and increased cortical thickening. Depending on the size of the deletion, the phenotype varies from isolated lissencephaly (ILS) (MIM#607432) to Miller–Dieker syndrome (MDS) (MIM#247200); the latter consists of severe grade ILS and additional characteristic dysmorphic features and malformations (Dobyns 1993). Deletions in MDS vary in size, from 0.1 to 2.9 Mb. The critical region differentiating ILS from MDS is approximately 400 Kb, and is referred to as the “MDS telemetric critical region” (Cardoso et al, 2003). Recently, 17p13.3 duplications involving the *PAFAH1B1* gene have been reported in patients with psychomotor retardation, hypotonia and dysmorphic features without lissencephaly or gross brain malformations (Bi et al, 2009; Roos et al, 2009; Bruno et al, 2010). The phenotype of transgenic mice conditionally overexpressing *PAFAH1B1* is indeed characterized by decreased brain size and neuronal migration abnormalities. All the submicroscopic rearrangements reported until now, are variable in size and have distinct breakpoints. Bruno et al. proposed to divide 17p13.3 microduplications in two different classes: class I duplications involving *YWHAE* but not *PAFAH1B1* showing a phenotype characterized by learning difficulties and/or autism with or without other congenital abnormalities; class II duplications always harboring *PAFAH1B1* that may also

include the genomic region encompassing the *CRK* and *YWHAE* genes, which are associated with developmental delay, psychomotor delay and associated hypotonia.

1.7 Microdeletion and microduplication in 16p11.2

The 16p11.2 region is a well-documented hot spot for recurrent rearrangements that are associated with autism-spectrum disorders (ASDs) and ID (Marshall 2008; Kumar 2008; Weiss 2008). This 555 kb CNV region, which is flanked by segmental duplications having >99% sequence identity, is presumed to have an elevated mutation rate due to its genomic architecture (Lupski 2007). Weiss et al. reported a recurrent microdeletion on chromosome 16p11.2 in five of 751 families with one or more cases with ASD, in three of 299 ASD patients, in five of 512 children referred for ID and/or autism (Weiss 2008). The reciprocal duplication was found in 11 patients and in five controls. In another study, the same deletion was detected in four of 712 autistic patients and none of 837 controls (Kumar 2008). The latter study identified the reciprocal duplication in one autism case and two controls. Similarly, Marshall et al. detected two de novo 16p11.2 deletions in 427 families with autism (Marshall 2008). The authors stated that deletions and duplications of 16p11.2 carry substantial susceptibility to autism, and that the deletions appear to account for approximately 1% of cases. Furthermore Walters et al; demonstrated that, in addition to the cognitive deficits or behavioural abnormalities, a 16p11.2 deletion give rise to a strongly-expressed obesity phenotype in adults, with a more variable phenotype in childhood (Walters 2011). The authors stated that the higher frequency of 16p11.2 deletions in the cohort ascertained for both phenotypes (2.9%), compared to cohorts ascertained for either phenotype alone (0.4% cognitive deficit and 0.6% obesity), confirms their impact on both obesity and developmental delay, adding to the evidence that these two phenotypes may be fundamentally interrelated.

1.8 Microdeletion and microduplication in 10q11.22

To date, interstitial deletions involving 10q11.2 have been reported in over 40 patients with variable abnormal phenotypes but also in individuals with a normal phenotype. The only clinical features common to a majority of affected individuals

were ID and DD. Stankiewicz and colleagues identified 24 unrelated individuals carrying a microdeletions at 10q11.21q11.23 ranging in size from \approx 1.9 to \approx 10.9 Mb. They also identified 17 individuals with reciprocal microduplications involving 10q11.21q21.1, ranging in size from \approx 0.3 to \approx 12 Mb (Stankiewicz 2010). A complex arrangement of six segmental duplication clusters have been identified in the 10q11.21q11.23 region, labelled LCR 10q11.2A-LCR10q11.2F. These segmental duplications range in size from 32 to 427 kb and have a complex evolutionary structure. Therefore, the complex structure of the LCR10s in this region appears to be involved in generating a variety of different genomic rearrangements. The finding of different sized rearrangements on chromosome 10q is similar to that observed for other recurrent genomic disorders, such as the Prader Willi/Angelman syndrome, Smith-Magenis syndrome, and the 15q24 deletion syndrome, where recombination within alternate LCRs can result in recurrent deletions and duplications of different size. CNVs overlapping the proximal LCRs are also frequent in control subjects. More recently a smaller duplication have been reported in patients showed the Zappella variant of Rett syndrome (Z-RTT) (Artuso 2011). The 10q11.22 duplication was considered a hypothetical modifier that can modulate the phenotype in patients matched for MECP2 mutation.

1.9 Phenotype variability in 22q11.2 deletion syndrome.

Microdeletion of chromosome 22q11.2 or 22q11.2 deletion syndrome (22q11.2DS) (MIM#188400/#192430) is the most common human deletion syndrome with an estimated prevalence of 1 in 4,000 live births (Goodship et al. 1998). The phenotypic spectrum encompasses several previously described syndromes including DiGeorge, velocardiofacial and conotruncal anomaly face syndromes as well as some individuals with other conditions such as Cayler cardiofacial syndrome. The phenotypic expression of the 22q11.2DS is known to be highly variable and ranges from a severe life-threatening condition to affected individuals with few associated features (Bassett et al. 2005; Kobrynski and Sullivan 2007; Ryan et al. 1997). The most frequent feature is a conotruncal heart defect, often associated with facial dysmorphisms, cleft palate, thymus hypoplasia, and

learning disability (McDonald 1999). Developmental delays and learning difficulties are very commonly associated, although severe intellectual disability is rare. Recurrent seizures are common and epilepsy may be present in about 5% of patients. Psychiatric conditions may be present in children and over 60% of patients develop treatable psychiatric disorders by adulthood (Bassett et al. 2005). In particular, due to the high frequency of schizophrenia in 22q11.2DS patients, the 22q11.2 region is considered to be one of the main schizophrenia susceptibility loci in humans (Bassett and Chow 2008; Insel 2010). Evidence from multiple studies indicates that about 1% of individuals with schizophrenia in the general population have 22q11.2 deletions (Bassett et al. 2010).

The high frequency of the 22q11.2 deletion can be explained by the presence of chromosome-specific low copy repeats flanking (LCR A and D) or within the typically deleted region (LCR A', B and C) (Shaikh TH 200). Most deletions (84–90%) encompass ~3 Mb, known as the typically deleted region. Smaller deletions, spanning 1.5 Mb, are found in about 7–14% of the cases (Carlson 1997; Saitta 2004). In addition, atypical deletions have also been described in a few patients (Garcia-Minaur 2002; O'Donnell 1997; Rauch A 1999). Shaikh et al. (Shaikh 2000) stated that 22q11.2 LCRs share 97.98% nucleotide sequence identity. The size and the homology among them seem to be related to the frequency of each type of deletion.

As clinical variability is not explained by differences in gene content within the deletion, allelic variation(s) in the non-deleted homologous region is considered a possible contributor to phenotypic variability. Most of the genes from the 22q11.2 deletion region are expressed in fetal and adult brain, thus are candidates for both the psychiatric phenotype of patients with 22q11.2 deletions and susceptibility to psychiatric disorders in the general population (Meechan et al. 2010).

2. AIM and OUTLINE OF THE STUDY

2) AIM AND OUTLINE OF THE STUDY

Accumulating evidence from a decade of array-CGH demonstrated that the single model attributing disease phenotype to a single pathogenic CNV does not fit all cases. We can thus distinguish two types of genomic disorders: syndromic forms where the phenotypic features are largely invariant and fully penetrant such as Williams syndrome (MIM#194050) and Angelman (MIM#105830) or Prader-Willi syndrome (MIM#176270), and those where the same genomic rearrangement associates with a set of diagnoses of different severity or with a complete normal phenotype such as the 22q11.2 microdeletion syndrome. In this latter type of genomic disorders, there is growing appreciation that CNVs can be viewed as contributing to the pathogenesis of “recessive” diseases, rather than simply functioning as dominant variants with reduced penetrance. Alternatively, CNVs can be responsible of complex disorders such as obesity in association with multiple high-penetrant alleles of low frequency. To confirm these alternative explanations of phenotypic variability, I focused my thesis on the investigation of three different genomic rearrangements: the 22q11.2 microdeletion, the 16p11.2 microdeletion/duplication and the 10q11.22 deletion/duplication.

The 22q11.2 microdeletion is known to be associated with a variety of phenotypes including velocardiofacial syndrome, isolated cardiac defect, schizophrenia and Van den Ende-Gupta syndrome (MIM#600920). For the latter, the presence of a recessive allele unmasked by the deletion has been recently demonstrated. In order to identify additional recessive alleles we performed targeted sequencing on three patients with a 22q11.2 deletion and an atypical phenotype (MURCS, severe intellectual deficit with polydactyly and Cayler) in collaboration with the University of Geneva.

Deletions and duplications of chromosome 16p11.2 were already reported as associated with reduced penetrance with ASDs and schizophrenia, two complex traits at the opposite ends of a single spectrum of psychiatric phenotypes. Thanks to a collaborative effort among several Medical Genetics Units, we show that deletions and duplications on chromosome 16p11.2 could also have an impact on the body mass index.

To date, interstitial deletions involving 10q11.2 have been reported in over 40 patients with variable abnormal phenotype, individual with a normal phenotype and two prenatal cases. The only clinical feature common to the majority of subjects was ID/DD. We recently reported that a small duplication on 10q11.22 including *GPRIN2* gene, a regulator of neurite outgrowth, and *PPYR1*, a gene involved in energy homeostasis, is a candidate modifier for Rett syndrome (MIM#312750). In the present study we explored the association of CNVs at 10q11.22 with ASD and body mass index (BMI)

Moreover we reported two unrelated girls carrying a duplication of the Miller-Dieker region at 17p13.3. So far only few cases with this duplication have been reported. Molecular cytogenetic analyses show that in both patients the 17p duplication is the result of an unbalanced translocation and therefore the resulting phenotype is more complex. However we further delineate the features associated with this novel microdeletion syndrome.

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 referred to the Medical Genetics Unit of the University Hospital of Siena. All patients were evaluated by clinical geneticists.

3.2 Array-based CGH

3.2.1 Samples preparation

Genomic DNA of normal controls was obtained from Promega. Genomic DNAs were extracted from peripheral blood samples 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 (Sambrook 1989). 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 purified using dedicated columns (DNA Clean and Concentrator, Zymo research, CA92867-4619, USA) and the appropriate DNA concentrations were determined 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 estimated 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 - GRCh build 37, Feb 2009 (<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 pre-run 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 (Ariani 2004).

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 μ l, 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 μ l 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 (Koolen 2004).

4. RESULTS

4.1 Reciprocal duplication of known deletion syndrome

Manuscript in preparation

Duplication of the Miller-Dieker region (17p13.3): two cases as a result of unbalanced translocations.

Mucciolo M.¹, Lo Rizzo C.^{1,2}, Ferreri M.I.³, Coviello G.⁴, Fantasia D.⁵, Chiappini E.⁶, Pollazzon M.¹, Marozza A.^{1,2}, Mencarelli M.A.^{1,2}, Ariani F.¹, Meloni I.¹, Simi P.³, Giovannucci-Uzielli M.L.^{6,7}, Mari F.^{1,2}, Hayek G.⁸, Renieri A.^{1,2}.

1 Medical Genetics, University of Siena, Siena, Italy

2 Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy

3 Cytogenetics and Molecular Genetic Unit, AOU Pisana, Ospedale Santa Chiara, Pisa, Italy

4 Radiology, Azienda Ospedaliera Universitaria Senese, Siena, Italy

5 Department of Oral Sciences, Nano and Biotechnologies/Section of Medical Genetics, University of Chieti, Italy

6 Firenze Dipartimento di Scienze della Salute della Donna e del Bambino, Università di Firenze, Firenze, Italy

7 Genetic Science, Piazza Savonarola 11, Firenze, Italy

8 Child Neuropsychiatry, Azienda Ospedaliera Senese, Siena, Italy

Corresponding author: Alessandra Renieri M.D., Ph.D.

ABSTRACT

Duplications of the Miller-Dieker region at 17p13.3 and involving the *PAFAH1B1* gene have been recently reported only in few cases so far. These cases were mostly due to de novo events. We report two unrelated girls carrying this duplication who exhibited intellectual deficit, microcephaly and facial dysmorphisms. Molecular cytogenetic analyses show that in both patients the 17p duplication is the result of an unbalanced translocation involving two different chromosomes: 9p24.2 in one case and 10q26.2 in the other. The facial features of our patients closely resemble those

previously reported, indicating that 17p13.3 duplication causes a quite distinctive facial phenotype. The overall phenotype of these two cases is complicated by the presence of a second copy number variation, and some phenotypic features of our patients can be attributed either to 9p deletion or 10q deletion. Overall, these new cases indicate that the 17p13.3 microduplication may be more frequent than thought and originates not only from de novo events. Moreover, we confirm the absence of gross anomalies of brain morphology in cases with *PAFAH1B1* gene duplications with respect to the *PAFAH1B1* deletion, as in Miller-Dieker syndrome.

KEYWORDS: 9pter deletion, 10qter deletion, 17pter duplication, *PAFAH1B1*, array-CGH, and Miller-Dieker syndrome.

INTRODUCTION

The use of array comparative genome hybridization (array-CGH) analysis for the investigation of children with intellectual disability (ID) has allowed the identification of numerous new microdeletion and microduplication syndromes, some of which have been clinically well characterized. Most of these rearrangements are the result of non-allelic homologous recombination between region of low copy repeats (LCRs) [1].

The short arm of chromosome 17 is prone to copy number variations (CNVs) due to a high density of LCRs [2]. Most of the rearrangements harboured on chromosome 17p lead to specular syndrome: the Charcot Marie Tooth syndrome type 1A due to a duplication in 17p12 and the Hereditary Neuropathy with liability to Pressure Palsies due to the reciprocal deletion [Chance et al., 1994; Reiter et al., 1996][3, 4]; the Smith-Magenis syndrome and the Potocki-Lupski syndrome due to a deletion and a duplication of the 17p11.2 region, respectively [5, 6]. Terminal deletions of chromosome 17p are associated with isolated lissencephaly when they include the *PAFAH1B1* gene, or with Miller-Dieker syndrome (MDS) when the 17p deletions also include the *YWHAE* gene [7-9].

Recently, isolated 17p13.3 duplications involving the *PAFAH1B1* gene have been reported in seven patients with psychomotor retardation, hypotonia and dysmorphic features without lissencephaly or gross brain malformations [10-12]. The phenotype

of transgenic mice conditionally overexpressing *PAFAH1B1* was indeed characterized by decreased brain size and neuronal migration abnormalities [10]. Bruno and colleagues identified two classes of co-locating microduplications in 17p13.3: class I duplications including *YWHAE* but not *PAFAH1B1*; and class II duplications always including *PAFAH1B1*, and sometimes including the genomic region encompassing the *CRK* and *YWHAE* genes [11]. Class I microduplications are associated with intellectual disability (ID), subtle dysmorphic facial features, subtle hand/foot malformations, and a tendency toward postnatal overgrowth [11]. Class II microduplications recently have been shown to be associated with mild to moderate ID and hypotonia. Some dysmorphic features, such as prominent forehead and pointed chin, are shared with class I duplications, while overgrowth, behavioural problems and hand/foot abnormalities are less often noted.

A complex rearrangement including the 17p13.3 microduplication has been reported in association with a second CNV in two cases. The rearrangement originated in a balanced translocation present in a parent [t(9;17) and t(X;17)] [13, 14].

The known 9p deletion syndrome was first described by Alfi *et al.* in 1973 [15]. This is an heterogeneous condition with variable deletion size characterized by ID, congenital malformations including trigonocephaly, congenital heart defect, anorectal and genital anomalies and dysmorphisms [16-19]. The critical region for the 9p deletion syndrome has been located between bands p22.3 and p24.1 [19]. The deletions of the more terminal part of chromosome 9p are rarer and some of them coexist in the same patient together with larger rearrangements in other chromosomes [20, 14, 21, 22]. Patients with deletions involving the 9p24.3 band show male to female sex reversal, possibly due to *DMRT1* and *DMRT2* haploinsufficiency [23, 24].

Terminal deletion of the long arm of chromosome 10 is a relatively frequent cytogenetic abnormality with clinical heterogeneity even among members of the same family [25]. Characteristic features of 10q deletion syndrome include peculiar facial features, cardiac and urogenital anomalies and neurodevelopmental deficit [26]. The critical region of the 10q deletion syndrome corresponds to a segment of ~600 Kb in 10q26.2 encompassing two genes, *DOCK1* and *C10orf90* [27].

This report describes two additional cases with a reciprocal duplication of the MDS region, suggesting that this condition may be less rare than previously thought. Both cases are the unbalanced result of two different balanced translocations: t(9;17)(p24.2;p13.3) and t(10;17)(q26.2;p13.2) and, therefore, their phenotypes are more complex than those of cases with isolated 17p13.3 microduplications. Nevertheless, the core phenotype of the 17p13.3 duplication is recognizable. This paper reviews the literature on the 17p13.3 region and further delineates the features associated with this novel microduplication syndrome.

PATIENTS AND METHODS

Patients

Written informed consent was obtained from the guardians of the patients included in this study. Participation in the study did not alter the standard of care.

Clinical Reports

Patient 1 is a 13 years and 5 months old girl, the second child of healthy unrelated parents (Fig.1a and Table 1). The girl was born at the 38th week of pregnancy by caesarean section due to fetal sufferance. Her birth weight was 3150 gr (50-75th percentile) and her length was 49 cm (50-75th percentile). Apgar score and head circumference (OFC) measurements were not available. She presented feeding difficulties and gastro esophageal reflux. The patient exhibited severe developmental delay. She never acquired sphincter control. She frequently suffered from respiratory infections during childhood. She presented drug-resistant epilepsy from the age of six months. Brain MRI performed at 10 years of age showed marked dilatation of the supratentorial ventricles and dilatation of the cisterna magna possibly due to a leptomenigeal cyst. Cardiac ultrasounds showed atrial septal defect and ventriculomegaly. Abdominal ultrasound was normal. No other major abnormalities were present. Physical examination at 12y1m (Fig.1b) showed: height, 128 cm (<<3rd percentile), weight, 27 kg (<<3rd percentile), microcephaly (OFC of 50 cm; <<3rd percentile), triangular face with pointed chin, upslanting palpebral fissures, sparse and V-shaped eyebrows, open mouth with protruding tongue, sialorrhea,

prominent nose, scoliosis and flat feet. The patient was able to walk independently, exhibited hand stereotypes, and was able to grasp. She also showed hyperactivity and continuously tried to catch the attention of the people around her. Standard karyotype from peripheral blood lymphocytes was normal.

Patient 2 is a 15 years and 4 months old girl, the second child of healthy unrelated parents (Fig.1c and Table 1). The mother had two spontaneous miscarriages in the first month of gestation. At the time of her birth, Patient 2's mother and father were 26 and 29 years old, respectively. The proband had a healthy older brother and two maternal cousins referred with psychomotor delay (not available for testing). The girl was born after a prolonged labour at term of an uneventful pregnancy. At birth, weight was 3300 gr (50th percentile) and length was 51 cm (50-75th percentile). Apgar score and OFC measurements were not available. A pale haemangioma of the forehead was observed. Patient 2 showed developmental delay: she began to sit alone at 1.5 year, crawled at 2 years, began to walk independently at 2.5 years, and said the first words at 5 years. She never acquired sphincter control and frequently suffered from respiratory infections during childhood. At 4 years the patient was surgically treated for strabismus. A radiological examination of skeletal development of the left-hand wrist showed mild bone-age delay (chronological age 5 years and 8 months, bone-age corresponding to 5 years and 1 month). A radiological survey of hands and feet performed at 11 years and 6 months showed aplasia of a phalanx of the fifth finger of both feet and a medial notch of the second phalanx of II finger of the left hand. Repeated EEGs were alternatively normal or showed a mild disorganization of the deep rhythm. Results of ophthalmological evaluation were normal except for mild myopia (-1.25/-1.50 diopters). A pelvic ultrasound showed mild irregularities of the morphology of the uterus. The following investigations were normal: abdominal and cardiac ultrasound, brain MRI and karyotype. Physical examination of Patient 2 at 11y1m (Fig.1d) demonstrated normal height (145 cm; 25-50th percentile) and weight (40 kg; 50-75th percentile), microcephaly (OFC of 48 cm; <<3rd percentile), triangular face, with pointed chin, synophrys, thickening in the medial part and V-shaped eyebrows, open mouth, high and narrow palate, and hypoplastic 5th toe, more evident on the right side. The patient showed ataxic gait, rocking of the trunk in upright position, unmotivated laughter and sialorrhea. At the

time of our examination Patient 2 had just begun to formulate sentences, always spoke to catch attention, displayed hyperactivity, and brought all objects to her mouth. Patient 2's mother exhibited isolated microcephaly (OFC 52 cm, <3rd percentile) and normal height (169 cm; 75-90th percentile).

Array-CGH analysis

Array-CGH analysis was performed using commercially available oligonucleotide microarrays containing about 44,000 60-mer probes (Human Genome CGH Microarray 44B Kit, Agilent Technologies, Santa Clara, California) according to the manufacturer's instructions and as previously reported [28]. The average spatial resolution of the probes was about 45 kb. Probe locations were assigned according to UCSC Genome Browser, GRCh37/hg19, Feb 2009 (<http://genome.ucsc.edu>).

Multiplex Ligation-dependant Probe Amplification (MLPA) analysis

We used a distinct commercially available MLPA kit, the SALSA P036D subtelomeric primer set (MRC-Holland, Amsterdam, The Netherlands). This kit contains oligonucleotide primer sets specific for the amplification of selected loci in the subtelomeric regions of all chromosome arms, except for the acrocentric chromosomes 13, 14, 15, 21 and 22 that effectively lack a short arm. For the latter, the manufacturer has included in this kit primer sets specific for loci adjacent to the centromere in the long arm of the acrocentric chromosomes, referred to as the 'acrocentric' primer. This kit was previously validated in other laboratories (data not shown) on series of patients with known subtelomeric ultra conserved regions (UCRs) [29, 30]. The target loci of this kit represent known functional genes or protein coding sequences. Each experiment was carried out according to the manufacturer's instructions.

Fluorescent *in situ* hybridization (FISH) analysis

Chromosomal preparations for the analysis were obtained according to standard techniques. FISH was performed with TelVision 9p and 17p probes (Vysis). Each experiment was carried out according to the manufacturer's instructions.

RESULTS

In Patient 1 array-CGH analysis detected the presence of two telomeric rearrangements: a ~6.9 Mb terminal deletion of chromosome 10 [arr 10q26.2q26.3(128,467,040-135,404,471)x1] and a ~5.5 Mb terminal duplication of chromosome 17 [arr 17p13.3p13.2(48,539-5,514,628)x3] (Fig. 2 a and b). Patient 2 had a ~4.4 Mb deletion on chromosome 9 [arr 9p24.3p24.2(204,193-4,600,751)x1] and a ~3 Mb duplication on chromosome 17 [arr 17p13.3(48,538-3,058,821)x3] (Fig.3 a and b). The array-CGH analysis also revealed a 50 Kb duplication in Xq28 [arr Xq28(148,690,284-148,728,581)x3 mat] in the proband and her mother, already reported in healthy individuals and thus probably not associated with a phenotype [31] (data not shown).

MLPA analysis confirmed the presence of a duplication of the area containing the *RPH3AL* probe on chromosome 17p13.3 in both patients, a deletion of the *PAOX* probe on chromosome 10q26.3 in Patient 1, and a deletion of the *DMRT1* probe on chromosome 9p24.3 in Patient 2 (data not shown).

FISH analysis of the parents of both patients, performed using telomeric probes for chromosomes 10 and 17 in family 1 and probes for chromosomes 9 and 17 in family 2, revealed a balanced translocation in Patient 1's father and Patient 2's mother (data not shown). Given the presence of microcephaly in the otherwise healthy mother of Patient 2, we also performed array-CGH analysis on the mother's DNA, to ascertain if the translocation was balanced. The analysis revealed no gains or losses at both breakpoints (data not shown).

DISCUSSION

These two cases, together with the others previously reported in the literature, concur to define the 17q13.3 microduplication syndrome [10, 11, 13, 14, 12]. We excluded from the analysis cases with class I microduplications centred on the *YWHAE* gene (cases 1, 2, 3 and 4 from Bi [10], and cases 9 and 11 from Bruno [11]), and we considered only cases with class II microduplications. All patients except two reported by Bi *et al.* [10] (patient 5 and 6), have large duplications, including both *PFAH1B1* and *YWHAE* genes (Fig.4).

When we compared our cases with those previously described (Tab. I), it appeared that the duplicated patients had many facial features in common. All patients, even at a younger age, had pointed chin; this characteristic was augmented in adolescence resulting in the triangular shape of the face (8/8). In all older patients a high nasal bridge became evident and most of them (7/9) had V shaped eyebrows.

Concerning the physical phenotype, microcephaly, or at least deceleration of head growth, was a consistent sign (9/10). Concurrently, some patients had a progressive reduction in height (3/8) and weight (4/8) growth. According to our review, degree of ID was variable, ranging from mild to severe (6/9). Very interestingly, we found that recurrent respiratory infections during childhood were reported in 7 patients. This characteristic was not emphasized previously in this microduplication syndrome. From a clinical point of view, the frequent respiratory infections, together with the deceleration of growth, could be used as an additional diagnostic handle of the syndrome.

Previous studies showed that transgenic mice over-expressing *PAFAH1B1* showed migration defect and reduced brain volume [10]. The latter sign is also present in humans, since most patients (9/10) showed microcephaly. On the other hand, neuronal migration defects were not detected by neuroimaging studies (Tab.1). Therefore, our data confirm that *PAFAH1B1* over-expression in humans does not cause neuronal migration defects or other gross brain malformations.

Since the 17p duplications of our patients originated from two unbalanced translocations, some of the clinical features can also be explained by 10q and 9p deletions. Terminal deletions of long arm of chromosome 10 are associated with broad/prominent nasal bridge, prominent nose, strabismus, thin upper lip, and fifth finger clinodactyly. Postnatal low weight and microcephaly are occasionally found [32]. The very prominent and abnormal shaped nose of Patient 1 may be the result of the combined effect of both 17p duplications and 10q deletion. The 10q deletion of Patient 1 includes the *DOCK1* gene, contributing to ID in 10q- syndrome [27] and *DPYSL4* (or *CRMP3*), a critical factor regulating dendrite arborization and spine morphology in the hippocampus [33].

Deletions of the terminal portion of the short arm of chromosome 9 are associated with ID due to *DOCK8* haploinsufficiency [34, 35] and a male to female sex

reversal, possibly due to *DMRT1* and *DMRT2* haploinsufficiency [23]. Although in female patients no urogenital anomalies are reported, we cannot completely rule out the hypothesis that the mild abnormal morphology of the uterus reported in our patient could be due to haploinsufficiency of the 9p region. Therefore, more accurate gynaecologic evaluation in the proband could be useful.

The rearrangements present in our patients originated from a balanced translocation present in a parent as demonstrated by FISH analysis. In family 2, the mother presented isolated microcephaly with normal intellectual functioning, and experienced two spontaneous miscarriages in the first month of gestation. In addition, the family history revealed that, two maternal cousins of the proband suffered from psychomotor delay. All these data indicated a segregation of the translocation in the maternal branch of the family. A similar translocation was previously reported by *Kohler* et al [14], in a family with two siblings showing an unbalanced translocation t(9;17)(p24.2;p13.3) that had originated from a balanced translocation present in the mother (Tab. 1). Family history also highlighted two spontaneous miscarriages and recurrent neonatal deaths; two of the fetuses showed the typical signs of MDS. The authors ascribed to the 17p deletion all of the early deaths in the family [14]. The same explanation may be given for the miscarriages reported in the families reported in this study.

The presence of microcephaly in both Patient 2 and her mother led us to consider disrupted genes at the breakpoints as possible candidate causes of microcephaly. The breakpoint at chromosome 17 did not disrupt genes, while the breakpoint at chromosome 9 interrupted the *C9orf68* gene, which has a sequence homology to *SPATA6*, encoding for a spermatogenesis-associated protein 6 precursor. A dosage alteration of genes located near the breakpoints due to a positional effect cannot be excluded as a possible cause for the microcephaly present both in the patient and her mother.

Overall, these new cases suggest that the 17p13.3 microduplication may be more frequent than thought. Our results confirm the absence of gross anomalies of brain morphology in cases with *PAFAH1B1* gene duplications in contrast to its haploinsufficiency. In the two cases reported here, the phenotype is more complex resulting from the combined effect of the duplication of the region involved in the

MDS and of a second CNV. Nevertheless, the core phenotype of the 17p13.3 duplication is recognizable and consists of V-shaped eyebrows, prominent nose, a high nasal bridge, a pointed chin evolving in a triangular face, decreased growth of the head, decreased height and weight, and recurrent infections.

ACKNOWLEDGEMENTS

We thank the biobank “*Cell lines and DNA bank of Rett syndrome, X linked Mental Retardation and other genetic diseases*” supported by Telethon grant GTB07001 to A.R. All authors disclose any potential sources of conflict of interest.

REFERENCES

- [1] Lupski J R, Stankiewicz P. (2005) Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet* 1:e49
- [2] Stankiewicz P, Shaw C J, Dapper J D, Wakui K, Shaffer L G, Withers M, Elizondo L, Park S S, Lupski J R. (2003) Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am J Hum Genet* 72:1101-16
- [3] Chance P F, Abbas N, Lensch M W, Pentao L, Roa B B, Patel P I, Lupski J R. (1994) Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Hum Mol Genet* 3:223-8
- [4] Reiter L T, Murakami T, Koeuth T, Pentao L, Muzny D M, Gibbs R A, Lupski J R. (1996) A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. *Nat Genet* 12:288-97
- [5] Chen K S, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault A C, Lee C C, Lupski J R. (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat Genet* 17:154-63
- [6] Potocki L, Chen K S, Park S S, Osterholm D E, Withers M A, Kimonis V, Summers A M, Meschino W S, Anyane-Yeboah K, Kashork C D, Shaffer L G, Lupski J R. (2000) Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 24:84-7
- [7] Mei D, Lewis R, Parrini E, Lazarou L P, Marini C, Pilz D T, Guerrini R. (2008) High frequency of genomic deletions--and a duplication--in the LIS1 gene in lissencephaly: implications for molecular diagnosis. *J Med Genet* 45:355-61
- [8] Reiner O, Carrozzo R, Shen Y, Wehnert M, Faustinella F, Dobyns W B, Caskey C T, Ledbetter D H. (1993) Isolation of a Miller-Dieker lissencephaly gene containing G protein beta-subunit-like repeats. *Nature* 364:717-21
- [9] Toyo-oka K, Shionoya A, Gambello M J, Cardoso C, Leventer R, Ward H L, Ayala R, Tsai L H, Dobyns W, Ledbetter D, Hirotsune S, Wynshaw-Boris A. (2003) 14-3-3epsilon is important for neuronal migration by binding to NUDEL: a molecular explanation for Miller-Dieker syndrome. *Nat Genet* 34:274-85
- [10] Bi W, Sapir T, Shchelochkov O A, Zhang F, Withers M A, Hunter J V, Levy T, Shinder V, Peiffer D A, Gunderson K L, Nezarati M M, Shotts V A, et al. (2009)

Increased LIS1 expression affects human and mouse brain development. *Nat Genet* 41:168-77

[11] Bruno D L, Anderlid B M, Lindstrand A, van Ravenswaaij-Arts C, Ganesamoorthy D, Lundin J, Martin C L, Douglas J, Nowak C, Adam M P, Kooy R F, Van der Aa N, et al. Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *J Med Genet* 47:299-311

[12] Roos L, Jonch A E, Kjaergaard S, Taudorf K, Simonsen H, Hamborg-Petersen B, Brondum-Nielsen K, Kirchhoff M. (2009) A new microduplication syndrome encompassing the region of the Miller-Dieker (17p13 deletion) syndrome. *J Med Genet* 46:703-10

[13] Hyon C, Marlin S, Chantot-Bastarud S, Mabboux P, Beaujard M P, Al Ageeli E, Vazquez M P, Picard A, Siffroi J P, Portnoi M F. A new 17p13.3 microduplication including the *PAFAH1B1* and *YWHAE* genes resulting from an unbalanced X;17 translocation. *Eur J Med Genet* 54:287-91

[14] Kohler A, Hain J, Muller U. (1994) Familial half cryptic translocation t(9;17). *J Med Genet* 31:712-4

[15] Alfi O, Donnell G N, Crandall B F, Derencsenyi A, Menon R. (1973) Deletion of the short arm of chromosome no.9 (46,9p-): a new deletion syndrome. *Ann Genet* 16:17-22

[16] Christ L A, Crowe C A, Micale M A, Conroy J M, Schwartz S. (1999) Chromosome breakage hotspots and delineation of the critical region for the 9p-deletion syndrome. *Am J Hum Genet* 65:1387-95

[17] Hauge X, Raca G, Cooper S, May K, Spiro R, Adam M, Martin C L. (2008) Detailed characterization of, and clinical correlations in, 10 patients with distal deletions of chromosome 9p. *Genet Med* 10:599-611

[18] Huret J L, Leonard C, Forestier B, Rethore M O, Lejeune J. (1988) Eleven new cases of del(9p) and features from 80 cases. *J Med Genet* 25:741-9

[19] Swinkels M E, Simons A, Smeets D F, Vissers L E, Veltman J A, Pfundt R, de Vries B B, Faas B H, Schrandt-Stumpel C T, McCann E, Sweeney E, May P, et al. (2008) Clinical and cytogenetic characterization of 13 Dutch patients with deletion 9p syndrome: Delineation of the critical region for a consensus phenotype. *Am J Med Genet A* 146A:1430-8

[20] Brisset S, Kasakyan S, L'Hermine A C, Mairovitz V, Gautier E, Aubry M C, Benkhalifa M, Tachdjian G. (2006) De novo monosomy 9p24.3-pter and trisomy 17q24.3-qter characterised by microarray comparative genomic hybridisation in a fetus with an increased nuchal translucency. *Prenat Diagn* 26:206-13

[21] Repetto G M, Wagstaff J, Korf B R, Knoll J H. (1998) Complex familial rearrangement of chromosome 9p24.3 detected by FISH. *Am J Med Genet* 76:306-9

[22] Saha K, Lloyd I C, Russell-Eggitt I M, Taylor D S. (2007) Chromosomal abnormalities and glaucoma: a case of congenital glaucoma associated with 9p deletion syndrome. *Ophthalmic Genet* 28:69-72

[23] Barbaro M, Balsamo A, Anderlid B M, Myhre A G, Gennari M, Nicoletti A, Pittalis M C, Oscarson M, Wedell A. (2009) Characterization of deletions at 9p affecting the candidate regions for sex reversal and deletion 9p syndrome by MLPA. *Eur J Hum Genet* 17:1439-47

[24] Muroya K, Okuyama T, Goishi K, Ogiso Y, Fukuda S, Kameyama J, Sato H, Suzuki Y, Terasaki H, Gomyo H, Wakui K, Fukushima Y, Ogata T. (2000) Sex-

determining gene(s) on distal 9p: clinical and molecular studies in six cases. *J Clin Endocrinol Metab* 85:3094-100

[25] Petit P, Devriendt K, Azou M, Gewillig M, Fryns J P. (1998) Terminal deletion of chromosome 10q26: delineation of two clinical phenotypes. *Genet Couns* 9:271-5

[26] Mulcahy M T, Pemberton P J, Thompson E, Watson M. (1982) Is there a monosomy 10qter syndrome? *Clin Genet* 21:33-5

[27] Yatsenko S A, Kruer M C, Bader P I, Corzo D, Schuette J, Keegan C E, Nowakowska B, Peacock S, Cai W W, Peiffer D A, Gunderson K L, Ou Z, Chinault A C, Cheung S W. (2009) Identification of critical regions for clinical features of distal 10q deletion syndrome. *Clin Genet* 76:54-62

[28] Pescucci C, Caselli R, Grosso S, Mencarelli M A, Mari F, Farnetani M A, Piccini B, Artuso R, Bruttini M, Priolo M, Zuffardi O, Gimelli S, Balestri P, Renieri A. (2007) 2q24-q31 deletion: report of a case and review of the literature. *Eur J Med Genet* 50:21-32

[29] Ahn J W, Ogilvie C M, Welch A, Thomas H, Madula R, Hills A, Donaghue C, Mann K. (2007) Detection of subtelomere imbalance using MLPA: validation, development of an analysis protocol, and application in a diagnostic centre. *BMC Med Genet* 8:9

[30] Rooms L, Reyniers E, van Luijk R, Scheers S, Wauters J, Ceulemans B, Van Den Ende J, Van Bever Y, Kooy R F. (2004) Subtelomeric deletions detected in patients with idiopathic mental retardation using multiplex ligation-dependent probe amplification (MLPA). *Hum Mutat* 23:17-21

[31] Sanlaville D, Prieur M, de Blois M C, Genevieve D, Lapierre J M, Ozilou C, Picq M, Gosset P, Morichon-Delvallez N, Munnich A, Cormier-Daire V, Baujat G, et al. (2005) Functional disomy of the Xq28 chromosome region. *Eur J Hum Genet* 13:579-85

[32] Courtens W, Wuyts W, Rooms L, Pera S B, Wauters J. (2006) A subterminal deletion of the long arm of chromosome 10: a clinical report and review. *Am J Med Genet A* 140:402-9

[33] Quach T T, Massicotte G, Belin M F, Honnorat J, Glasper E R, Devries A C, Jakeman L B, Baudry M, Duchemin A M, Kolattukudy P E. (2008) CRMP3 is required for hippocampal CA1 dendritic organization and plasticity. *Faseb J* 22:401-9

[34] Griggs B L, Ladd S, Saul R A, DuPont B R, Srivastava A K. (2008) Deducator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. *Genomics* 91:195-202

[35] Ruusala A, Aspenstrom P. (2004) Isolation and characterisation of DOCK8, a member of the DOCK180-related regulators of cell morphology. *FEBS Lett* 572:159-66

Table I. Summary of clinic features of patients with 17p13.3 microduplication

	Patient 1 (Roos et al)	Patient 2 (Roos et al)	Patient 3 (Roos et al)	Patient 5 (Bi et al)	Patient 6 (Bi et al)	Patient 7 (Bi et al)	Sibling 1 (Kohler et al)	Sibling 2 (Kohler et al)	Patient 10 (Bruno et al)	Patient (Hyon et al)	Present case 1	Present case 2
Sex	M	F	M	M	M	F	M	M	M	F	F	F
Age	14y	28m	22m	32m	17y4m	10y5m	20y	17y	6y6m	13y	13y4m	15y4m
Duplication size (Mb)	1,8	3	4	0,151 + 0,58	063 (0,16 triplication)	3,6	NA	NA	2,07	4.2	5,5	3
Inheritance	De Novo	De Novo	De Novo	Maternally Inherited	De Novo	De Novo	Originated from a balanced translocation in the mother	Originated from a balanced translocation in the mother	De novo	NA	Originated from a balanced translocation in the father	Originated from a balanced translocation in the mother
Normal birth auxological parameters	Yes	Yes	Yes	Yes	NA	Yes	NA	NA	Yes	Yes	Yes	Yes
Deceleration of head growth	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	NA	NA	Yes	Yes
Deceleration of height	No	No	No	Yes	Yes	No	NA	NA	NA	NA	Yes	No
Deceleration of weight	No	Yes	No	Yes	Yes	No	NA	NA	NA	NA	Yes	No
High nasal bridge after childhood	Yes	NA	NA	NA	Yes	NA	Yes	Yes	NA	No	Yes	Yes
Pointed chin	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes
Triangular face in older patients	Yes	NA	NA	NA	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes
V shaped eyebrows	Yes	Yes	Yes	No	Yes	NA	No	Yes	NA	NA	Yes	Yes
Level of ID at school age and after	Mild	NA	NA	NA	Mild to moderate	Mild	Severe	Severe	Normal	Mild	Moderate to Severe	Mild
Brain imaging results	Normal	Normal	Dilated lateral ventricles and CC agenesis	NA	Right subaracnoid cyst and gross dysgenesis of CC (especially affecting the splenium), cerebellar atrophy and mild cerebral volume loss	Thinning of the splenium of the CC and mild cerebellar volume loss	NA	Normal	NA	Normal	Marked dilatation of the supratentorial ventricles: dilatation of the cisterna magna possibly due to leptomeningeal cyst	Normal
Recurrent upper airway infections	Yes	Yes	Yes	NA	Yes	NA	NA	NA	Yes	NA	Yes	Yes

NA, not available or not applicable; CC, corpus callosum

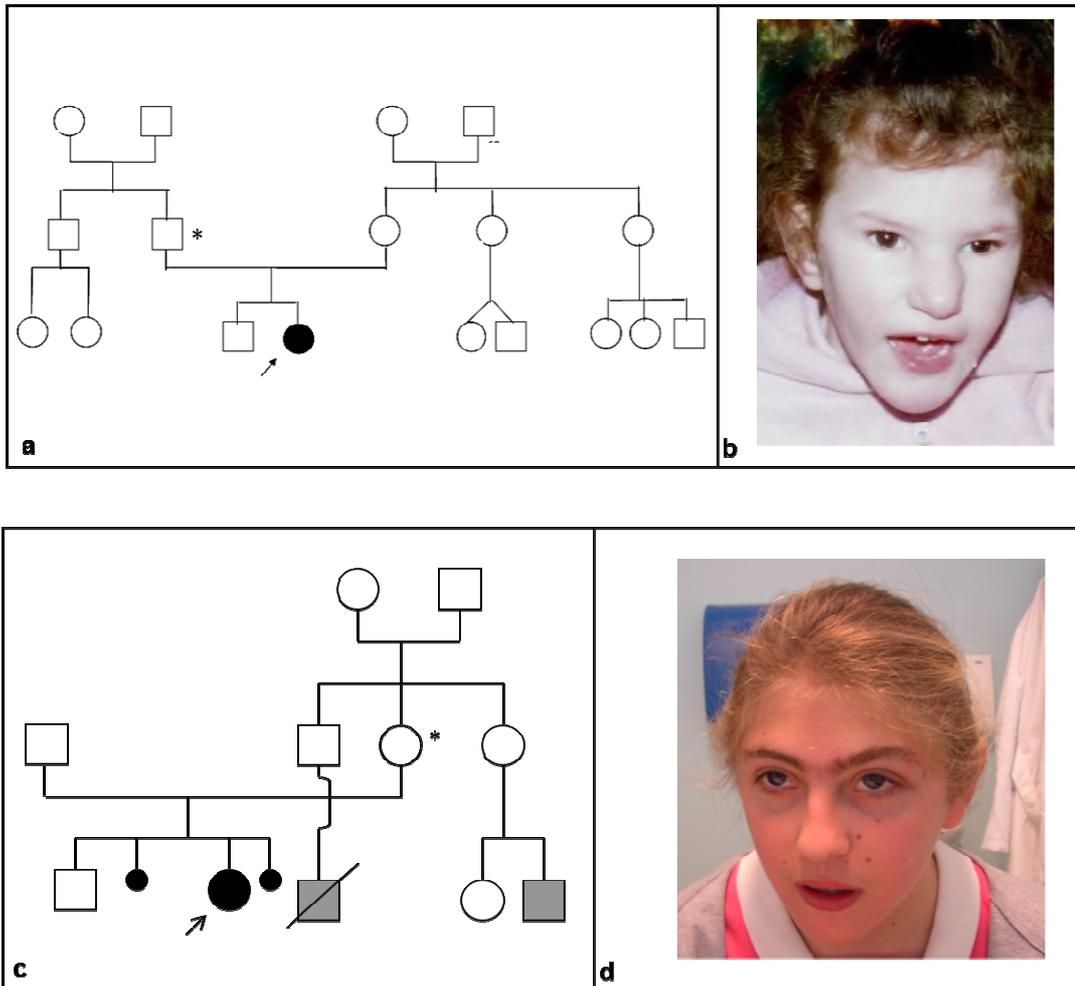


Fig.1. Pedigree (a, c) and pictures (b, d) of both patients. a) Pedigree of Patient 1. b) Frontal view of patient 1 at the age of 12 years and 1 month showing prominent nose with high and broad nasal bridge, open mouth and triangular face. c) Pedigree of Patient 2. Grey symbols refer to the two cousins with ID. d) Frontal view of Patient 2 at the age of 11 years and 6 months showing V-shaped eyebrows with synophris, high nasal bridge and triangular face. An asterisk in both pedigrees indicates carriers of the balanced translocation. Arrows indicate the patient.

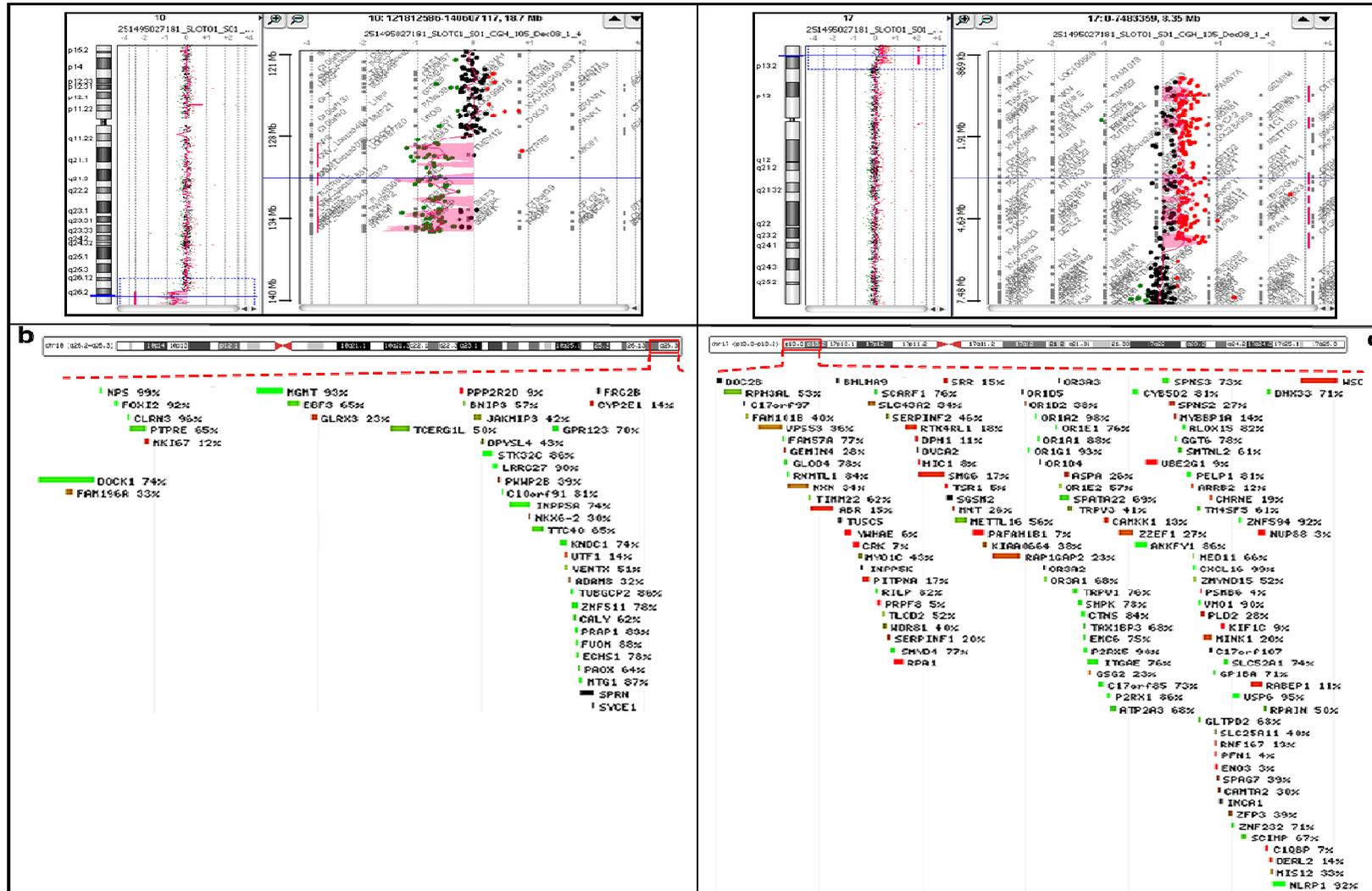


Fig.2. Array-CGH result of Patient 1. Array-CGH ratio profile of chromosome 10 (a) and 17 (c). On the left of each panel is illustrated the chromosome ideogram, and on the right, the log2 ratio of chromosome probes plotted as a function of chromosomal position. Image from DECIPHER database (b, d) showing the genes in the rearranged regions.

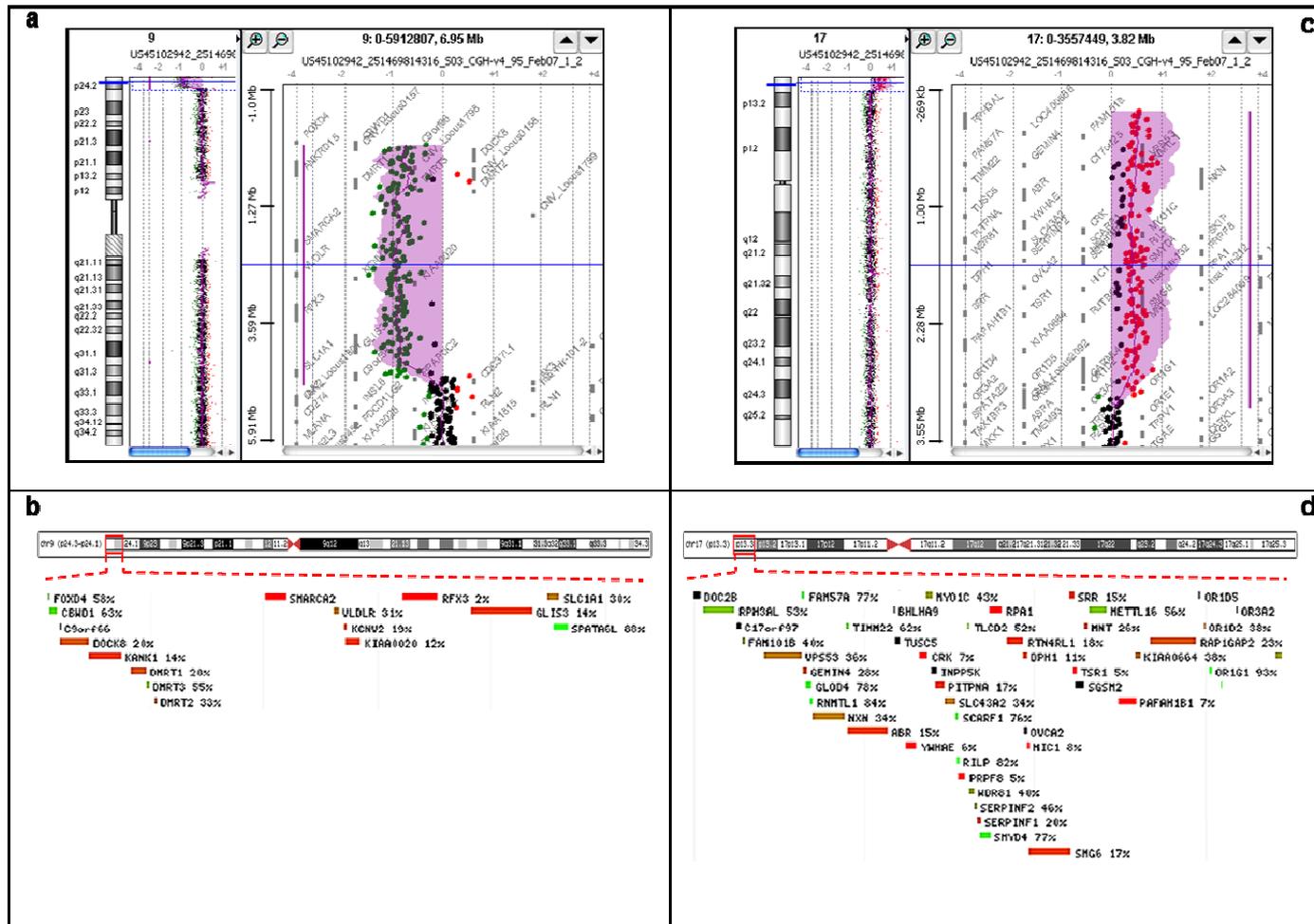


Fig.3 Array-CGH result of Patient 2. Array-CGH ratio profile of chromosome 9 (a) and 17 (c). On the left of each panel is depicted the chromosome ideogram, and on the right, the log2 ratio of chromosome probes plotted as a function of chromosomal position. Image from DECIPHER database (b, d) showing the genes in the rearranged regions.

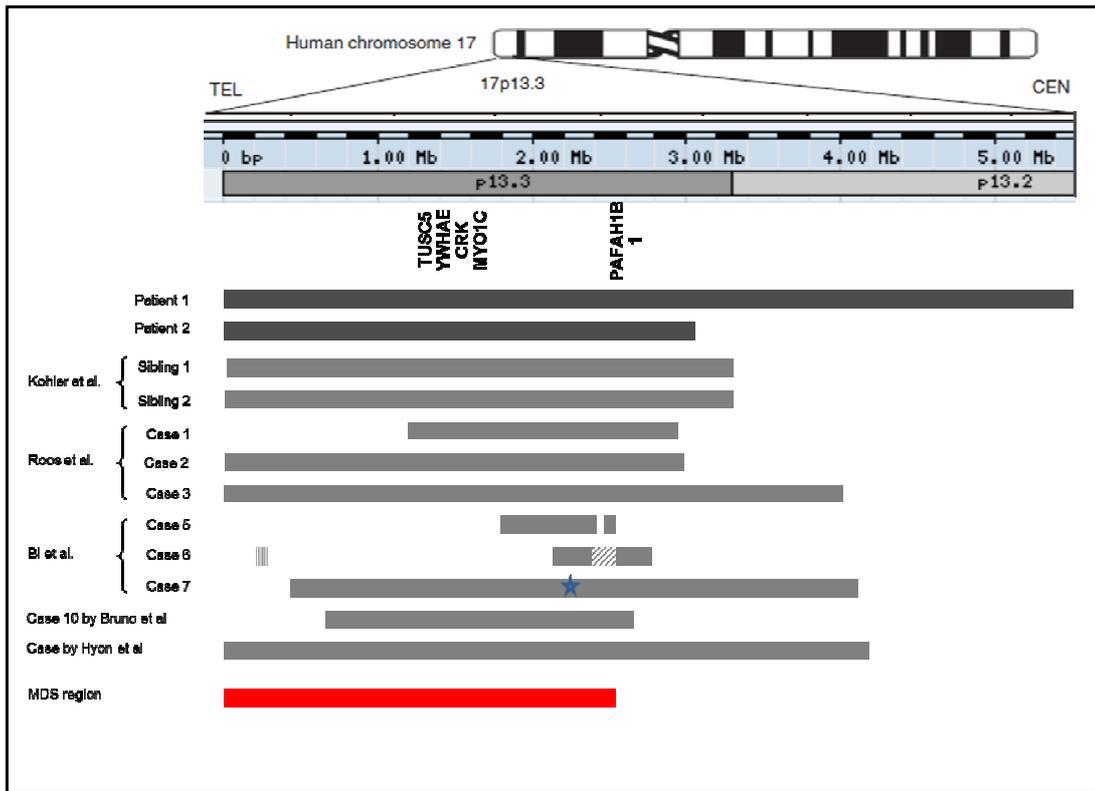


Fig.4. The extent of the duplicated area in the two patients compared to the patients reported in the literature. Upper panel: ideogram of chromosome 17. Lower panel: dark grey bars show duplicated region in Patients 1 and 2 reported in this study. Light grey bars show duplications of patients reported by: Kohler [Kohler et al., 1994], these two cases were not characterized at molecular level but only cytogenetically (dotted lines indicate undefined breakpoints); Roos [Roos et al., 2009], Bi (vertical lines indicate a deletion, diagonal lines a triplication, and the asterisk indicates a small 4Kb deletion) [Bi et al., 2009], Bruno [Bruno et al., 2010] and Hyon [Hyon et al., 2011].

4.2 Microdeletion and microduplication in 16p11.2

Nature. 2011 Aug 31;478(7367):97-102

Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus

S. Jacquemont*, A. Reymond*, F. Zufferey, L. Harewood, R.G. Walters, Z. Kutalik, D. Martinet, Y. Shen, A. Valsesia, N.D. Beckmann, G. Thorleifsson, M. Belfiore, S. Bouquillon, D. Champion, N. de Leeuw, B.B.A. de Vries, T. Esko, B.A. Fernandez, F. Fernandez-Aranda, J.M. Fernandez-Real, M. Gratacòs, A. Guilmatre, J. Hoyer, M.R. Jarvelin, R.F. Kooy, A. Kurg, C. Le Caignec, K. Mannik, O.S. Platt, D. Sanlaville, M.M. Van Haelst, S. Villatoro Gomez, F. Walha, B. Wu, Y. Yu, A. Aboura, M.C. Addor, Y. Alembik, S.E. Antonarakis, B. Arveiler, M. Barth, N. Bednarek, F. Bena, S. Bergmann, M. Beri, L. Bernardini, B. Blaumeiser, D. Bonneau, A. Bottani, O. Boute, H.G. Brunner, D. Cailley, P. Callier, J. Chiesa, J. Chrast, L. Coin, C. Coutton, J.M. Cuisset, J.C. Cuvellier, A. David, B. de Freminville, B. Delobel, M.A. Delrue, B. Demeer, D. Descamps, G. Didelot, K. Dieterich, V. Disciglio, M. Doco-Fenzy, S. Drunat, B. Duban-Bedu, C. Dubourg, J.S. El-Sayed Moustafa, P. Elliott, B.H.W. Faas, La. Faivre, A. Faudet, F. Fellmann, A. Ferrarini, R. Fisher, E. Flori, L. Forer, D. Gaillard, M. Gerard, C. Gieger, S. Gimelli, G. Gimelli, H.J. Grabe, A. Guichet, O. Guillin, A.L. Hartikainen, D. Heron, L. Hippolyte, M. Holder, G. Homuth, B. Isidor, S. Jaillard, Keren, A. Kloss-Brandstatter, N.V.A.M. Knoers, D.A. Koolen, P.M. Kroisel, F. Kronenberg, A. Labalme, E. Landais, E. Lapi, V. Layet, S. Legallic, B. Leheup, B. Leube, S. Lewis, J. Lucas, K.D. MacDermot, P. Magnusson, C. Marshall, M. Mathieu-Dramard, M.I. McCarthy, T. Meitinger, M.A. Mencarelli, G. Merla, A. Moerman, V. Mooser, F. Morice-Picard, **M. Mucciolo**, M. Nauck, N. Coumba Ndiaye, A. Nordgren, L. Pasquier, F. Petit, R. Pfundt, G. Plessis, E. Rajcan-Separovic, G.P. Ramelli, A. Rauch, R. Ravazzolo, A. Reis, A. Renieri, C. Richart, J.S. Ried, C. Rieubland, W. Roberts, K.M. Roetzer, C. Rooryck, M. Rossi, E. Saemundsen, V. Satre, C. Schurmann, E. Sigurdsson, D.J. Stavropoulos, H. Stefansson, C. Tengström, U. Thorsteinsdóttir, F.J. Tinahones, R. Touraine, L. Valleé, E. van Binsbergen, N. Van der Aa, C. Vincent-Delorme, S. Visvikis-Siest, P. Vollenweider, H. Vořlzke, A.T. Vulto-van Silfhout, G. Waeber, C. Wallgren-Pettersson, R.M. Witwicki, S. Zwolinski, J. Andrieux, X. Estivill, J.F. Gusella, O. Gustafsson, A. Metspalu, S.W. Scherer, K. Stefansson, A.I.F. Blakemore, J.S. Beckmann & P. Froguel.

Nature. 2011 Aug 31;478(7367):97-102.

Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus

A list of authors and their affiliations appears at the end of the paper

Both obesity and being underweight have been associated with increased mortality^{1,2}. Underweight, defined as a body mass index (BMI) ≤ 18.5 kg per m² in adults and ≤ -2 standard deviations from the mean in children, is the main sign of a series of heterogeneous clinical conditions including failure to thrive³⁻⁵, feeding and eating disorder and/or anorexia nervosa^{6,7}. In contrast to obesity, few genetic variants underlying these clinical conditions have been reported^{8,9}. We previously showed that hemizyosity of a ~600-kilobase (kb) region on the short arm of chromosome 16 causes a highly penetrant form of obesity that is often associated with hyperphagia and intellectual disabilities¹⁰. Here we show that the corresponding reciprocal duplication is associated with being underweight. We identified 138 duplication carriers (including 132 novel cases and 108 unrelated carriers) from individuals clinically referred for developmental or intellectual disabilities (DD/ID) or psychiatric disorders, or recruited from population-based cohorts. These carriers show significantly reduced postnatal weight and BMI. Half of the boys younger than five years are underweight with a probable diagnosis of failure to thrive, whereas adult duplication carriers have an 8.3-fold increased risk of being clinically underweight. We observe a trend towards increased severity in males, as well as a depletion of male carriers among non-medically ascertained cases. These features are associated with an unusually high frequency of selective and restrictive eating behaviours and a significant reduction in head circumference. Each of the observed phenotypes is the converse of one reported in carriers of deletions at this locus. The phenotypes correlate with changes in transcript levels for genes mapping within the duplication but not in flanking regions. The reciprocal impact of these 16p11.2 copy-number variants indicates that severe obesity and being underweight could have mirror aetiologies, possibly through contrasting effects on energy balance.

Copy-number variants (CNVs) at the 16p11.2 locus have been associated with cognitive disorders including autism (deletions) and schizophrenia (duplications)¹¹⁻¹³, conditions that have been suggested to lie at opposite ends of a single spectrum of psychiatric phenotypes¹⁴. We and others have reported that a deletion of this region spanning 28 genes (Supplementary Table 1) increases the risk of morbid obesity 43-fold (Supplementary Fig. 1)^{10,15}. We hypothesized that the reciprocal duplication, with its resulting increase in gene dosage, may influence BMI in a converse manner. The duplication was identified in 73 out of 31,424 patients with DD/ID, a frequency consistent with previous reports¹³ (Table 1). Four additional cases were identified among 1,080 patients affected by bipolar disease or schizophrenia. Compared to its prevalence in seven European population-based genome-wide association study (GWAS) cohorts¹⁶⁻¹⁸ (31 out of 58,635 individuals), the duplication was significantly more frequent in both the DD/ID cohorts ($P = 4.23 \times 10^{-13}$; odds ratio = 4.4, 95% confidence interval = 2.9-6.9) and the psychiatric cohorts ($P = 3.6 \times 10^{-3}$; odds ratio = 7.0, 95% confidence interval = 1.8-19.9) (Table 1), strengthening previous reports of similar associations^{12,13}. Our data do not support a two-hit model¹⁹ for the effects of 16p11.2 duplications or deletions (Supplementary Text and Supplementary Table 2).

We compared available data on weight, height and BMI for 106 independent duplication carriers (including published cases) to data for reference populations matched for gender, age and geographical location (Table 2, Methods and Supplementary Tables 3 and 4). The duplication was strongly associated with lower weight (mean Z-score -0.56 ; $P = 4.4 \times 10^{-4}$) and lower BMI (mean Z-score -0.47 ; $P = 2.0 \times 10^{-3}$) (Table 2 and Supplementary Table 5). Birth parameters ($n = 48$) were normal, indicating a postnatal effect. Adults carrying the duplication had a relative risk of being clinically underweight (BMI < 18.5) of 8.3 (95% confidence interval = 4.4-15.9, $P = 1.53 \times 10^{-10}$) (see Methods). Concordantly, none of the 3,544 patients in our obesity cohorts^{10,15} carried the duplication (Table 1).

To investigate these associations further, we carried out separate analyses of carrier patients (DD/ID and psychiatric) and non-medically ascertained carriers (population-based cohorts, plus 11 transmitting parents and three other affected first-degree relatives for whom data were available) (Table 2). Each category had significantly lower weight and BMI, with similar effect sizes. However, the proportion of underweight cases (BMI ≤ -2 s.d.) was higher in the first group than in the second group (17 out of 76 compared to 2 out of 40; $P = 0.017$). Note that the impact of the duplication on underweight status might be underestimated here owing to prescription of antipsychotic treatments that are often associated with weight gain²⁰ (Supplementary Table 6).

Having demonstrated an association of the duplication with being underweight, we investigated the implications of gender for the resulting phenotypes (Fig. 1, Supplementary Fig. 2 and Supplementary Table 7). In DD/ID patients, the impact of the duplication on being underweight is stronger in males; the effect in females is in the same direction, but is smaller and not statistically significant (Table 2). A similar and significant difference ($P = 0.0168$) was observed in adult carriers (all groups combined): the relative risk of being underweight for males is 23.2 (95% confidence interval = 9.1-59.3, $P = 4.6 \times 10^{-11}$); for females it is only 4.7 (95% confidence interval = 1.9-11.8, $P = 9.9 \times 10^{-4}$). A gender bias was also observed in the ascertainment of DD/ID duplication carriers, in which we have an excess of males (51 males:33 females, $P = 0.044$). By contrast, carriers from the general population showed a strong overrepresentation of females (10 males:21 females, $P = 0.035$) (Supplementary Text). A similar bias was observed among transmitting parents (7 males:23 females, $P = 5.53 \times 10^{-4}$). Thus, there is an overrepresentation of males in the medically ascertained group, and a depletion in the non-medically ascertained one. We suggest that males may be more likely than females to present severe phenotypes, and that this may account for the gender bias because severely affected males may be less likely to be recruited to adult population cohorts or to be reproductively successful.

As previously reported²¹, the duplication was also associated with reduced head circumference (mean Z-score -0.89 ; $P = 7.8 \times 10^{-6}$) (Fig. 1), 26.7% presenting with microcephaly (head circumference ≤ -2 s.d.), whereas carriers of the reciprocal deletion had an increased head circumference (mean Z-score $+0.57$; $P = 1.79 \times 10^{-5}$) (Supplementary Fig. 3 and Supplementary Table 8): an additional instance of a mirror phenotype associated with reciprocal changes in copy number at this locus. Notably, head circumference Z-scores correlate

Table 1 | 16p11.2 rearrangements in cases and controls

Ascertainment	Cohorts	Duplication		Deletion		Total	
		n	P†	n	P†		
Neuro-developmental disorders	Unspecified DD/ID* from 28 cytogenetic centres	72		113		30,323	
	ADHD‡, deCODE	0		1		591	
	Childhood autism‡, deCODE	0		2		159	
	Childhood autism spectrum disorder‡, deCODE	1		3		351	
	TOTAL	73	4.23 × 10⁻¹³	119	5.43 × 10⁻³²	31,424	
	Rearrangement frequency (95% CI)	0.23% (0.18–0.29)		0.38% (0.31–0.45)			
Family history	First-degree relatives of probands	30		35		43/62‡	
	Schizophrenia, deCODE	0		1		657	
	Bipolar disease, Rouen	1		0		156	
	Schizophrenia, schizo-affective, Rouen	3		0		267	
	TOTAL	4	3.57 × 10⁻³	1	3.78 × 10⁻¹	1,080	
	Rearrangement frequency (95% CI)	0.37% (0.01–0.73)		0.09% (0–0.27)			
Underweight Obesity	Eating disorder, Spain	13		0		441	
	Obesity, Spain	0		2		653	
	Adult obesity, France	0		4		705	
	Childhood obesity, France & UK	0		7		1,574	
	Obesity bariatric surgery, France	0		2		141	
	Obesity discordant siblings, Sweden	0		2		159	
	Obesity and cognitive delay, France & UK	0		9		312	
	TOTAL	0	4.21 × 10⁻¹	26	2.52 × 10⁻¹⁹	3,544	
		Rearrangement frequency (95% CI)		0.73% (0.45–1.01)			
	Population-based cohorts	NFBC1966 Finnish	4		3		5,319
		CoLaus Swiss	5		0		5,612
EGCUT Estonian		2		1		2,994	
deCODE Iceland		17		18		36,601	
SHIP Germany		1		2		4,070	
KORA F3+F4 Germany		2		1		3,458	
Paediatric family study		0		0		581	
TOTAL		31		25		58,635	
		Rearrangement frequency (95% CI)	0.05% (0.03–0.07)		0.04% (0.03–0.06)		

CI, Confidence interval; ADHD, attention-deficit hyperactivity disorder. *Not a disease-specific cohort. Detailed distribution is provided in the online methods. †Fisher's exact test, compared to the combined frequency in general population groups. ‡There was no overlap between these 3 cohorts. §Atypical duplication (see Supplementary Fig. 5). ¶Total number of parental pairs tested for duplication/deletion. 13 out of 43 duplications and 27 out of 62 deletion cases were de novo.

positively with those of BMI in carriers of both the duplication ($\rho = 0.37$; $P = 2.65 \times 10^{-3}$) and the deletion ($\rho = 0.42$; $P = 1.9 \times 10^{-3}$) (Supplementary Methods). This indicates that head circumference and BMI may be regulated by a common pathway, or that a causal relationship exists between these two traits in these patients. Alternatively, the two phenotypes may arise from distinct genes and pathways. A full list of malformations and secondary phenotypes reported in duplication carriers ascertained for DD/ID is available in Supplementary Table 9.

In view of the importance of modified eating behaviours in obesity and being underweight, the clinical reports of duplication carriers were screened for evidence of such modified behaviours. In 11 out of 77 clinically ascertained cases, clinicians had spontaneously reported low food intake and selective and restrictive eating behaviour, again mirroring one of the phenotypes—hyperphagia—seen in deletion carriers¹⁰ (Supplementary Table 6) and indicating that the duplication may increase the risk of eating disorders. Consequently, we carried out

multiplex ligation-dependent probe amplification (MLPA, Supplementary Table 10) to screen for 16p11.2 rearrangements in 441 patients diagnosed with eating disorders, including anorexia nervosa, bulimia and binge eating disorder (Table 1 and Supplementary Text). No duplications of the entire region were identified, but one out of 109 anorexia nervosa patients carried an atypical 136-kb duplication that encompasses the sialophorin (SPN) and quinolinate phosphoribosyltransferase (QPRT) genes (Supplementary Fig. 4). This single, smaller duplication does not allow us to draw any firm conclusions, but together with other atypical rearrangements, it may, in the future, be essential for establishing the roles of the 28 genes within the region.

Large genomic structural variants are known to affect the expression of genes not only within the affected region but also at a distance^{22–25}. Therefore, it is possible that the phenotypes observed in 16p11.2 deletion and duplication individuals are due to effects on the expression of genes mapping outside the rearranged interval, rather than to gene dosage within the 600-kb deletion or duplication. We measured

Table 2 | Comparisons of the height, weight and BMI distributions in duplication carriers and controls.

Strata	Combined†			DD/ID or psychiatric‡			Non-medically ascertained‡			
	Mean Z-score	P-value	n*	Mean Z-score	P-value	n*	Mean Z-score	P-value	n*	
BMI	All	-0.47	2.0 × 10⁻³	102	-0.56	4.1 × 10⁻³	76	-0.45	6.0 × 10⁻³	40
	Male	-0.54	2.1 × 10⁻²	52	-0.71	1.3 × 10⁻²	43	-0.31	2.0 × 10⁻¹	14
Weight	Female	-0.4	1.8 × 10⁻²	50	-0.37	8.3 × 10⁻²	33	-0.52	4.2 × 10⁻³	26
	All	-0.56	4.4 × 10⁻⁴	104	-0.65	1.3 × 10⁻³	78	-0.61	3.0 × 10⁻³	40
Height	Male	-0.64	5.8 × 10⁻³	53	-0.79	4.4 × 10⁻³	44	-0.57	8.8 × 10⁻²	14
	Female	-0.47	1.7 × 10⁻²	51	-0.47	6.5 × 10⁻²	34	-0.63	8.6 × 10⁻³	26
Height	All	-0.24	4.8 × 10⁻²	103	-0.33	3.6 × 10⁻²	77	-0.15	1.8 × 10⁻¹	40
	Male	-0.34	4.5 × 10⁻²	52	-0.4	4.6 × 10⁻²	43	-0.29	1.2 × 10⁻¹	14
Female	-0.14	2.6 × 10⁻¹	51	-0.24	2.1 × 10⁻¹	34	-0.07	3.7 × 10⁻¹	26	

The available BMI, weight and height data for duplication carriers were transformed to Z-scores using gender- and age-matched reference populations, and one-tailed t-tests were carried out to determine whether the mean Z-scores deviated from zero. Significant differences were identified by reference to cutoffs controlling the false discovery rate at 5% (see Methods). BMI, 0.022; weight, 0.032; height, 0.025. Significant results are indicated in bold. Data were not available for all subjects. *Relatives of probands were excluded as required, to avoid including more than one member of the same family in a single analysis. †Including 24 cases from the literature (Supplementary Table 3). ‡Population-based cases and first-degree relatives of probands.

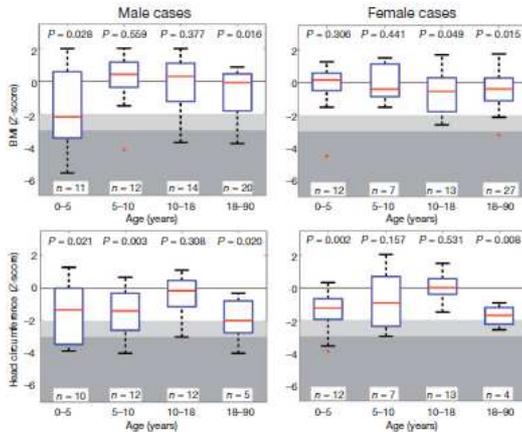


Figure 1 | Effect of the chromosome 16p11.2 duplication on BMI and head circumference. Z-score values of BMI and head circumference in carriers of the 16p11.2 duplication, stratified by gender and age group. The most severe effect is observed in children at 0–5 years of age. Boxplots represent the fifth, twenty-fifth, median, seventy-fifth and ninety-fifth percentile for each age group. Light grey and dark grey backgrounds represent ≤ -2 and ≤ -3 s.d., respectively, corresponding to the WHO definition of moderately and severely underweight⁴⁰. BMI is decreased in adolescent and adult females.

relative transcript levels of 27 genes mapping within or near to the rearrangement, using lymphoblastoid cell lines (Supplementary Tables 1 and 11): six from deletion carriers, five from duplication carriers and ten from gender- and age-matched controls (Supplementary Table 12). Expression levels correlated positively with gene dosage for all genes in the CNV region (Fig. 2), consistent with published partial results from adipose tissue¹⁰. Mean relative transcript levels in deletion and duplication carriers were, respectively, 67% and 214% of the levels measured in controls (Supplementary Table 13). Although genes proximal (centromeric) to the rearrangement interval

showed no significant variation in relative transcript levels between patients and controls (Fig. 2), distal (telomeric) genes showed a marked alteration in relative expression. However, their expression levels, including that of *SH2B1* (for which gene dosage and a nearby single nucleotide polymorphism (SNP) have been associated with obesity^{15,26}), were similarly upregulated in cell lines of both deletion and duplication carriers, showing no apparent correlation between transcript level and either copy number or phenotype (Fig. 2). Although lymphoblastoid cells may not recapitulate obesity-relevant tissues, previous experiments have shown a high degree of correlation between expression levels in different tissues and cell lines²², indicating that the same pathways may be similarly disrupted in different cell lineages. Thus, any involvement of these distal genes in the control of BMI in these subjects seems unlikely.

Our study demonstrates the power of very large screens (>95,000 samples: to our knowledge the largest of its kind so far) to characterize the clinical and molecular correlates of a rare functional genomic variant. We demonstrate unambiguously that carrying the 16p11.2 duplication confers a high risk of being clinically underweight, and show that reciprocal changes in gene dosage at this locus result in several mirror phenotypes. As in the schizophrenia/autism¹⁴ and microcephaly/macrocephaly²¹ dualisms, abnormal eating behaviours, such as hyperphagia and anorexia, could represent opposite pathological manifestations of a common energy-balance mechanism, although the precise relationships between these mirror phenotypes remain to be determined. We speculate that head circumference (which correlates with brain volume²⁷), and thus neuronal circuitry, may affect cognitive function and energy balance in patients with 16p11.2 rearrangements (possibly through eating behaviour). Consistent with this are previous reports that a subgroup of children with microcephaly show a concomitant reduction in weight percentile²⁸. Our findings also support the observation that severe overweight and underweight phenotypes correlate with lower cognitive functioning^{4,29}. Thus, abnormal food intake may be a direct result of particular neurodevelopmental disorders. Although it is possible that the 16p11.2 region encodes distinct genes specific for each trait, a more parsimonious hypothesis is that these clinical manifestations of dysfunction of the central nervous system are all secondary to the disruption of a single neurodevelopmental step that is sensitive to gene dosage. Further resolution of this issue may

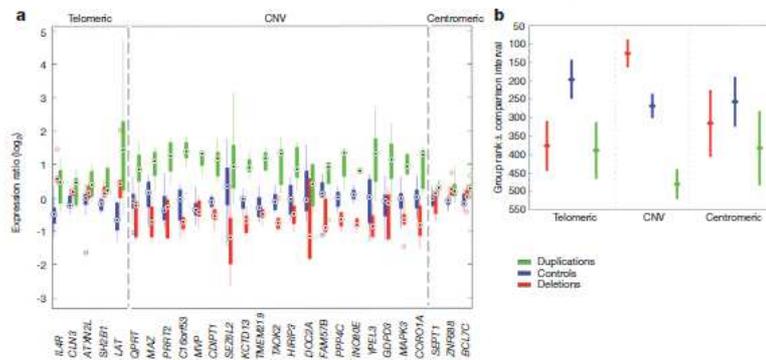


Figure 2 | Transcript levels for genes within and near to the 16p11.2 rearrangements. a, Relative expression levels of 27 genes mapping to 16p11.2 in deletion ($n = 6$) and duplication ($n = 5$) carriers (red and green, respectively), and in control cell lines ($n = 10$, blue). Grey lines denote the extent of the 16p11.2 CNV (29.5–30.1 megabases (Mb)). Complete lists of genes mapping within the rearranged interval, and of the quantitative PCR assays, are in Supplementary Tables 1 and 11, respectively. For the possible relevance of each of these genes to obesity/leanness and/or developmental delay/cognitive deficits, see ref. 10. b, Rank comparison (Kruskal–Wallis test)

between the expression of 27 genes mapping to 16p11.2 in deletion and duplication carriers (red and green, respectively) and in control cell lines (blue). Genes are labelled as telomeric, centromeric or within the rearranged interval (CNV). Dots correspond to the mean group rank and bars indicate the comparison interval. Groups with non-overlapping intervals are significantly different (P -values were adjusted for multiple testing issues using a Bonferroni correction, where the number of tests is the number of pairwise comparisons; the resulting adjusted P -value was less than 0.05).

require the identification of additional patients with rare atypical rearrangements in this region.

METHODS SUMMARY

Underweight is defined in adults as BMI ≤ 18.5 . In individuals younger than 18 years of age, it is defined as a Z-score ≤ -2 .

Statistics. Two-tailed Fisher's exact test was used to compare frequencies of the rearrangement in patients and controls. Z-scores were computed for all data using gender-, age- and geographically-matched reference populations. One-tailed Student's *t*-test was performed to test BMI, height, weight and head circumference in duplication carriers for Z-scores of less than zero. We used Kruskal–Wallis tests for differences in gene expression patterns. *P*-values were adjusted using a Bonferroni correction, considering the number of pairwise comparisons; the resulting adjusted *P*-value was less than 0.05. The relative risk of being underweight was calculated as the ratio of the fraction of underweight individuals among duplication carriers versus our control group.

Discovery of CNVs. Carriers of 16p11.2 duplication and deletion were identified through various procedures: (1) comparative genomic hybridization with Agilent 44K, 60K, 105K, 180K, 244K arrays; (2) Illumina Human370, Human370, HumanHap550, Human610 and IM BeadChips; (3) Affymetrix 6.0, 500K genotyping arrays; (4) quantitative multiplex PCR of short fluorescent fragments (QMPSF); (5) fluorescent *in situ* hybridization (FISH); (6) MLPA. CNV analyses of GWAS data were carried out using cnvHap, a moving-window average-intensity procedure, a Gaussian mixture model, circular binary segmentation, QuantiSNP, PennCNV, BeadStudio GT module and Birdseed. At least two independent algorithms were used for each cohort.

Expression analyses. Lymphoblastoid cell lines were established from carriers and controls. SYBR Green quantitative PCR was performed to assess relative expression of genes.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 February; accepted 29 July 2011.

Published online 31 August 2011.

- Berrington de Gonzalez, A. *et al.* Body-mass index and mortality among 1.46 million white adults. *N. Engl. J. Med.* **363**, 2211–2219 (2010).
- Flegal, K. M., Graubard, B. I., Williamson, D. F. & Gail, M. H. Cause-specific excess deaths associated with underweight, overweight, and obesity. *J. Am. Med. Assoc.* **298**, 2028–2037 (2007).
- Olsen, E. M. *et al.* Failure to thrive: the prevalence and concurrence of anthropometric criteria in a general infant population. *Arch. Dis. Child.* **92**, 109–114 (2007).
- Corbett, S. S. & Dnewell, R. F. To what extent is failure to thrive in infancy associated with poorer cognitive development? A review and meta-analysis. *J. Child Psychol. Psychiatr.* **45**, 641–654 (2004).
- Rudolf, M. C. & Logan, S. What is the long term outcome for children who fail to thrive? A systematic review. *Arch. Dis. Child.* **90**, 925–931 (2005).
- Bravender, T. *et al.* Classification of eating disturbance in children and adolescents: proposed changes for the DSM-V. *Eur. Eat. Disord. Rev.* **18**, 79–89 (2010).
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR* 39–134 and 583–596 (Am. Psychiatric Assoc., 2000).
- Scherag, S., Hebebrand, J. & Hinney, A. Eating disorders: the current status of molecular genetic research. *Eur. Child Adolesc. Psychiatr.* **19**, 211–226 (2010).
- Wang, K. *et al.* A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol. Psychiatry* doi:10.1038/mp.2010.107 (2010).
- Walters, R. G. *et al.* A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* **463**, 671–675 (2010).
- Marshall, C. R. *et al.* Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 477–488 (2008).
- McCarthy, S. E. *et al.* Microduplications of 16p11.2 are associated with schizophrenia. *Nature Genet.* **41**, 1223–1227 (2009).
- Weiss, L. A. *et al.* Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.* **358**, 667–675 (2008).
- Crespi, B., Stead, P. & Elliot, M. Evolution in health and medicine Sackler colloquium: comparative genomics of autism and schizophrenia. *Proc. Natl Acad. Sci. USA* **107** (Suppl 1), 1736–1741 (2010).
- Bochukova, E. G. *et al.* Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* **463**, 666–670 (2010).
- Firmann, M. *et al.* The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc. Disord.* **8**, 6 (2008).
- Sabatti, C. *et al.* Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nature Genet.* **41**, 35–46 (2009).
- Neils, M. *et al.* Genetic structure of Europeans: a view from the north-east. *PLoS ONE* **4**, e5472 (2009).
- Girirajan, S. & Eichler, E. E. Phenotypic variability and genetic susceptibility to genomic disorders. *Hum. Mol. Genet.* **19**, R176–R187 (2010).
- Pramyothin, P. & Khaodh, L. Metabolic syndrome with the atypical antipsychotics. *Curr. Opin. Endocrinol. Diabetes Obes.* **17**, 460–466 (2010).
- Shinawi, M. *et al.* Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioral problems, dysmorphism, epilepsy, and abnormal head size. *J. Med. Genet.* **5**, 332–341 (2009).
- Merla, G. *et al.* Submicroscopic deletion in patients with Williams–Beuren syndrome influences expression levels of the nonhemizygous flanking genes. *Am. J. Hum. Genet.* **79**, 332–341 (2006).
- Stranger, B. E. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
- Hentchens, C. N. *et al.* Segmental copy number variation shapes tissue transcriptomes. *Nature Genet.* **41**, 424–429 (2009).
- Ricard, G. *et al.* Phenotypic consequences of copy number variation: insights from Smith–Magenis and Potocki–Lupski syndrome mouse models. *PLoS Biol.* **8**, e1000543 (2010).
- Wilder, C. J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature Genet.* **41**, 25–34 (2009).
- Courchesne, E., Carper, R. & Akshoomoff, N. Evidence of brain overgrowth in the first year of life in autism. *J. Am. Med. Assoc.* **290**, 337–344 (2003).
- Baxter, P. S., Rigby, A. S., Rotsaert, M. H. & Wright, I. Acquired microcephaly: causes, patterns, motor and IQ effects, and associated growth changes. *Pediatrics* **124**, 590–595 (2009).
- Li, Y., Dai, Q., Jackson, J. C. & Zhang, J. Overweight is associated with decreased cognitive functioning among school-age children and adolescents. *Obesity (Silver Spring)* **16**, 1809–1815 (2008).
- de Onis, M., Blossner, M., Borghi, E., Frongillo, E. A. & Morris, R. Estimates of global prevalence of childhood underweight in 1990 and 2015. *J. Am. Med. Assoc.* **291**, 2600–2606 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Vital-IT high-performance computing centre of the Swiss Institute of Bioinformatics. S.J. is recipient of a bourse de relèvement académique de la Faculté de Biologie et Médecine de l'Université de Lausanne. This work was supported by the Leenaards Foundation Prize (S.J., D.M. and A.Reymond), the Jérôme Lejeune Foundation (A.Reymond), the Telethon Action Suisse Foundation (A.Reymond), the Swiss National Science Foundation (A.Reymond, J.S.B., S.B. and S.E.A.), a Swiss National Science Foundation Sinergia grant (S.J., D.M., S.B., J.S.B. and A.Reymond), the European Commission an EU-Playoity Integrated Project grant 037627 (A.Reymond, S.B., X.E., H.G.B. and S.E.A.), the Ludwig Institute for Cancer Research (AV.), the Swiss Institute of Bioinformatics (S.B. and Z.K.), an Imperial College Department of Medicine PhD studentship (J.S.E.-S.M.), the Comprehensive Biomedical Research Centre, Imperial College Healthcare NHS Trust, and the National Institute for Health Research (P.E.), the Wellcome Trust and the Medical Research Council (A.I.F.B. and P.F.), the Instituto de Salud Carlos III (ISCIII)-FIS, the German Mental Retardation Network funded through a grant of the German Federal Ministry of Education and Research (NGFNplus 01GS08160) (A.Reis), European Union-FEDER (PI081714, PS09/01778) (F.F.A., M.G. and X.E.), SAF2008-02278 (C.R.), the Belgian National Fund for Scientific Research, Flanders (N.V.A. and R.F.K.), the Dutch Organisation for Health Research and Development (ZON-MW grant 917-86-319) and Hersenstichting Nederland (B.B.A.d.V.), grant 81000346 from the Chinese National Natural Science Foundation (Y.G.Y.), the Simons Foundation Autism Research Initiative, Autism Speaks and NIH grant GM061354 (J.F.G.), and Oesterreichische Nationalbank (ONB) grant 13059 (A.K.-B.). Y.S. holds a Young Investigator Award from the Children's Tumor Foundation and a Catalyst Award from Harvard Medical School. B.L.W. holds a Fudan Scholar Research Award from Fudan University, a grant from Chinese National '973' project on Population and Health (2010CB529601) and a grant from the Science and Technology Council of Shanghai (09JC1402400). E.R.S. and S.L., recipients of the Michael Smith Foundation for Health Research Scholar award, acknowledge the CHR MOP 74502 operational grant. The Estonian Genome Center of the University of Tartu (EGCUT) received support from the EU Centre of Excellence in Genomics and FP7 grants 201413 and 245536, and from Estonian Government SF0180142s08, SF0180026s09 and SF0180027s10 (A.M., K.M. and A.K.). D.S. thanks the Direction Générale de l'Organisation des Soins from the French Ministry of Health for their support in the development of several array-CGH platforms, and the Centres Labelisés Anomalies du Développement in France. The Helmholtz Zentrum Munich and the State of Bavaria financed the KORA study, also supported by the German National Genome Research Network (NGFN-2 and NGFNPlus: 01GS0823), the German Federal Ministry of Education and Research (BMBF), and the Munich Center of Health Sciences (MC Health, LMUinnovativ). CIBEROBN and CIBERESP are initiatives of ISCIII (Spain). S.W.S. holds the GlaxoSmithKline-Canadian Institutes of Health Chair in Genetics and Genomics at the University of Toronto and the Hospital for Sick Children, and is supported by Genome Canada and the McLaughlin Centre. Funding for deCODE came in part from NIH grant MH071425 (K.S.), EU grant HEALTH-2007-2-2-10-223423 (Project PsychCNV) and EU grant IMI-JU-NewMeds. NFB01966 received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, BioCenter, University of Oulu, Finland (75617), the European Commission (EURO-BLCS, Framework 5 award QL6-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706-02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, and the Medical Research Council, UK (G0500539, G0600705, PrevMetSyn/SALVE). The DNA extractions, sample quality controls, biobank up-keeping and aliquotting was performed in the National Public Health Institute, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biozentrum Helsinki. We thank M. Hass, Z. Jaros, M. Jussila, M. Koironen,

P. Rantakallio, M. C. Rudolf, V. Soo, O. Tornwall, S. Vaara, T. Ylitalo and the French DHOS national CGH network for their help, as well as all participating patients and clinicians. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions S.J. A.Reymond, P.F. and J.S.B. wrote the manuscript with contributions from F.Z., L.H., R.G.W., N.D.B., Z.K., A.I.F.B. and A.V. L.H., A.V. and A.Reymond produced and analyzed the expression data. Z.K., A.V., R.G.W. and N.D.B. conducted the statistical analyses, guided by S.J., A.Reymond, P.F. and J.S.B. S.J., A.Reymond, F.Z., L.H., D.M., Y.S., G.T., M.B., S.B., D.C., N.d.L., B.B.A.d.V., B.A.F., F.F.A., M.G., A.G., J.H., A.K., C.L.C., K.M., O.S.P., D.S., M.M.V.H., S.V.G., A.T.V.-v.S., F.W., B.-L.W., Y.Y., J.A., X.E., J.F.G., A.M., S.W.S., K.S., U.T., A.I.F.B., J.S.B., P.F. and all other authors phenotyped and/or genotyped patients and/or individuals of the general population. S.J., A.Reymond and J.S.B. designed the study. All authors commented on and approved the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. J.S.B. and P.F. are equally contributing senior authors and F.Z., L.H., R.G.W. and Z.K. are equally contributing second authors. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.S.B. (Jacques.Beckmann@chuv.ch).

Sébastien Jacquemont^{1*}, Alexandre Reymond^{2*}, Flore Zufferey¹, Louise Harewood², Robin G. Walters³, Zoltán Kutalik⁴, Danielle Martinet¹, Yiping Shen^{5,7}, Armand Valsesia^{4,5,8}, Noam D. Beckmann⁹, Gudmar Thorleifsson¹⁰, Marco Belfiore⁶, Sonia Bouquillon¹⁰, Dominique Campion¹¹, Nicole de Leeuw¹², Bert B. A. de Vries¹², Tõnu Esko^{13,14}, Bridget A. Fernandez¹⁵, Fernando Fernández-Aranda¹⁶, José Manuel Fernández-Real¹⁷, Mónica Gratacós¹⁸, Audrey Guilmatre¹¹, Juliane Hoyer¹⁹, Marjo-Riitta Jarvelin^{20,21,22}, R. Frank Kooy²³, Ants Kurg¹³, Cédric Le Cañgne²⁴, Katrin Männik¹³, Orah S. Platt⁶, Damien Sanlaville²⁵, Mieke M. Van Haest^{26,26}, Sergi Villatoro Gomez¹⁸, Daina Walha²⁷, Bai-lin Wu²⁷, Yongguo Yu^{6,28}, Azzedine Aboura²⁹, Marie-Claude Addor³⁰, Yves Alembik³⁰, Stylianos E. Antonarakis³¹, Benoit Arveiler^{30,32}, Magalie Barth³⁴, Nathalie Bednarek³⁵, Frédérique Béna³¹, Sven Bergmann⁴⁵, Mylène Berri³⁶, Laura Bernardini³⁷, Bettina Blaumeiser³⁸, Dominique Bonneu³⁴, Armand Bottani³¹, Odile Boute³⁸, Han G. Brunner¹², Dorothée Calley³³, Patrick Callier³⁹, Jean Chiesa⁴⁰, Jacqueline Chrast⁴¹, Lachlan Coin⁴¹, Charles Coutton⁴¹, Jean-Marie Cuisset⁴², Jean-Christophe Cuvelier⁴², Albert David⁴⁴, Bénédicte de Freminville⁴³, Bruno Delobel⁴⁴, Marie-Ange Delrue^{32,33}, Bénédicte Demeer⁴⁵, Dominique Descamps⁴⁶, Gérard Didelot⁴⁷, Klaus Dieterich⁴⁷, Vittoria Disciglio^{48,110}, Martine Doco-Fenzy⁴⁹, Séverine Drunat⁴⁹, Bénédicte Duban-Bedu⁴⁴, Christèle Dubourg^{50,51}, Julia S. El-Sayed Moustafa⁵², Paul Elliott^{32,53}, Brigitte H. W. Faas¹², Laurence Faivre⁵⁴, Anne Faudet⁵⁵, Florence Fellmann¹, Alessandra Ferrarini¹, Richard Fisher⁵⁶, Elisabeth Fiori⁵⁷, Lukas Forer⁵⁷, Dominique Gaillard⁴⁹, Marlon Gerard⁵⁸, Christian Gieger⁵⁸, Stefania Gimelli⁵¹, Giorgio Gimelli⁵¹, Hans J. Graber⁶⁰, Agnès Guichet⁶⁴, Olivier Guillin⁶¹, Anna-Liisa Hartikainen⁶¹, Delphine Heron^{62,63}, Loyse Hippolyte¹, Muriel Holder²⁸, Georg Homuth⁶⁴, Bertrand Isidor²⁴, Sylvie Jaillard^{65,65}, Zdenek Jaroš⁶⁶, Susana Jimenez-Murcia⁶⁶, Géraldine Joly Helas⁶⁷, Philippe Jonveaux⁶⁸, Satu Kaksonen⁶⁸, Boris Keren⁶⁹, Anita Kloss-Brandstätter⁶⁷, Niine V. A. M. Knoppers^{26,69}, David A. Koolen¹², Peter M. Krose⁷⁰, Florian Kronenberg⁶⁷, Audrey Labalme⁷¹, Emilie Landais⁴⁵, Elisabetta Lapi⁷², Valérie Layet⁷², Solenn Legalic¹¹, Bruno Leheup⁷³, Barbara Leube⁷⁴, Suzanne Lewis⁷⁵, Josefine Lucas⁶⁵, Kay D. MacDermot⁶⁶, Pall Magnusson⁷⁷, Christian Marshall⁷⁸, Michèle Mathieu-Dramard⁴⁵, Mark I. McCarthy⁷⁹, Thomas Meitinger^{82,111}, Maria Antonietta Mencarelli⁴⁸, Giuseppe Merlo⁸³, Alexandre Moerman²⁸, Vincent Mosser⁸⁴, Fanny Morice-Picard^{82,83}, Mafalda Mucciolo⁴⁸, Matthias Nauck⁸⁵, Ndeye Coumba Ndiaye⁸⁶, Ann Nordgren⁸⁷, Laurent Pasquier⁸⁸, Florence Petit⁸⁸, Rolf Pfundt¹², Ghislaine Plesse⁸⁹, Evica Rajcan-Separovic⁹⁰, Gian Paolo Ramelli⁹¹, Anita Rauch⁹², Roberto Ravazzolo⁹³, Andre Reis¹³, Alessandra Renieri⁴⁸, Cristóbal Richar⁹⁴, Janina S. Ried⁹⁵, Claudine Rieubland⁹⁵, Wendy Roberts⁹⁶, Katharina M. Roetzer⁹⁷, Caroline Rooyck^{32,33}, Massimiliano Rossi⁹⁸, Evald Saemundsen⁹⁹, Véronique Sastre⁴¹, Claudia Schumann¹⁰⁰, Engelbert Sigurdsson⁹⁸, Dimitri J. Stavropoulos⁹⁹, Hreinn Stefansson¹⁰¹, Carola Tengström¹⁰², Unnur Thorsteinsdóttir^{99,101}, Francisco J. Tinahones¹⁰², Renaud Touraine⁴³, Louis Valée⁴², Elen van Binsbergen⁹⁶, Nathalie Van der Aar²³, Catherine Vincent-Delorme¹⁰³, Sophie Visvikis-Sies⁹⁶, Peter Vollenweider¹⁰⁴, Henry Völzke¹⁰⁵, Anneke T. Vulto-van Silfhout¹², Gérard Waeber¹⁰⁴, Carina Wallgren-Pettersson¹⁰⁶, Robert M. Witwicki¹⁰⁷, Simon Zvolinski¹⁰⁸, Joris Andrieux¹⁰, Xavier Estvil¹⁰, James F. Gusella¹⁰⁹, Omar Gustafsson^{110,11}, Andres Metspalu^{111,11}, Stephen W. Scherer¹⁰⁶, Karl Stefansson¹⁰⁹, Alexandra I. F. Blakemore¹, Jacques S. Beckmann^{1,4} & Philippe Froguel^{31,109}

¹Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois, 1011 Lausanne, Switzerland. ²Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland. ³Department of Genomics of Common Disease, Imperial College London, London W12 0NN, UK. ⁴Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland. ⁵Swiss Institute of Bioinformatics, University of Lausanne, 1015 Lausanne, Switzerland. ⁶Laboratory Medicine, Children's Hospital Boston, Boston, Massachusetts 02115, USA. ⁷Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. ⁸Ludwig Institute for Cancer Research, University of Lausanne, 1015 Lausanne, Switzerland. ⁹aeCODE Genetics, Sturlugata 8, IS-101, Reykjavik, Iceland. ¹⁰Laboratoire de Génétique Médicale, Hôpital Jeanne de Flandre, CHRU de Lille, 59000 Lille, France. ¹¹INSERM U614, University of Rouen and Centre Hospitalier du Rouvray, 76000 Rouen, France. ¹²Department of Human Genetics, Nijmegen Centre for Molecular Life Sciences and Institute for Genetic and Metabolic Disorders, Radboud University Nijmegen Medical Centre, 6500 HB

Nijmegen, The Netherlands. ¹³Institute of Molecular and Cell Biology, University of Tartu, 51010 Tartu, Estonia. ¹⁴Estonian Genome Center, University of Tartu, 51010 Tartu, Estonia. ¹⁵Discipline of Genetics and Medicine, Memorial University of Newfoundland, St John's A1B 3X6, Newfoundland, Canada. ¹⁶Department of Psychiatry, University Hospital of Bellvitge-IDIBELL, Ciber Fisiopatología Obesidad y Nutrición (CIBEROBN), 08907 Barcelona, Spain. ¹⁷Department of Diabetes, Endocrinology, and Nutrition, Hospital Universitari de Girona Dr. Josep Trueta, Institut d'Investigació Biomèdica de Girona, Ciber Fisiopatología Obesidad y Nutrición (CIBEROBN), Instituto Salud Carlos III, 17007 Girona, Spain. ¹⁸Genes and Disease Program, Center for Genomic Regulation (CRG-UPF), CIBER en Epidemiología y Salud Pública (CIBERESP), 08003 Barcelona, Catalonia, Spain. ¹⁹Institute of Human Genetics, Friedrich-Alexander University Erlangen-Nuremberg 91054 Erlangen, Germany. ²⁰Department of Epidemiology and Biostatistics, Imperial College London, School of Public Health, London W2 1PG, UK. ²¹Department of Child and Adolescent Health, National Institute for Health and Welfare, Box 310, 90101 Oulu, Finland. ²²Institute of Health Sciences, University of Oulu, and Biocenter Oulu, University of Oulu, Box 5000, 90014 University of Oulu, Finland. ²³Department of Medical Genetics, University and University Hospital Antwerp, B-2650 Antwerp, Belgium. ²⁴Service de Génétique Médicale, CHU Nantes, 44093 Nantes, France. ²⁵Service de Cytogénétique Constitutionnelle, Hospices Civils de Lyon, CHU de Lyon and Neuroscience Research Center, TIGER team, UCBL 1, Lyon, F-69600, France. ²⁶Department of Medical Genetics, University Medical Center, 3584 EA Utrecht, The Netherlands. ²⁷Children's Hospital and Institutes of Biomedical Science, Fudan University, 200032 Shanghai, China. ²⁸Shanghai Children's Medical Center, 200127 Shanghai, China. ²⁹Department of Genetics, APHP-Robert DEBRE University Hospital, 75019 Paris, France. ³⁰Service de Génétique Médicale, CHU Strasbourg, Hôpital de Hautepierre, 67000 Strasbourg, France. ³¹Service de Génétique Médicale, University Hospitals of Geneva, 1205 Geneva, Switzerland. ³²Laboratoire Maladies Rares-Génétique et Métabolisme, Université Bordeaux 2, 33076 Bordeaux, France. ³³Service de Génétique Médicale, CHU de Bordeaux, 33076 Bordeaux, France. ³⁴Service de Génétique, CHU Angers, 49933 Angers, France. ³⁵Service Pédiatrie, CHU Hôpital AIX de Champagne, 51100 Reims, France. ³⁶Laboratoire de génétique, CHU Nancy, 54511 Vandœuvre les Nancy, France. ³⁷Mendel Laboratory, IRCCS Casa Sollievo della Sofferenza Hospital, 71013 San Giovanni Rotondo, Italy. ³⁸Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU de Lille, France. ³⁹Laboratoire de Cytogénétique, CHU le Bocage, 21070 Dijon, France. ⁴⁰Laboratoire de Cytogénétique, CHU Caremeau, 30029 Nîmes, France. ⁴¹Laboratoire de Génétique Chromosomique, CHU de Grenoble, BP 217, 38043 Grenoble, France. ⁴²Service de Neurologie pédiatrique, Hôpital Roger Salengro, CHRU de Lille, 59037 Lille, France. ⁴³CHU-Hôpital Nord, Service de Génétique, CHU Saint Etienne, 42055 St Etienne, France. ⁴⁴Centre de Génétique Chromosomique, Hôpital Saint-Vincent de Paul, GHICL, 59160 Lille, France. ⁴⁵Service de Génétique Médicale, CHRU Amiens, 80000 Amiens, France. ⁴⁶Centre hospitalier de Béthune, 62408 Béthune, France. ⁴⁷Service de Génétique Clinique, CHU Grenoble, 38043 Grenoble, France. ⁴⁸Medical Genetics, Department of Biotechnology, University of Siena, 53100 Siena, Italy. ⁴⁹Service de Génétique, HMB, CHU REIMS, IFR 53, 51092 Reims, France. ⁵⁰UMR 6061 CNRS, IFR 140 GFAS, Université de Rennes 1, 35043 Rennes, France. ⁵¹Service de Génétique Moléculaire, CHU, 35033 Rennes, France. ⁵²Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK. ⁵³MRC-HPA Centre for Environment and Health, Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK. ⁵⁴Centre de Génétique, CHU Dijon, 21079 Dijon, France. ⁵⁵Département de Génétique et Cytogénétique, APHP-GH Pitié-Salpêtrière, 75013 Paris, France. ⁵⁶Institute of Human Genetics, International Centre for Life, Newcastle Upon Tyne NE1 4EP, UK. ⁵⁷Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, 6020 Innsbruck, Austria. ⁵⁸Institute of Genetic Epidemiology, Helmholtz Center Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ⁵⁹Laboratorio di Citogenetica, G. Gaslini Institute, 16147 Genova, Italy. ⁶⁰Department of Psychiatry and Psychotherapy, Ernst-Moritz-Arndt University Greifswald, 17475 Greifswald and D-18437 Stralsund, Germany. ⁶¹Institute of Clinical Medicine, Department of Obstetrics and Gynecology, University of Oulu, 90570 Oulu, Finland. ⁶²Embryologie AP-HP, Université Pierre et Marie Curie, 75005 Paris, France. ⁶³Département de Génétique Cytogénétique, Hôpital Pitié-Salpêtrière, 75013 Paris, France. ⁶⁴Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt University Greifswald, D-17487 Greifswald, Germany. ⁶⁵Laboratoire de Cytogénétique et Biologie Cellulaire, CHU Rennes, 35033 Rennes, France. ⁶⁶Abteilung für Kinder- und Jugendheilkunde, Landeskrankenhaus Waldviertel Zwettl, 3910 Zwettl, Austria. ⁶⁷Department of Genetics, CHU Rouen, 76183 Rouen, France. ⁶⁸The Habilitation Unit of Folkhalsan, Folkhalsan, SF 00250 Helsinki, Finland. ⁶⁹Department of Human Genetics, University Medical Center, 3584 EA Utrecht, The Netherlands. ⁷⁰Institute of Human Genetics, Medical University of Graz, 8010 Graz, Austria. ⁷¹Medical Genetics Unit, Children's Hospital Anna Meyer, 50139 Firenze, Italy. ⁷²Unité de Génétique, Groupe Hospitalier du Havre, 76600 le Havre, France. ⁷³Service de Médecine Infantile III et Génétique Clinique, CHU-Nancy et PRES de l'Université de Lorraine UHP Nancy, 54511 Vandœuvre les Nancy, France. ⁷⁴Institute of Human Genetics and Anthropology, Heinrich-Heine University Hospital Duesseldorf, D-40001 Duesseldorf, Germany. ⁷⁵Department of Medical Genetics, The University of British Columbia and Child and Family Research Institute, Vancouver V6H 3N1, Canada. ⁷⁶North West Thames Regional Genetics Service, Northwick Park & St Marks Hospital, Harrow HA1 3UJ, UK. ⁷⁷Child and Adolescent Psychiatry, Landspítali University Hospital, IS-105 Reykjavik, Iceland. ⁷⁸The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada. ⁷⁹Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Old Road, Headington, Oxford OX3 7LJ, UK. ⁸⁰Wellcome Trust Centre for Human Genetics, University of Oxford, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. ⁸¹Oxford NIHR Biomedical Research Centre, Churchill Hospital, Old Road, Headington, Oxford OX3 7LJ, UK. ⁸²Institute of Human Genetics, Helmholtz Center Munich, German Research Center for Environmental Health and Institute of Human Genetics, Technical University Munich, 85764 Neuherberg, Germany. ⁸³Medical Genetics Unit, IRCCS Casa Sollievo della Sofferenza Hospital, 71013 San Giovanni

Rotondo, Italy.⁸⁴Genetics, GlaxoSmithKline R&D, 720 Swedeland Road, King of Prussia, Pennsylvania 19406, USA.⁸⁵Institute of Clinical Chemistry and Laboratory Medicine, Ernst-Moritz-Arndt University Greifswald, D-17475 Greifswald, Germany.
⁸⁶Cardiovascular Genetics Research Unit, EA4373, Université Henri Poincaré, 54000 Nancy, France.⁸⁷Department of Molecular Medicine and Surgery, Karolinska Institutet, 171 76 Stockholm, Sweden.⁸⁸Service de Génétique-CLAD Ouest, 35203 Rennes cedex 2, France.⁸⁹Service de Génétique, CHU Clémenceau, 14033 Caen, France.⁹⁰Department of Pathology, University of British Columbia and Child and Family Research Institute, Vancouver, British Columbia V5Z 4H4, Canada.⁹¹Division of Pediatrics, Ospedale San Giovanni, 6500 Bellinzona, Switzerland.⁹²Institute of Medical Genetics, University of Zurich, 8603 Schwerzenbach, Switzerland.⁹³Department of Pediatrics and CEBR, G. Gaslini Institute, University of Genova, 16126 Genova, Italy.⁹⁴Department of Internal Medicine, University Hospital Juan XXIII, Universitat Rovira y Virgili, Ciber Fisiopatología Obesidad y Nutrición (CIBEROBN), Instituto Salud Carlos III, 43005 Tarragona, Spain.⁹⁵Division of Human Genetics, Department of Paediatrics, Inselspital, University of Bern, 3010 Bern, Switzerland.⁹⁶The Autism Research Unit, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada.⁹⁷State Diagnostic and Counseling Center, 200 Kopavogur, Iceland.⁹⁸University of Iceland and Landspítali University Hospital, 101 Reykjavik, Iceland.⁹⁹Department of Pediatric Laboratory Medicine, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada.¹⁰⁰Genetic Services, Rinneke Research Foundation, Kumputie 1, SF-02980 Espoo, Finland.¹⁰¹Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland.¹⁰²Department of Endocrinology and Nutrition, Clinic Hospital of Virgen de la Victoria, Ciber Fisiopatología Obesidad y Nutrición (CIBEROBN), Instituto Salud Carlos III, 29010 Málaga, Spain.¹⁰³Centre de Maladies Rares, Anomalies du Développement Nord de France, CH Arras-CHRU Lille, 59000 Lille, France.¹⁰⁴Department of Internal Medicine, Centre Hospitalier Universitaire Vaudois, 1011 Lausanne, Switzerland.¹⁰⁵Institute for Community Medicine, Ernst-Moritz-Arndt University Greifswald, D-17475 Greifswald, Germany.¹⁰⁶Department of Medical Genetics, Haartman Institute, University of Helsinki and Folkhälsan Institute of Genetics, 00251 Helsinki, Finland.¹⁰⁷Department of Psychiatry, Oslo University Hospital, N-0407 Oslo, Norway.¹⁰⁸The Hospital for Sick Children, University of Toronto, Toronto, Ontario M5G 1L7, Canada.¹⁰⁹CNRS 8090-Institute of Biology, Pasteur Institute, 59800 Lille, France.¹¹⁰UOC Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy.¹¹¹Institute of Human Genetics, Technische Universität München, Klinikum rechts der Isar, 81675 München, Germany.
 *These authors contributed equally to this work.

METHODS

Study cohorts. For the description of these cohorts, refer to Supplementary Information.

CNV detection. Cases ascertained for intellectual disabilities and developmental delay were identified through standard medical diagnostic procedures. CNV analyses of GWAS data were variously carried out using *cnvHap*³¹; a moving-window average-intensity procedure; a Gaussian mixture model (Valesia *et al.*, submitted); circular binary segmentation^{32,33}; QuantiSNP³⁴; PennCNV³⁵; BeadStudio GT module (Illumina Inc.); and Birdseed³⁶ (see below). At least two independent algorithms were used for each cohort.

Patients referred for intellectual disabilities and developmental delay. All diagnostic procedures (CGH, quantitative PCR and/or quantitative multiplex PCR of short fluorescent fragments) were carried out according to the relevant guidelines of good clinical laboratory practice for the respective countries. All rearrangements in probands were confirmed by a second independent method and karyotyping was performed in all cases to exclude a complex rearrangement. **Northern Finland 1966 birth cohort (NFBC).** CNV calling has been previously described¹⁰. In brief, data were normalized using Illumina BeadStudio, then GC effects on ratios were removed by regressing on GC and GC2, and wave effects were removed by fitting a Loess function³⁷. CNV analysis was done using *cnvHap*³¹. All called 16p11.2 duplications were validated by direct analysis of log₂ ratios. Data for each probe were normalized by first subtracting the median value across all samples (so that the distribution of ratios for each probe was centred on zero), and then dividing by the variance across all samples (to correct for variation in the sensitivity of different probes to copy-number variation). All CNV calls were confirmed by MLPA.

deCODEgenetics. Illumina Human317, Human370, HumanHap550, Human610 and 1M BeadChips were used for CNV analysis. BeadStudio (version 2.0) was used to call genotypes, normalize the signal intensity data and establish the log₂ ratio (LRR) and B allele frequency (BAF) at every SNP according to standard Illumina protocols. All samples passed a standard SNP-based quality control procedure with a SNP call rate greater than 0.97. PennCNV³⁵, a free, open-source tool, was used for detection of CNVs. The input data for PennCNV are LRR, a normalized measure of the total signal intensity for the two alleles of the SNP, and BAF, a normalized measure of the allelic intensity ratio of the two alleles. These values are derived with the help of control genotype clusters (HapMap samples), using the Illumina BeadStudio software. PennCNV employs a hidden Markov model to analyse the LRR and BAF values across the genome. CNV calls are made on the basis of the probability of a given copy state at the current marker, as well as on the probability of observing a copy-state change from the previous marker to the current one. PennCNV uses a built-in correction model for GC content³⁸.

Cohorte Lausannoise (CoLaus). Data normalization and CNV calling have been previously described¹⁰. Data normalization included allelic cross-talk calibration^{39,40}, intensity summarization using robust median average, and correction for any PCR amplification bias. Wave effects were corrected by fitting a Loess function³⁷. CNV calling was done using a Gaussian mixture model (Valesia *et al.*, submitted) that fits four components (deletion, copy-neutral, one additional copy and two additional copies) to copy-number ratios. The final copy number at each probe location is determined as the expected (dosage) copy number. The method has been validated by comparing test data sets with results from the CNAT⁴¹ and CBS^{32,33} algorithms, and by replicating a subset of CoLaus subjects on Illumina arrays. Only duplications found by both Gaussian mixture model and CBS were considered.

Estonian genome center of the University of Tartu (EGCUT). Genotypes were called by BeadStudio software GT module v3.1 or GenomeStudio GT v1.6 (Illumina Inc.). Values for LRR and BAF produced by BeadStudio were formatted for further CNV analysis and break-point mapping with Hidden-Markov-Model-based softwares QuantiSNP (ver.1.1)³⁴ and PennCNV³⁵ or CNVPartition 2.4.4 (Illumina Inc.). All analyses were carried out using the recommended settings, except changing EMiters to 25 and L to 1,000,000 in QuantiSNP. For PennCNV, the Estonian-population-specific SNP allele frequency data was used. All detected duplications were confirmed by quantitative PCR.

Study of health in Pomerania (SHIP). Raw intensities were normalized using Affymetrix power tools (Affymetrix); CNV analysis was done using Birdseye from the Birdsuite software package³⁶ and PennCNV³⁵. PennCNV predictions with confidence scores less than 10 were removed. Birdsuite predictions were filtered as in ref. 15; CNVs were kept if their linkage disequilibrium (LOD) score was >10, length >1 kb, number of probes ≥5 and size per number of probes <10,000.

Kooperative Gesundheitsforschung in der Region Augsburg (KORA) F3 and F4. Genotyping for KORA F3 was performed using the Affymetrix 500K array set, consisting of two chips (Sty1 and Nsp1). The KORA F4 samples were genotyped with the Affymetrix human SNP array 6.0. For both studies, genomic DNA from blood samples was used for analysis. Hybridization of genomic DNA was done in accordance with the manufacturer's standard recommendations. Genotyping was

done in the Genome Analysis Centre of the Helmholtz Centre Munich. Genotypes were determined using BRLMM clustering algorithm (Affymetrix 500K array set) and Birdseed2 clustering algorithm (Affymetrix array 6.0). For quality control purposes, we applied a positive control and a negative control DNA every 48 samples (KORA F3) or 96 samples (KORA F4). On the chip level, only subjects with overall genotyping efficiencies of at least 93% were included. In addition, the called gender had to agree with the gender in the KORA study database. After exclusions, 1,644 individuals remained in KORA F3 and 1,814 in KORA F4 for further analysis.

MLPA analysis. We used MLPA to determine changes in the copy number of a region of about 2 Mb on chromosome 16p11.2. Briefly, we designed, using hg18, nine probes within the targeted region, one control probe outside the rearranged region and seven control probes targeting unique position in the genome (Supplementary Table 10). Assays were performed with MRC-Holland reagents according to the manufacturer's protocol⁴². The analysis of the amplification products was performed by capillary electrophoresis in the DNA Analyser 3730XL and using the GeneMapper software v3.7 (Applied Biosystems). The calculations were performed independently for each experiment: we first normalized the MLPA data to minimize the amount of experimental variation, summing all signal values of each control probe for each sample, and then dividing each signal value of each sample by the sum. The normalized signal values were compared to signal values from all other samples in the same experiment, dividing the normalized signal values by the average calculated from all the samples in the same experiment. The product of this calculation is termed dosage quotient (DQ). A DQ value of less than 0.65 or more than 1.25 was considered as copy-number loss or gain, respectively, as previously described^{44–46}.

Custom array-CGH for the short arm of chromosome 16. DNA samples were labelled with Cy3 and cohybridized to custom-made Nimblegen arrays with Cy5-labelled DNA from the CEPH cell line GM12042. These arrays contained 71,000 probes spread across the short arm of chromosome 16 from 22.0 Mb to 32.7 Mb (at a median space of 45 bp between 27.5 Mb and 31.0 Mb), and 1,000 control probes situated in invariable regions of the X chromosome. DNA labelling, hybridization and washing were performed according to Nimblegen protocols. Scanning was performed using an Agilent G2565BA microarray scanner. Image processing, quality control and data extraction were performed using the NimbleScan software v.2.5.

Defining underweight. Underweight was defined throughout the study as BMI ≤ 18.5 kg per m² in adults and ≤ -2 s.d. in children^{36,47,48}.

Weight, height, BMI and head circumference Z-scores as a function of age. For paediatric cases, weight, height, BMI and head circumference Z-scores were determined for paediatric cases (0–18 years of age) using clinical growth charts specific to the country of origin. Children were ascertained from nine different countries. If charts were only available in percentiles, those measures were transformed into Z-scores using gender-, age- and geographically-matched reference populations (see Statistics).

For the USA and Canada, data from the Center for Disease Control and National Center for Health Statistics (CDC/NCHS) were used to calculate Z-scores⁴⁹.

For the French paediatric population, we used French national growth charts^{50,51}. For the Swiss paediatric population, we used Swiss national growth charts⁵². For Dutch participants, Dutch national growth charts were used⁵³. For Italian, German, Finnish and Austrian cases (*n* = 6), height, weight and BMI Z-scores were estimated using WHO growth charts⁵⁴.

To check for discrepancies generated by the use of different growth charts, height, weight and BMI Z-scores were recalculated using WHO growth charts for all cases under five years of age, regardless of origin (<http://www.who.int/childgrowth/standards/en/54>). Z-scores obtained using the WHO data were not significantly different. These growth standards, developed by the World Health Organization multicentre growth reference study, describe normal child growth from birth to 5 years under optimal environmental conditions. These standards can be applied to all children everywhere, regardless of ethnicity, socioeconomic status and type of feeding^{55,56}.

If necessary, percentile values were transformed to Z-scores by the inverse-normal density function. When growth charts were unavailable, we used reported LMS parameters (median (M), generalized coefficient of variation (S) and skewness (L)) to obtain Z-scores via the formula:

$$Z\text{-score} = \begin{cases} \frac{(X/M)^L - 1}{L \cdot S}, & L > 0 \\ \ln(X/M)/S, & L = 0 \end{cases}$$

in which X is the observed value.

In adults (>18 years of age), we estimated LMS parameters when these were unavailable from the available sex-, age- and origin-matched Swiss (CoLaus),

Estonian or French control populations. For cases identified from population-based cohorts, *Z*-scores were directly inferred from the cohort.

Gene expression. We established lymphoblastoid cell lines from deletion and duplication carriers, as well as from controls (Supplementary Table 12), by transforming peripheral blood mononuclear cells with Epstein–Barr virus. Patients and controls were enrolled after obtaining appropriate informed consent via the physicians in charge, and approval by the ethics committee of the University of Lausanne. More control cell lines were obtained from Coriell Institute for Medical Research (<http://www.coriell.org/>) (Supplementary Table 12). SYBR Green real-time quantitative PCR (RT–PCR) was performed as previously described^{22,29}. Briefly, 1 µg of total RNA from lymphoblastoid cell lines was converted to complementary DNA using Superscript VILO (Invitrogen) primed with a mixture of oligo(dT) and random hexamers. Oligos were designed using the PrimerExpress program (Applied Biosystems) with default parameters (Supplementary Table 11). Non-intron-spanning assays were tested for genomic contamination in standard \pm reverse transcriptase reactions. The amplification efficiency of each primer pair was tested in a cDNA dilution series, as previously described³⁴. A full list of genes mapping in the rearranged interval, and exclusion criteria, are presented in Supplementary Table 1. All RT–PCR reactions were performed in a 10-µl final volume and triplicates per sample. The setup in a 384-well plate format was performed using a Freedom EVO robot (TECAN) and assays were run in an ABI 7900 sequence detection system (Applied Biosystems) with the following amplification conditions: 50 °C for 2 min, 95 °C for 10 min, and 45 cycles of 95 °C 15 s, then 60 °C for 1 min. A final incubation of 95 °C for 15 s followed by 60 °C for 15 s was carried out to establish a dissociation curve. Each plate included the appropriate normalization genes to control for any variability between plate runs. Raw threshold cycles (Ct) values were obtained using SDS2.4 (Applied Biosystems). To calculate the normalized relative expression ratio of individuals carrying the CNV and of controls, we used Biogazelle qBase Plus software³⁵ including geNorm³⁶. This program identified appropriate normalization genes (*EEF1A1*, *RPL13*, *GUSB* and *TBP*) having a gene-stability measure of $M = 0.25$. We note that one gene, *LAT*, showed a very high expression profile in one of the duplication samples (DASYL, Supplementary Table 13), reaching a relative expression value of 27.3 (s.e.m. = 1.37), compared to an average expression for other duplications of 1.89 (s.e.m. = 0.51). We cannot exclude that this finding is genuine (and confirmed it in a second experiment), but it was removed from further analyses as an outlier to give a more accurate overview of expression profiles for these genes.

In silico analysis was performed to check for brain, and specifically hypothalamus, expression of genes in the rearranged 16p11.2 interval (Supplementary Table 1). This was done using Allen Brain Atlas Resources, available from <http://www.brain-map.org>.

Cases with major neurological signs. Major neurological signs were defined by moderate to severe hypotonia, hypertonia, ataxia, spasticity, hyperreflexia, hyporeflexia and/or extra-pyramidal signs, and by the presence of epilepsy.

Statistics. Student's *t*-test: one-tailed *t*-tests were performed to test whether duplication carriers have *Z*-score values lower than zero for BMI, height and weight. We found this analysis more suitable than linear regression analysis, correcting for confounding factors such as sex and age, because these anthropometric traits have a highly nonlinear dependence on these factors, as can be observed in control populations.

Kruskal–Wallis test: this was used to test differences in the gene expression pattern between deletion and duplication carriers and control individuals. Because expression values are not necessarily normally distributed, this test is more adequate than a classical one-way analysis of variance. To test pairwise differences, we computed the difference in mean gene rank with its 95% confidence interval (as provided by the multcompare function in Matlab). Correction for multiple testing was done using a Bonferroni adjustment.

Multiple testing: we determined false-discovery-rate-based thresholds for association *P*-values for each phenotype, to correct for multiple testing. For each phenotype, we replaced the observed *Z*-scores with numbers randomly drawn from a standard normal distribution and performed the same *t*-tests for the same strata. The procedure was repeated 1,000 times. For various *P*-value thresholds, we asked how many tests would be declared significant for the null set on average (over the 1,000 random draws). The false discovery rate was estimated as the ratio of this number and the actual number obtained for the observed *Z*-scores. Thus, we controlled the dependence between nested tests.

Relative risk among adults, we defined underweight as a BMI <18.5 (WHO criteria). The estimated relative risk is the ratio of the fraction of underweight individuals among duplication carriers versus our control group. The standard error of log(relative risk) and its significance were calculated as previously described⁶¹. In our control group (population-based cohorts), the frequency of

being underweight is 1.9% (38 males and 148 females out of 9,470). Owing to the fact that being underweight decreases with age in the general population, we resampled our control group to ensure precise age-matching.

31. Coia, L. J. *et al.* cnvHap: an integrative population and haplotype-based multiplatform model of SNPs and CNVs. *Nature Methods* **7**, 541–546 (2010).
32. Olshen, A. B., Venkatraman, E. S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5**, 557–572 (2004).
33. Venkatraman, E. S. & Olshen, A. B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657–663 (2007).
34. Colella, S. *et al.* QuantiSNP: an objective Bayes Hidden-Markov model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res.* **35**, 2013–2025 (2007).
35. Wang, K. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
36. Korn, J. M. *et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nature Genet.* **40**, 1253–1260 (2008).
37. Marioni, J. C. *et al.* Breaking the waves: improved detection of copy number variation from microarray-based comparative genomic hybridization. *Genome Biol.* **8**, R228 (2007).
38. Diskin, S. J. *et al.* Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res.* **36**, e126 (2008).
39. Bengtsson, H., Irizary, R., Carvalho, B. & Speed, T. P. Estimation and assessment of raw copy numbers at the single locus level. *Bioinformatics* **24**, 759–767 (2008).
40. Bengtsson, H., Ray, A., Spellman, P. & Speed, T. P. A single-sample method for normalizing and combining full-resolution copy numbers from multiple platforms, labs and analysis methods. *Bioinformatics* **25**, 861–867 (2009).
41. Huang, J. *et al.* Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum. Genomics* **1**, 287–299 (2004).
42. Wang, B. *et al.* Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **316**, 1194–1198 (2007).
43. Schouten, J. P. *et al.* Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* **30**, e57 (2002).
44. Bunyan, D. J. *et al.* Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. *Br. J. Cancer* **91**, 1155–1159 (2004).
45. Ferrández, L. *et al.* Comparative study of three diagnostic approaches (FISH, STRs and MLPA) in 30 patients with 22q11.2 deletion syndrome. *Clin. Genet.* **68**, 373–378 (2005).
46. Slater, H. R. *et al.* Rapid, high throughput prenatal detection of aneuploidy using a novel quantitative method (MLPA). *J. Med. Genet.* **40**, 907–912 (2003).
47. Mei, Z. *et al.* Validity of body mass index compared with other body-composition screening indexes for the assessment of bodyfatness in children and adolescents. *Am. J. Clin. Nutr.* **75**, 978–985 (2002).
48. Physical status: the use and interpretation of anthropometry. Report of a WHO expert committee. *World Health Organ. Tech. Rep. Ser.* **854**, 1–452 (1995).
49. Kuczmarski, R. J. *et al.* CDC growth charts: United States. *Adv. Data* **314**, 1–27 (2000).
50. Sempé, M., Pedron, G. & Roy-Pernot, M. P. *Auxologie, Méthode et Séquences* (Thérapiex, 1979).
51. Roland-Cachera, M. F. *et al.* Body mass index variations: centiles from birth to 87 years. *Eur. J. Clin. Nutr.* **45**, 13–21 (1991).
52. Prader, A., Largo, R. H., Molinari, L. & Issler, C. Physical growth of Swiss children from birth to 20 years of age. First Zurich longitudinal study of growth and development. *Helv. Paediatr. Acta, Suppl.* **52**, 1–125 (1989).
53. Fredriks, M. *Growth Diagrams, 1997. Fourth Dutch Nation-wide Survey*. 233–242 (Bohn Stafleu Van Loghum, 1997).
54. de Onis, M. *et al.* Development of a WHO growth reference for school-aged children and adolescents. *Bull. World Health Organ.* **85**, 660–667 (2007).
55. de Onis, M., Garza, C., Onyango, A. W. & Borghi, E. Comparison of the WHO child growth standards and the CDC 2000 growth charts. *J. Nutr.* **137**, 144–148 (2007).
56. Mei, Z., Ogden, C. L., Flegal, K. M. & Grummer-Strawn, L. M. Comparison of the prevalence of shortness, underweight, and overweight among US children aged 0 to 59 months by using the CDC 2000 and the WHO 2006 growth charts. *J. Pediatr.* **153**, 622–628 (2008).
57. Molina, J. *et al.* Abnormal social behaviors and altered gene expression rates in a mouse model for Potocki–Lupski syndrome. *Hum. Mol. Genet.* **17**, 2486–2495 (2008).
58. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) method. *Methods* **25**, 402–408 (2001).
59. Hellebrans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **8**, R19 (2007).
60. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, 1–11 (2002).
61. Morris, J. A. & Gardner, M. J. Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. *Br. Med. J. (Clin. Res. Ed.)* **296**, 1313–1316 (1988).

4.3 Microdeletion and microduplication in 10q11.22

Unpublished results

Mirror effects for Autism Spectrum Disorder due to gene dosage at 10q11.22 affecting *GPRIN2* and *PPYR1*.

Introduction

Autism spectrum disorders (ASDs) represent a group of neurodevelopmental disorders that are characterized by impaired reciprocal social interactions, delayed or aberrant communication, and stereotyped, repetitive behaviours, often with restricted interests (Hu 2011). The prevalence for these disorders is now estimated at 1% (Gillbert 1999, Forbonne 2003, Kogan 2009). With a concordance rate as high as 90% in monozygotic twins and 2-10% in dizygotic twin pairs (Folstein 2001), ASD is among the most heritable of neuropsychiatric conditions. Although autism or autism features often occur in single gene disorders such as Tuberous Sclerosis (MIM#191100) and Fragile X syndrome (MIM#300624) (Gillberg and Coleman 2000), these disorders only explain around 2-5% of the autism cases (Baker 1998, Carney 2003; Kielinen 2004, Volkmar 2005, Hatton 2006). Thus a considerable amount of effort has been devoted to identifying genetic mutations or variants that associate with these disorders.

Until recently, karyotyping has been the standard method for the detection of cytogenetic aberrations in patients with developmental disorders. The development of whole-genome screening methodologies for the detection of CNVs, such as array-CGH, provides a much higher resolution than karyotyping leading to the identification of novel microdeletion and microduplication syndromes, such as deletions and duplications in chromosome band 15q13.2q13.3, 16p11.2, and 17p11.2, often associated with an autism phenotype (Ballif et al., 2007; Potocki et al., 2007; Weiss et al., 2008; Miller et al., 2009). The discovery of an increasing number of genomic disorders, allowed the identification of NAHR as the predominant underlying molecular mechanism using the segmental duplication or LCRs as recombination substrates (Stankiewicz and Lupski 2010). LCRs have been defined as human DNA fragments >1 Kb in size and of 90% DNA sequence identity that can mediate constitutional and somatic genomic rearrangements (Stankiewicz and Lupski 2010). The constantly increasing resolution of the arrays has further

improved the detection of copy number abnormalities down to single genes and is likely to provide new advances in the autism genetics field. Although clinical genetic laboratories are familiar with recurrent copy-number changes mediated by segmental duplication architecture, population studies suggest that the vast majority of copy-number variation is not recurrent (Itsara 2009). Even if array-CGH offers the sensitivity of high-resolution genome-wide detection of clinically significant CNVs, the additional challenge of interesting variants of uncertain clinical significance can impose a burden on clinicians and laboratories (Vos 2009).

We recently reported that a small duplication on 10q11.22 including *GPRIN2* gene, a regulator of neurite outgrowth, and *PPYRI*, a gene involved in energy homeostasis, is a candidate modifier for Rett syndrome (Artuso 2011). Specifically, duplications were found in the Zappella variant, the Rett variant with recovery of speech, and lacking the typical growth delay, underweighting and autistic features. Since *PPYRI* knockout mice display underweight and reduced white adipose tissue (Sainsbury et al.) we supposed that an over-expression of *PPYRI* due to gene duplication may be responsible for the higher body weight characterizing Zappella variant. In Artuso et al, we concluded that duplication at 10q11.22 may play a role in protecting from both underweighting and autistic features in Rett patients (Artuso 2011).

In the present study, we explored the association of CNV at 10q11.22 with ASD in a cohort of 1394 patients referred for a wide range of referring diagnoses, including Syndromic Intellectual Disability (SID), Non Syndromic Intellectual Disability (NSID), ASD and MCA.

Materials and methods

Cases and controls.

This study collected patients with SID, NSID, ASD and MCA, obtained from 2 sources. Patients were ascertained by the Medical Genetics Unit of Siena, Italy (n=304), and by the Laboratory of Genetics Diagnosis, IRCCS Oasi SS Maria of Troina, Italy (n=1090), 320 control subjects were collected for this study. Experiments were performed on genomic DNA extracted from peripheral blood samples from each patient after informed consent approved by the local Institutional

Review Board. Moreover two additional centres have been contacted: the Laboratoire Génétique Chromosomique, Hôpital Couple enfant, CHU Grenoble, France, and the Unidad de Neurología Infanto-juvenil Hospital Universitario Quiron Centro CADE, Madrid, Spain. We are still collected data from these centres

Array-CGH

Array-CGH analysis was performed using commercially available oligonucleotide microarrays containing about 105.000 60-mer probes (Human Genome CGH Microarray 105K Kit respectively, Agilent Technologies, Santa Clara, California) as previously reported by Pescucci et al.

Multiplex Ligation Probe Amplification (MLPA)

MLPA probes were designed according to protocols available at MRC Holland website (<http://www.mrc-holland.com/pages/indexpag.html>). Two and three MLPA probes targeted the *GPRIN2* and the *PPYR1* genes, respectively. MLPA analysis was carried out essentially as described by Schouten et al. PCR products were identified and quantified by capillary electrophoresis on an ABI 3130 genetic analyzer, using the Gene Mapper software from Applied Biosystems, Foster City, CA. In order to process efficiently the MLPA deletion/ duplication data, a spreadsheet was generated in Microsoft Excel. First, the data corresponding to each sample (patient's and control's DNAs) were normalized by dividing each probe's signal strength (i.e., the area of each peak) by the average signal strength yielded by the 10 control probes to generate for each peak a Relative Peak Area (RPA) value. The RPA value for each probe in the patient's sample was then compared to that of a control's sample by dividing, for each peak, the patient's RPA by the control's RPA. The latter ratio was then used to define the following categories: (i) $\square 1$, for the non-deleted/non-duplicated gene region, (ii) $\square 0,5$ if deleted, (iii) $\square 1,5$ if duplicated.

Statistical Analysis

To assess the significance of the frequency of recurrent 10q11.2 CNVs in ASD or SID/NSID patients and controls, a Fischer's exact test was used.

Results

Identification of CNVs in 10q11.22

In a first analysis performed by array-CGH in our cohort of patients, we identified 12 individuals sharing a overlapping CNV in 10q11.22 (3 deletions and 9 duplications). The patients carrying the deletion were classified as ASD while the duplicated patients were classified as SID and NSID. The identified CNVs ranged in size from ~0.17 to ~1.16 Mb (Tab.1 and Fig.1). The smallest region of overlap of approximately 0.17 Mb included only two genes: the pancreatic polypeptide receptor 1 (*PPYRI*) and the G protein regulated inducer of neurite outgrowth 2 (*GPRIN2*) (Fig.1).

In order to investigate a possible association of CNVs at 10q11.22 (46,976,157-47,148,490) with ASD we collected additional patients. Among our cohort we selected 292 patients that have been previously analysed by array-CGH 44K and were negative for deletions and duplications in 10q11.22. Because the 44K slides have only one probe located in the 10q11.22 region, we decided to reanalyze these cohort of patients by MLPA. An additional cohort of 1090 patients was collected from Troina (Italy). We divided the collected patients in ASD group (398 individuals) and SID/NSID group (984 individuals). Moreover we included in the MLPA analysis a cohort of 320 control subjects. We identified 7 deletions in the ASD group, 4 deletions in the SID/NSID group while no deletions were found in the control group. We also analysed the ASD, SID/NSID and control group looking for duplications in the 10q11.22 region. Seven duplications were found in the ASD group, 43 were found in the SID/NSID group and 10 in the control group. Combining the results obtained by MLPA and by array-CGH analysis, we obtained a total of 10/401 deletions and 7/401 duplications in the ASD group, 4/993 and 52/993 deletions and duplications respectively in the SID/NSID group and 10/320 duplications in the control group.

Statistical analysis of our preliminary results shows a significantly correlation between the presence of the 10q11.22 deletion and the ASD phenotype if compared both with SID/NSID and control group (tab.2). On the contrary the duplication is more frequent between SID/NSID and in control cases than in ASD (tab.2).

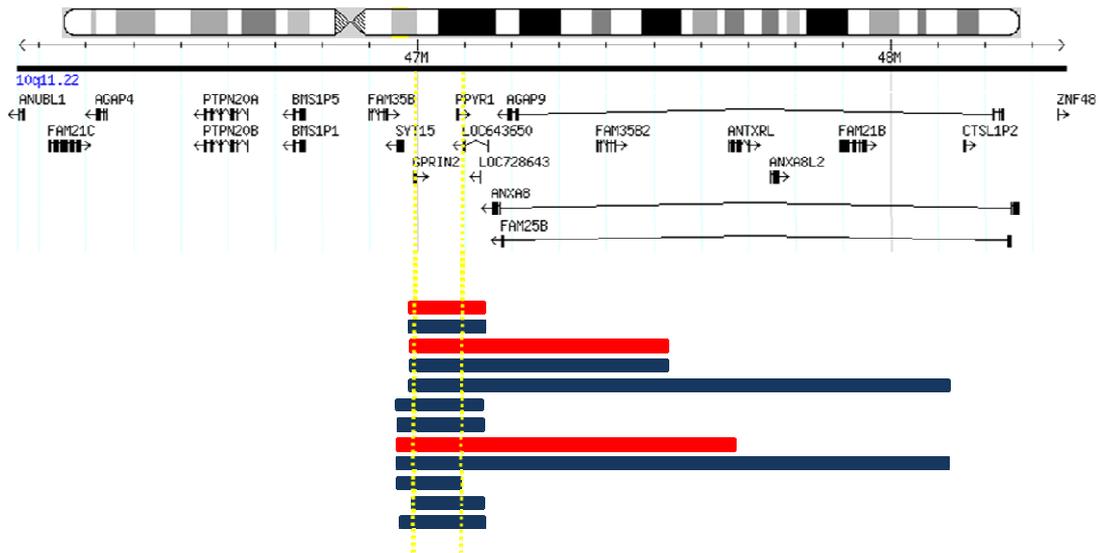


Fig.1 The extent of the deleted (red) and duplicated (blue) area in the twelve patients analysed by array-CGH. Upper panel: ideogram of chromosome 10. Dotted line: small region of overlap.

Identification of additional copy number changes in patients

Three individuals with deletions and 3 individuals with duplications had secondary copy number alterations. Four of the additional CNVs in patients #79, #681, #283 and #384 were inherited from phenotypically normal parents. The parental origins of the additional CNVs in patients #1275 and #1410 were unknown. Moreover, three duplicated patients (#139, #368 and #601) showed a mutation in *MECP2* gene responsible of both the classical and the preserved speech variant form (Zappella variant) of Rett syndrome.

Tab.1 Deletions and duplications in 10q11.22 identified by array-CGH.

<i>Patient</i>	<i>Coordinates (hg19)</i>	<i>Size (Mb)</i>	<i>Gain / Loss</i>	<i>Microarray platform</i>	<i>Additional CNVs / single gene mutation</i>
#384	46,976,157-47,148,490	0,17	Loss	Agilent 105K	arr 14q21.2(41,018,728-41,310,931)x1
#1453	46,976,157-47,547,592	0,57	Loss	Agilent 105K	
#1905	46,951,237-47,678,024	0,73	Loss	Agilent 105K	
#424	46,951,237-47,148,490	0,19	Gain	Agilent 105K	
#39	46,951,237-47,148,490	0,19	Gain	Agilent 105K	
#1391	46,951,237-48,115,466	1,16	Gain	Agilent 105K	
#1410	46,951,237-47,086,737	0,13	Gain	Agilent 105K	arr 3p22.3(35,290,648-35,361,705)x3
#2307	46,988,690-47,148,490	0,16	Gain	Agilent 244K	
#2202	46,964,973-47,148,490	0,18	Gain	Agilent 244K	
#139	46,976,157-47,148,490	0,17	Gain	Agilent 105K	MECP2 (c.1157del32)
#368	46,976,157-48,115,466	1,14	Gain	Agilent 105K	MECP2 (p.R133C)
#601	46,976,157-47,547,592	0,57	Gain	Agilent 105K	MECP2 (c.1163del26)

Tab.2 Fisher's exact test of 10q11.22 deletions and duplications

	ASD	SID/NSID	Control	ASD vs. SID/NSID	ASD vs. Control
Deletion	10	4	0	p=0,001	p=0,002
Duplication	7	52	10	p=0,001	p=0,09
Total cohort	401	993	320		

Association of obesity frequency in deleted and duplicated cohort

Auxological parameters were available only for 41 cases of the 80 showing a CNV in chromosome 10. We had height and weight measures for 11 patients carrying a 10q11.22 deletion and for 30 patients showing the duplication. Three deleted patients were overweight, 7 were normal and 1 was underweight. Among the duplicated cohort 17 cases were overweight, 10 were normal and 3 were underweight. We considered the frequency of both overweight and underweight in the two cohorts. The underweight frequency was almost the same in the deleted and duplicated cohort (9% and 10% respectively). Otherwise the overweight phenotype was more frequent in the duplicated cohort (56,6%) than in the deleted one (27,3%). We performed the same analysis taking in account only the patients carrying the smallest rearrangements. We collected 1 deleted case (#384) and 4 duplicated cases (#139, #1410, #2307, #2202,). We excluded two patients (#424 and #39) because carried a duplication including also *SYT15* gene. The deleted patients had a normal BMI, while 75% (3/4) of the duplicated patients were overweight.

Tab.3 Auxological parameters of deleted patients

<i>Patient</i>	<i>*Age</i>	<i>Gender</i>	<i>Height</i>	<i>Weight</i>	<i>OFC</i>	<i>BMI</i>	
#1227	19y3m	M	182 cm	126 Kg	58 cm	38	O
#1187	24y3m	F	158 cm	76 Kg	56 cm	30,4	OW
#79	9y6m	F	138 cm	33 Kg	53 cm	17,3	N
#1275	13y8m	M	174 cm	140 Kg	58,5 cm	46,2	O
#384	12y6m	M	157 cm	43 Kg	57 cm	17,4	N
#1453	1y9m	M	80,5 cm	10 Kg	48 cm	15,6	N
#1905	4y4m	F	105 cm	13 Kg	47 cm	11,8	UW
#03099	NA	F	NA	NA	NA	21,4	N
#02873	NA	F	NA	NA	NA	17,2	N
#03632	NA	M	NA	NA	NA	17,3	N
#03660	NA	F	NA	NA	NA	16,4	N

OW (overweight); UW (underweight); O (obese); N (normal); NA (not available)

* Age at the clinical evaluation.

Tab.4 Auxological parameters of duplicated patients

<i>Patient</i>	<i>*Age</i>	<i>Gender</i>	<i>Height</i>	<i>Weight</i>	<i>OFC</i>	<i>BMI</i>	
#681	14y11m	M	182 cm	55 Kg	57 cm	16,6	UW
#2060	7y4m	M	127 cm	36 Kg	53 cm	22,3	O
#283	3y8m	M	102 cm	17 Kg	51 cm	16,3	N
#424	9y	M	153 cm	44 Kg	54,5 cm	18,8	OW
#39	12y9m	M	153 cm	49 Kg		20,9	N
#1391	11y10m	F	158 cm	63 Kg	54,5 cm	25,2	OW
#1410	4y	F	108 cm	22,1 Kg	51 cm	18,8	OW
#1139	9y5m	F	135,5 cm	29 Kg	51 cm	15,9	N
#2307	10y3m	M	146 cm	36 Kg	50,5 cm	16,9	N
#2202	13y3m	F	160 cm	70 Kg	55,5 cm	27	OW
#139	NA	F	NA	NA	NA	NA	OW
#368	NA	F	NA	NA	NA	NA	OW
#601	10y	F	154 cm	60 Kg	54 cm	25,3	OW
#01269	NA	F	NA	NA	NA	23,5	N
#01860	NA	M	NA	NA	NA	28	OW
#02169	NA	F	NA	NA	NA	15,3	N
#02193	NA	F	NA	NA	NA	30	OW
#03270	NA	F	NA	NA	NA	17,1	OW
#03284	NA	M	NA	NA	NA	23,3	OW
#03284S	NA	F	NA	NA	NA	22,4	OW
#03324	NA	F	NA	NA	NA	22,2	OW
#03431	NA	F	NA	NA	NA	30	OW
#03857	NA	M	NA	NA	NA	24,8	OW
#03877	NA	M	NA	NA	NA	19,4	N
#02910	NA	M	NA	NA	NA	21,4	N
#02910F	NA	M	NA	NA	NA	17	UW
#02651	NA	F	NA	NA	NA	15,6	N
#02594	NA	M	NA	NA	NA	21,2	N
#02155	NA	M	NA	NA	NA	11,5	UW
#02980	NA	F	NA	NA	NA	28,6	OW

OW (overweight); UW (underweight); O (obese); N (normal); NA (not available)

* Age at the physical evaluation.

Discussion

In a first analysis we observed 12 individuals sharing a 10q11.22 CNV. Three were deleted and classified as ASD, the other 9 cases were duplicated and classified as SID/NSID. Although features of developmental delay and dysmorphisms are already documented (Stankeiwicz 2011), an analysis of CNV-phenotype association has not been carried out and this CNV has not been classified as pathogenic. To investigate the nature of this CNV we collected additional patients from Italy (Siena and Troina), France and Spain. The group of patients reported herein represents the largest collection of individuals with microdeletions or microduplications within chromosome 10q11.22 reported in the literature. In the present study we divided our cohort according to the technique used for the analysis (array-CGH and MLPA) and each cohort was additionally divided into two groups: the ASD group and the SID/NSID group. We compared the frequency of 10q11.22 rearrangements in the ASD group in the SID/NSID group and in control group, in order to determine whether the deletion predisposes individuals to an abnormal phenotype. In total we identified 10 and 4 deletions in the ASD and SID/NSID group respectively but no deletions were found in the control group. The reciprocal duplication has also been reported in literature (Stankaiwicz 2011). Therefore we checked our cohort also for the presence of duplications in 10q11.22. We found 7 and 52 duplications in the ASD and SID/NSID group respectively. Unlike the deletion, the duplication has been detected in 10/320 control subjects suggesting that the duplication had a less penetrance. Moreover some individuals carried additional genomic imbalances (tab.1) which could modify the phenotype of these patients.

Stankeiwicz et al. recently reported 24 cases with deletions and 17 cases with duplications at 10q11.21q21.1. The ~66% of the reported rearrangements were flanked by large, directly oriented segmental duplications of 98% sequence identity, suggesting that NAHR caused these genomic rearrangements. Rearrangement in 10 of 12 individuals may have been caused by NAHR between LCR 10q11.2A and LCR 10q11.2B and 2 by LCR 10q11.2A and LCR 10q11.2C. The smallest overlapping genomic imbalance in 10q11.22 was mapped to be ~170 kb. Only two genes are located in this region, *GPRIN2* and *PPYR1* (fig.1).

GPRIN2 is highly expressed in the cerebellum and interacts with activated members of the Gi subfamily of G protein α subunit and functions together with *GPRIN1* to regulate neurite outgrowth (Iida and Kozasa 2004). The fact that *GPRIN2* is exclusively expressed in the cerebellum suggests that it could be involved in the ASD phenotype when deleted. The differences in deletion and duplication prevalence between the ASD group and the control populations are statistically significant for deletion ($p=0,002$) and not significant for duplications ($p=0,09$). Because the 10q11.22 deletion is significantly enriched in the ASD population when compared to the SID/NSID population ($p=0,001$), we propose that this microdeletion is probably clinically relevant and responsible for the ASD phenotype.

The *PPYRI* gene is a key regulator of energy homeostasis and directly involved in the regulation of food intake. *PPYRI*, also named as neuropeptide Y receptor or pancreatic polypeptide 1, is a member of the seven transmembrane domain-G-protein coupled receptor family. Genetic variation studies have reinforced the potential influence of *PPYRI* on body weight in humans. Pancreatic polypeptide is the preferential *PPYRI* agonist. Peripheral administration of pancreatic polypeptide inhibits gastric emptying and decreases food intake in humans (Sha 2009). This effect is mediated by direct action on local *PPYRI* within the arcuate nucleus. Sha et al, demonstrated that subjects with 10q11.22 loss had 12.4% higher BMI value, and subjects with 10q11.22 gain had 5.4% lower BMI value when compared to normal diploid subjects. *PPYRI* null animals showed, for instance, an opposite result. Knockout mice displayed lower body weight and reduced white adipose tissue accompanied with increased plasma levels of pancreatic polypeptide (Sainsbury et al. 2002).

In order to confirm a correlation between *PPYRI* and body weight, we checked the deleted and duplicated patients of our cohort for BMI. Out of a total of 80 cases with 10q11.22 imbalances, weight and height information were available for 41 (11 deleted and 30 duplicated cases) (tab.3 and tab.4). We noticed that among the deleted patients only the 27,3% (3/11) were overweight, while the 56,6% (17/30) of duplicated patients showed a high BMI. Even if not statistically significant the overweight phenotype was more frequent among the duplicated cohort than among the deleted one. Moreover one of the deleted cases showing overweight (#1275)

presented a second rearrangement that could be responsible for the high BMI. Because we were interested in the influence of *PPYRI* gene on BMI, we took into account only patients carrying the smallest rearrangement. Surprisingly we noticed that none of the deleted patients were overweight, while 75% (3/4) of the duplicated patients had an increased BMI. Our data suggested that a higher level of *PPYRI* expression due to gene duplication may correlate with the overweight reported in our cases. These results are in contrast with those reported by Sha that showed an association between the 10q11.22 loss and a higher body mass index value in the Chinese population. A possible explanation could be represented by the different size of the rearranged region. The CNV reported by Sha is larger with respect to the small region of overlap reported here and includes two additional genes, *SYT15* and *LOC728643*. These two genes have not been reported to have relation with any obesity phenotype. *Syt15* mRNA has been found in different tissues (i.e. heart, lung, skeletal muscle and testis) but unlike other Syt family members was absent in the brain. Moreover *Syt15* C2 domains lack Ca²⁺-dependent phospholipid binding activity. These results suggested that *Syt15* may be involved in constitutive membrane trafficking in selected non-neuronal tissues (Fukuda 2003). However it is still unknown whether the interactions of the four genes may lead to the BMI variation.

In conclusion, our results suggested that recurrent reciprocal microdeletions and microduplications within 10q11.22 represent novel genomic disorders consisting of ASD and SID/NSID phenotype respectively. The duplication was observed also in several controls, suggesting that the duplication confers either no phenotype at all or a range of phenotypes of varying severity. Moreover contrasting result in BMI association analysis exist between patients with the deletion and the reciprocal duplication. In fact an high BMI was more frequently observed in microduplicated than in deleted patients.

Overall our findings have important implications for genetic counselling. CVNs such as those described in this report are often associated with unpredictable and variable phenotypic outcomes and pose diagnostic and counselling difficulties. However, the analysis of additional patients and controls with 10q11.22 rearrangements is required

to reinforce this hypothesis and to obtain a better insight in the potential pathology associated with the observed microdeletion and microduplication events.

References

1. Artuso R, Papa FT, Grillo E, Mucciolo M, Yasui DH, Dunaway KW, Disciglio V, Mencarelli MA, Pollazzon M, Zappella M, Hayek G, Mari F, Renieri A, Lasalle JM, Ariani F. Investigation of modifier genes within copy number variations in Rett syndrome. *J Hum Genet.* 2011 Jul;56(7):508-15.
2. Baker P, Piven J, Sato Y. 1998. Autism and tuberous sclerosis complex: Prevalence and clinical features. *J Autism Dev Disord* 28(4): 279–285.
3. Ballif BC, Hornor SA, Jenkins E, Madan-Khetarpal S, Surti U, Jackson KE, Asamoah A, Brock PL, Gowans GC, Conway RL, et al. 2007. Discovery of a previously unrecognized microdeletion syndrome of 16p11.2-p12.2. *Nat Genet* 39(9): 1071–1073.
4. Carney RM, Wolpert CM, Ravan SA, Shahbazian M, Ashley-Koch A, Cuccaro ML, Vance JM, Pericak-Vance MA. 2003. Identification of MeCP2 mutations in a series of females with autistic disorder. *Pediatr Neurol* 28(3): 205–211.
5. Folstein SE, Rosen-Sheidley B. 2001. Genetics of autism: Complex aetiology for a heterogeneous disorder. *Nat Rev Genet* 2(12): 943–955.
6. Fombonne E. 2003. Epidemiological surveys of autism and other pervasive developmental disorders: An update. *J Autism Dev Disord* 33(4): 365–382.
7. Fukuda M. Molecular cloning and characterization of human, rat, and mouse synaptotagmin XV. *Biochem Biophys Res Commun.* 2003 Jun 20;306(1):64-71.
8. Gillberg C, Coleman M. 2000. *The Biology of the Autistic Syndromes.* London, UK: Mac Keith Press, Distributed by Cambridge University Press.
9. Gillberg C, Wing L. 1999. Autism: Not an extremely rare disorder. *Acta Psychiatr Scand* 99(6): 399–406.
10. Hatton DD, Sideris J, Skinner M, Mankowski J, Bailey DB Jr, Roberts J, Mirrett P. 2006. Autistic behavior in children with fragile X syndrome: Prevalence, stability, and the impact of FMRP. *Am J Med Genet Part A* 140A(17): 1804–1813.
11. Hu VW, Addington A, Hyman A. Novel autism subtype-dependent genetic variants are revealed by quantitative trait and subphenotype association analyses of published GWAS data. *PLoS One.* 2011 Apr 27;6(4):e19067.
12. Iida N, Kozasa T. Identification and biochemical analysis of GRIN1 and GRIN2. *Methods Enzymol.* 2004;390:475-83.
13. Itsara, A., Cooper, G.M., Baker, C., Girirajan, S., Li, J., Absher, D., Krauss, R.M., Myers, R.M., Ridker, P.M., Chasman, D.I., et al. (2009). Population analysis of large copy number variants and hotspots of human genetic disease. *Am. J. Hum. Genet.* 84, 148–161.
14. Kielinen M, Rantala H, Timonen E, Linna SL, Moilanen I. 2004. Associated medical disorders and disabilities in children with autistic disorder: A population-based study. *Autism* 8(1): 49–60.

15. Kogan MD, Blumberg SJ, Schieve LA, Boyle CA, Perrin JM, Ghandour RM, Singh GK, Strickland BB, Trevathan E, van Dyck PC. 2009. Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. *Pediatrics* 124(5): 1395–1403.
16. Miller DT, Shen Y, Weiss LA, Korn J, Anselm I, Bridgemohan C, Cox GF, Dickinson H, Gentile J, Harris DJ, et al. 2009. Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders. *J Med Genet* 46(4): 242–248.
17. Pescucci C, Caselli R, Grosso S, et al. 2q24-q31 deletion: report of a case and review of the literature. *Eur J Med Genet* 2007;50(1):21-32.
18. Potocki L, Bi W, Treadwell-Deering D, Carvalho CM, Eifert A, Friedman EM, Glaze D, Krull K, Lee JA, Lewis RA, et al. 2007. Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am J Hum Genet* 80(4): 633–649.
19. Sainsbury A, Schwarzer C, Couzens M, Fetissov S, Furtinger S, Jenkins A, Cox HM, Sperk G, Hökfelt 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 Jun 25;99(13):8938-43. Epub 2002 Jun 18.
20. Schouten JP, McElgunn CJ, Waaijer R, et al. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*. 2002 Jun 15;30(12):e57
21. Sha BY, Yang TL, Zhao LJ, Chen XD, Guo Y, Chen Y, Pan F, Zhang ZX, Dong SS, Xu XH, Deng HW. Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population. *J Hum Genet*. 2009 Apr;54(4):199-202. Epub 2009 Feb 20.
22. Stankiewicz P, Kulkarni S, et al. Recurrent deletions and reciprocal duplications of 10q11.21q11.23 including CHAT and SLC18A3 are likely mediated by complex low-copy repeats. *Hum Mutat*. 2012 Jan;33(1):165-79. doi: 10.1002/humu.21614.
23. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med*. 2010;61:437-55. Review.
24. Volkmar FR, Paul R, Klin A, Cohen D. 2005. *Handbook of Autism and Pervasive Developmental Disorders*. New Jersey: Wiley.
25. Vos, J., van Asperen, C.J., Wijnen, J.T., Stiggelbout, A.M., and Tibben, A. (2009). Disentangling the Babylonian speech confusion in genetic counseling: an analysis of the reliability and validity of the nomenclature for BRCA1/2 DNA-test results other than pathogenic. *Genet. Med.* 11, 742–749.
26. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, et al. 2008. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 358(7): 667–675.

4.4 Microdeletion unmasking recessive phenotype.

Unpublished results

Recessive likely pathogenic variants unmasked by microdeletion syndromes with unusual phenotypes

Introduction

Microdeletion of chromosome 22q11.2 or 22q11.2 deletion syndrome (22q11.2DS) (MIM#188400/#192430) is the most common human deletion syndrome with an estimated prevalence of 1 in 4,000 live births (Goodship 1998). Up to 93% of cases occurs de novo, whereas in the remaining 7% the deletion is found to be inherited from a parent.

The high frequency of the 22q11.2 deletion can be explained by the presence of chromosome-specific low copy repeats flanking (LCR A and D) or within the typically deleted region (LCR A, B and C) (Edelmann 1999, Shaikh 2000). Since LCRs present chromosome-specific repeated DNA sequences, they can be prone to misalignment during meiosis and unequal recombination exchanges, resulting in chromosome rearrangements in the 22q11.2 region. Shaikh et al. stated that 22q11.2 LCRs share 97.98% nucleotide sequence identity. The size and the homology among them seem to be related to the frequency of each type of deletion. The 3 Mb deletion is the most frequent one (90% of cases), since it is mediated by the largest LCRs, A and D, which share 250 kb of duplicated sequence in a complex arrangement. On the other hand, the 1.5 Mb deletion (8%) is flanked by LCRs A and B, which share a common block of 135 kb. Some smaller or atypical deletions have been reported but there is no evidence for specific genotype–phenotype correlations. It has been argued that the 1.5Mb deletions contain all key genes responsible for the syndrome (Carlson et al., 1997).

The phenotypic spectrum encompasses several previously described syndromes including DiGeorge, velocardiofacial and conotruncal anomaly face syndromes as well as some individuals with other conditions such as Cayler cardiofacial syndrome. The phenotypic expression of the 22q11.2DS is known to be highly variable and ranges from a severe life-threatening condition to affected individuals with few associated features (Bassett et al. 2005; Kobrynski and Sullivan 2007; Ryan et al. 1997). Abnormal development of the pharyngeal arches and pharyngeal pouches

gives rise to the cardinal physical manifestations of the syndrome: conotruncal anomaly, hypocalcemia due to dysfunctional parathyroid glands, palatal abnormalities and paediatric immunodeficiency that may be secondary to hypo/aplasia of the thymus (Lindsay et al. 2001; Scambler 2000). Major heart defects are present in about 40% of cases while minor anomalies, e.g., of the aortic arch, may be identified only on cardiac ultrasonography. Overt cleft palate is rare, whereas submucous cleft palate associated with velopharyngeal insufficiency is characteristic of 22q11.2DS. In contrast, the facial features are considered a constant manifestation of the syndrome (Guyot et al. 2001), although the overall facial appearance is not always readily identifiable even to informed clinicians.

Developmental delays and learning difficulties are very commonly associated, although severe intellectual disability is rare. Recurrent seizures are common, especially those related to hypocalcemia, and epilepsy may be present in about 5% of patients. Psychiatric conditions may be present in children and over 60% of patients develop treatable psychiatric disorders by adulthood (Bassett et al. 2005). This risk is a major concern for families. In particular, due to the high frequency of schizophrenia in 22q11.2DS patients, the 22q11.2 region is considered to be one of the main schizophrenia susceptibility loci in humans (Bassett and Chow 2008; Insel 2010). Evidence from multiple studies indicates that about 1% of individuals with schizophrenia in the general population have 22q11.2 deletions (Basset et al, 2010).

The commonly deleted region in 22q11.2 encompasses approximately 45 genes and most of them are expressed in fetal and adult brain, thus are candidates for both the psychiatric phenotype of patients with 22q11.2 deletions and susceptibility to psychiatric disorders in the general population (Meechan et al. 2010). As clinical variability is not explained by differences in gene content within the deletion, allelic variation(s) in the non-deleted homologous region is considered a possible contributor to phenotypic variability.

In order to identify possible recessive alleles we performed targeted sequencing on three patients with a 22q11.2 deletion and an atypical phenotype (MURCS, severe intellectual deficit with polydactyly and Cayler syndrome) in collaboration with the University of Geneva.

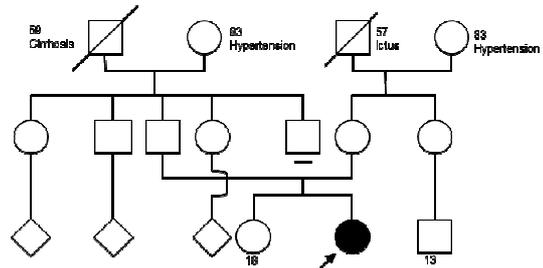
Case #1

At birth: cleft palate, polydactyly in both hands and both feet, ventricular septal defect, bilateral congenital leukoma and iris and retinal coloboma.

Psychomotor retardation and a period of regression.

16y: long face, long nose, narrow and up-slanting palpebral fissures, short stature, hypotelorism.

Karyotype: inv15
MECP2: normal



Additional information

Father:

- No DNA
- Normal karyotype
- Coloboma
- Microcythemia
- Dialysis since he was 39

Mother:

- Inv15
- No del22q11.2
- No *MED15* mutation

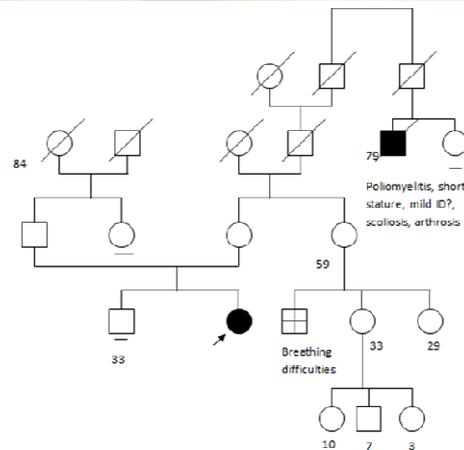
Case #2

At birth: weight 3100 Kg (50° cnt), length 45 cm (<10° cnt), head circumference reported to be normal.

Speech delay, frequent infections and fractures, growth curve always underweight.

22y: short stature (1,42 cm, <5° cnt), obesity (BMI 30,7), head circumference of 52 cm (around 3° cnt), long face, tubular nose with bulbous tip, high nasal bridge and small ears (5.2 cm, <-2SD), flat feet, nasal voice.

Bicornuate uterus, renal agenesis, hypothyroidism, aortic arch anomalies, C2-C3 fusion.



Additional information

Father:

- Duodenal ulcer
- Radio-dermatitis of the hands
- Renal cysts
- Episodes of macrohematuria

Mother:

- Reduced motility of the neck
- Carpal tunnel surgery
- Fibromatosus uterus
- Uterine myomas

Materials and methods.

Target sequence

The libraries for paired end sequencing were prepared with an Illumina library prep kit and captured with a custom made Agilent capture kit designed for the 3Mb deletion region. The kit was able to capture the 3 Mb of the classical VCFS region plus ~200 Kb upstream and downstream to the breakpoints. It didn't capture the repeated regions. Briefly, 3 µg of DNA were sheared using the Covaris instrument. After fragmentation, the ends were irregular with 3' and 5' overhangs, so the "ends repairing" was performed. This step converted all the ends into blunt ends using T4 polymerase and Klenow DNA polymerase enzymes. The latter enzyme had a 3' to 5' exonuclease activity, removing 3' overhangs. The polymerase instead refilled the 5' overhangs. Finally a T4 polynucleotide kinase phosphorylated the 5' ends. The phosphorylation of the 5' ends was a necessary step for the ligation of the index-specific paired-end adapter. The capture process continued with the hybridization, in which biotinylated fragments were added. This fragments were complementary to the fragment of interest and can be isolated using streptavidin coated beads. The biotinylated baits were then removed and the index tags were added. The final step was the pooling of the sample. The samples were sequenced in a HiSeq2000.

The pipeline

The obtained reads were aligned to a reference genome with BWA. On average, the samples had 99% of the target region covered at least 8x. SNVs and small indels were called using Samtools, that recognized the data in a format that described the base pair information at each chromosomal position; and Pindel2, that identified long insertions or deletions. Finally the variants were annotated with Annovar.

Results

We obtained from the sequencing a total of 440 millions reads. The percentage of reads of each sample was about the same and the 99% of each sample had an 8 fold coverage.

We decided to start with the analysis of the coding regions. We proceeded by applying different and consecutive filters. We removed the synonymous variants, the variants already reported as segmental duplication or already reported in the SNP database or in the 1000 genome project (tab.1). We found only 1 variant in Case#1 (tab.2). It was a non-synonymous variant occurred in the *MED15* gene. Because we didn't find any mutation in the coding regions in Case#2 and #3, we proceeded with the analysis of the genome data of the 22q11 region. We divided the data in 2 files, one containing all the variants called by Samtools and the other one with the variants called by Pindel. Again we applied different filters. We removed all the variants outside of the patient's deletion; the variants already reported as segmental duplication or already reported in the SNP database or in the 1000 genome project (tab.3). We found only 1 variant in Case#2 (tab.4). The mutation was located in a non-coding RNA occurring between *SEPT5* and *GP1BB* genes. In Case#3 we didn't find mutation.

In both Case#1 and #2, we confirmed the identified mutation by Sanger. For Case#1 we had the DNA from the mother but not from the father. The mother sequence was normal and therefore we can't define the mutation's origin. For Case#2 we had DNA from both parents and we confirmed that the mutation was inherited from the mother.

Tab.1 22q11.2 exome data analysis

	Total # variants	Filters	Final # variants
Case#1	76	Synonymous, SegDup, dbSNP, allele frequency >1%	1
Case#	44	Synonymous, SegDup, dbSNP, allele frequency >1%	0
Case#3	77	Synonymous, SegDup, dbSNP, allele frequency >1%	0

Tab.2 22q11.2 exome result

Gene	ExonicFunc	AChange	SIFT	PolyPhen2	Ref	Obs	Otherinfo
MED15	Nonsynonymous SNV	NM_015889:c.G2141A.p.R714H	0	0,973	G	A	hom

Tab.3 22q11.2 genome data analysis

	Variants calling software	Total # variants	Filters	Final # variants
Case#2	Samtools	2268	Qual score, SegDup, allele frequency >1%, conservation	1
	Pindel	1634	Qual score <500, allele frequency >1%, SegDup, conservation	0
Case#3	Samtools	2777	Qual score, SegDup, allele frequency >1%, conservation	0
	Pindel	1683	Qual score <500, allele frequency >1%, SegDup, conservation	0

Tab.4 22q11.2 genome result

Func	Gene	Conserved	Chr	Start	End	Ref	Obs	Otherinfo	Qual score
ncRNA_exonic	SEPT5-GP1BB	340;Name=lod=32	chr22	19710889	19710889	G	A	hom	222

Discussion

By array CGH we have identified a 22q11 deletion in three patients with an atypical phenotype. We defined the phenotype atypical because it was in part coincident with the clinical features reported for the 22q11 deletion, but each patient had additional physical characteristics not reported in the deletion syndrome. From literature we know that the phenotype associated with the 22q11.2 microdeletions is highly variable but to date, no consistent correlations have been detected between deletion extent and phenotype. In addition, intrafamilial variability, even in monozygotic twins, has been found. This suggests that other factors might be involved in the expression of these malformations, including genetic and environmental factors (Uliana 2007). Because our patients showed a classical 22q11 deletion we hypothesised that one of the genes located in the 22q11 region can be mutated on the non deleted allele and that this gene can be responsible for the additional clinical features or that this gene was not directly responsible for the phenotype but altered the expression of a second gene.

In Case#3 we found no obvious likely pathogenic mutation, but the analysis is still ongoing. Case#3 was suspected for Cayler syndrome. The main characteristic of this syndrome is the asymmetric crying facies, a minor congenital anomaly seen in 0.5-1% of newborns. It is caused by either agenesis or hypoplasia of the depressor anguli oris muscle. This unilateral facial weakness is first noticed when the infant cries or smiles, affecting only one corner of the mouth (Garzena 2000). However the Cayler syndrome belongs to the group of conditions linked by microdeletion in the long arm of chromosome 22 (Giannotti 1994).

In Case#1 we found a mutation in the *MED15* gene. *MED15* is part of the Mediator complex (Blazek et al 2005). This complex is involved in the regulated transcription

of nearly all RNA polymerase II-dependent genes. It serves as a scaffold for the assembly of a functional preinitiation complex with RNA polymerase II and the general transcription factors. The Mediator is characterized by the presence of 4 module termed head, body, leg and kinase. The head module is essential for Mediator function, as mutations within it disrupt RNA polymerase II binding (Ranish et al. 1999). The body complex confers structural integrity to the Mediator, while the leg or tail region of Mediator seems involved in both activation and repression of transcription. The kinase module is an additional subcomplex reversible associated with the Mediator and has implicated in transcription repression. The *MED15* is part of the leg module.

In a not really recent study (Berti 2001), the authors demonstrated that *MED15* was expressed during embryogenesis with a high level in the frontonasal mass, pharyngeal arches and limb bud. They suggest a role in the regulation of developmental pathways underlying the morphogenesis of the derivative organs. Because our patients showed polydactyly in both hands and both feet surgically treated, the expression of *MED15* in limb bud and its regulation function can be involved in this clinical manifestation. Moreover, Kato et al. isolated the *Xenopus* homologue of *MED15* and demonstrated that was widely expressed during embryogenesis with high level in neuronal tissue (Kato 2002).

In Case#2 we found a mutation in a non-coding RNA. It represents a natural occurring read-through transcription between septin 5 (*SEPT5*) and glycoprotein Ib beta polypeptide (*GP1BB*). It arises from inefficient use of an imperfect polyA signal in the *SEPT5* gene and is candidate for nonsense-mediated mRNA decay (NMD). Septins constitute a family of GTP-binding proteins implicated in a variety of cellular processes from cell polarity to cytokinesis (Kinoshita 2001). In addition, septins seems to regulate exocytosis in post-mitotic cells such as neurons and platelets (Roeseler 2009). *SEPT5* deficiency seems to exert pleiotropic effects on a selected set of affective behaviours and cognitive processes as shown for *SEPT5* knockout mice which demonstrated delayed acquisition of rewarded goal approach (Suzuki 2009). Moreover the septin *SEPT5/7/11* complex is critical for dendrite branching and dendritic-spine morphology. Mutations in the *GP1BB* gene, together with mutations in *GP1BA* and *GP1X*, seem instead involved in the Bernard-Soulier

syndrome (BSS) (Savoia 2010). *GP1BB*-deficient mouse model of BSS displays macrothrombocytopenia and a severe bleeding phenotype, but no neurological impairments. Our patient didn't show a severe developmental delay nor the BSS phenotype. Furthermore, from the encode data it appeared that the non-coding RNA *SEPT5-GP1BB* was an highly conserved element that can represent an enhancer and therefore can regulate the expression of a distal gene. To confirm this hypothesis we have planned a luciferase assay.

Because of the presence of uterus and renal anomalies and the presence of a C2-C3 fusion, Case#2 was suspected for MURCS (Mullerian Renal Cervico-thoracic Somite anomalies) association. The most common associated malformations of MURCS, involve the upper part of urinary tract (40% of patients) and the cervicothoracic spine (30-40% of patients) (Pittok 2005). The MURCS association may be attributed to alterations in blastema giving rise to the cervicothoracic somites and the pronephric ducts, the ultimate spatial relationships of which are already determined by the end of the fourth week of fetal development (Duncan 1979). From literature we know that the smallest common deleted region among the deletions overlapping 22q11.2 and associated with MURCS is the most frequent 3 Mb 22q11.2 deletion associated with DiGeorge syndrome (Morcel 2011). This strongly suggests that the MURCS association is an additional component of the 22q11.2 deletion phenotype. In order to assess a correlation between the non-coding RNA mutated in our case and MURCS association, we have collected 9 additional patients with MURCS in which test the presence of the mutation.

In conclusion we have reported three patients with a 22q11.2 deletion and an atypical phenotype and in two of them we have found a mutation. Both the probands can present a deletion of chromosome 22 and, independently from it, additional clinical characteristics. Alternatively, the atypical phenotype of these two patients can represent an additional feature owing to the 22q11 deletion. Investigation of the gene located within this interval will be important in the search for genotype-phenotype correlation in future studies in this cytogenetic syndrome. However it stilled the possibility that a gene responsible for the phenotype variability was located outside of the 22q11 region or on a different chromosome.

References

1. Bassett AS, Chow EW, Husted J, Weksberg R, Caluseriu O, Webb GD, Gatzoulis MA. Clinical features of 78 adults with 22q11 deletion syndrome. *Am J Med Genet Part A*. 2005; 138:307–313.
2. Bassett AS, Chow EW. Schizophrenia and 22q11.2 deletion syndrome. *Curr Psychiatry Rep*. 2008; 10:148–157.
3. Bassett AS, Costain G, Fung WLA, Russell KJ, Pierce L, Kapadia R, Carter RF, Chow EW, Forsythe PJ. Clinically detectable copy number variations in a Canadian catchment population of schizophrenia. *J Psychiatr Res*. 2010; 44:1005–1009.
4. Berti L, Mittler G, Isolation and characterization of a novel gene from the DiGeorge chromosomal region that encodes for a mediatorsubunit. *Genomics*. 2001 Jun 15;74(3):320-32.
5. Blazek E, Mittler G, Meisterernst M. The mediator of RNA polymerase II. *Chromosoma*. 2005 Mar;113(8):399-408. Epub 2005 Feb 3.
6. Carlson C, Sirotkin H, Pandita R, Goldberg R, McKie J, et al: Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am J Hum Genet* 61: 620–629 (1997).
7. Duncan PA, Shapiro LR, Stangel JJ, Klein RM, Addonizio JC: The MURCS association: Mullerian duct aplasia, renal aplasia, and cervicothoracic somite dysplasia. *J Pediatr* 1979, 95:399-402.
8. Edelmann L, Pandita RK, Morrow BE. Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am J Hum Genet* 1999;64(4):1076–1086.
9. Garzena E, Ventriglia A, Patenella GA, Simonitti A, Becchino L. Congenital malformations and asymmetric crying facies. *Acta Biomed Ateneo Parmese* 2000;71:507–509.
10. Giannotti A, Digilio M, Marino B, Mingarelli R, Dallapiccola B. Cayler cardiofacial syndrome and 22q11: Part of the CATCH22 phenotype. *Am J Med Genet* 1994;53:303–304.
11. Goodship J, Cross I, LiLing J, Wren C. A population study of chromosome 22q11 deletions in infancy. *Arch Dis Child*. 1998; 79:348–351.
12. Guyot L, Dubuc M, Pujol J, Dutour O, Philip N. Craniofacial anthropometric analysis in patients with 22q11 microdeletion. *Am J Med Genet*. 2001; 100:1–8.
13. Insel TR. Rethinking schizophrenia. *Nature*. 2010; 468:187–193.
14. Kato Y, Habas R, Katsuyama Y, Näär AM, He X. A component of the ARC/Mediator complex required for TGF beta/Nodal signalling. *Nature*. 2002 Aug 8;418(6898):641-6. Epub 2002 Jul 24.
15. Kinoshita M, Noda M. Roles of septins in the mammalian cytokinesis machinery. *Cell Struct Funct* 2001; 26: 667–670.
16. Kinoshita, A., Noda, M. and Kinoshita, M. (2000) Differential localization of septins in the mouse brain. *J. Comp. Neurol.*, 428, 223–239.
17. Kobrynski LJ, Sullivan KE. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. *Lancet*. 2007; 370:1443–1452.
18. Lindsay EA, Baldini A. Recovery from arterial growth delay reduces penetrance of cardiovascular defects in mice deleted for the DiGeorge syndrome region. *Hum Mol Genet*. 2001; 10:997–1002.

19. Meechan DW, Maynard TM, Tucker ES, Lamantia AS. Three phases of DiGeorge/22q11 deletion syndrome pathogenesis during brain development: patterning, proliferation, and mitochondrial functions of 22q11 genes. *Int J Dev Neurosci.* 2011 May;29(3):283-94.
20. Morcel K. et al. Utero-vaginal aplasia (Mayer-Rokitansky-Küster-Hauser syndrome) associated with deletions in known DiGeorge or DiGeorge-like loci. *Orphanet Journal of Rare Diseases* 2011, 6:9
21. Pittock ST, Babovic-Vuksanovic D, Lteif A: Mayer-Rokitansky-Kuster-Hauser anomaly and its associated malformations. *Am J Med Genet A* 2005, 135:314-316.
22. Ranish JA, Yudkovsky N, Hahn S (1999) Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev* 13:49–63
23. Roeseler S, Sandrock K, Bartsch I, et al. Septins, a novel group of GTP-binding proteins: relevance in hemostasis, neuropathology and oncogenesis. *Klin Padiatr* 2009; 221: 150–155.
24. Ryan AK, Goodship JA, Wilson DI, Philip N, Levy A, Seidel H, Schuffenhauer S, Oechsler H, Belohradsky B, Prieur M, Aurias A, Raymond FL, Clayton-Smith J, Hatchwell E, McKeown C, Beemer FA, Dallapiccola B, Novelli G, Hurst JA, Ignatius J, Green AJ, Winter RM, Brueton L, Brondum-Nielsen K, Scambler PJ, et al. Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study. *J Med Genet.* 1997; 34:798–804.
25. Savoia A, et al. Clinical and genetic aspects of Bernard-Soulier syndrome: searching for genotype/phenotype correlations. *Haematologica* 2011;96(3):417-423
26. Scambler PJ. The 22q11 deletion syndromes. *Hum Mol Genet.* 2000; 9:2421–2426.
27. Shaikh TH, Kurahashi H, Saitta SC, O’Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, et al. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* 2000;9 (4):489–501.
28. Suzuki G, et al. *SEPT5* deficiency exerts pleiotropic influence on affective behaviors and cognitive functions in mice. *Human Molecular Genetics*, 2009, Vol. 18, No. 9
29. Uliana V, Giordano N, Caselli R, Papa FT, Ariani F, Marcocci C, Gianetti E, Martini G, Papakostas P, Rollo F, et al: Expanding the phenotype of 22q11 deletion syndrome: the MURCS association. *Clin Dysmorphol* 2008, 17:13-17.

5. DISCUSSION

Genomic rearrangements describe mutational changes that alter genome structure (e.g., duplication, deletion, insertion, and inversion). These are different from the traditional mutation caused by Watson–Crick base pair alterations. Each of these rearrangements, excepting inversions, result in copy number variation (CNV) or change from the usual copy number of two for a given genomic segment or genetic locus of our diploid genome. Genomic rearrangements can represent polymorphisms that are neutral in function, or may produce abnormal phenotypes. The pathological conditions caused by genomic rearrangements are collectively defined as genomic disorders (Lupski 1998 and 2009). Due to the limited resolution of conventional cytogenetic techniques, the majority of genomic disorders were missed in the past, because the genomic rearrangements were not cytogenetically visible. However, high-resolution array comparative genomic hybridization (aCGH) techniques have revolutionized the approach to diagnosis of genomic disorders, and enabled the screen of the entire human genome for CNVs. Therefore a growing number of submicroscopical deletions and duplications causing complex neurodevelopmental disorders have been identified and recently the reciprocal duplication syndromes have been reported for almost all microdeletion syndromes. Many of the known microdeletion syndromes and their corresponding microduplication syndromes occur on the basis of non-allelic homologous recombination in low copy repeats.

Duplications or deletions of regions on chromosome 17 have been implicated in a number of genomic disorders in humans (Lupski and Stankiewicz, 2005). Chromosome 17 has the second highest gene content amongst all chromosomes. It harbors several dosage-sensitive genes, including *PMP22*, *PAFAH1B1*, *YWHAE*, *RAI1*, and *NFI*, which have been implicated in a number of genomic disorders (Lupski, 2009). Genomic studies have elucidated the mechanisms underlying genomic rearrangements in chromosome 17 and their contribution to the clinical phenotypes. Based on NAHR mechanism, a CNV generation is the prediction that a deletion can have a reciprocal duplication. Hence a genomic disorder caused by deletion could, in theory, also have a corresponding duplication-associated disorder. However, intrachromatid NAHR can only result in deletion and so the frequency of

deletion versus duplication is not equal, with a higher deletion frequency. Existing knowledge supports the notion that the deletion phenotype is anticipated to be more severe than the duplication phenotype. Decreased expression resulting from a gene deletion causes a phenotype usually similar to that observed with loss-of-function point mutations of a “dosage-sensitive” gene. Increased expression, resulting from gene duplication may convey clinical findings that are different, and sometimes divergent from the deletion phenotype (Bi 2009).

We reported two cases with a duplication of the Miller-Dieker region. Both cases are the unbalanced result of two different balanced translocations: $t(9;17)(p24.2;p13.3)$ and $t(10;17)(10q26.2;p13.3)$; and therefore their phenotypes are more complex than the phenotype of cases with isolated 17p13.3 duplications (result 4.1). Previous studies highlighted that transgenic mice over-expressing *PAFAH1B1* showed migration defect and reduced brain volume (Bi 2009). The last sign is also present in humans since most patients showed microcephaly. Therefore, our data confirm that *PAFAH1B1* over-expression in humans does not cause neuronal migration defects or other gross brain malformations. Comparing our cases with those previously described in literature, it appeared that they share some facial and physical features such as pointed chin, triangular face high nasal bridge and a deceleration of head growth. Interestingly, recurrent respiratory infections during childhood were reported in all patients. Since the 17p duplications of our patients harboured from two unbalanced translocations, the phenotype is more complex resulting by the combined effects of the duplication of the 17p13.3 region and of the 9p and 10q deletions. Nevertheless we contribute to further delineate the features associated with this novel microduplication syndrome.

Autism spectrum disorders (ASDs), typically apparent by the age of 3 years, encompass a broad range of developmental disorders that are marked by limitations in one of three behavioural/developmental domains: social interaction; language, communication, and imaginative play; and range of interest and activities (Muhle 2004). The ASDs range from phenotypically mild to severe and include autism, atypical autism, Asperger syndrome, and pervasive developmental disorders. The heritability of autism may as high as 90%, making it one of the most heritable

complex disorders. About 10% of cases are associated with a Mendelian syndrome (e.g. fragile X syndrome and tuberous sclerosis complex). There are two hypothesis for the genetic aetiology of autism. The first theory, referred as the “common gene/common disease” hypothesis, is that common diseases result from the additive or multiplicative effects of genetic and environmental factors. Common genetic variants confer only a small increased risk to a given individual, but because of the high frequency with which these variants are found, each has a large attributable risk among the population (Weiss 2009). An alternative to the “common gene/common disease” hypothesis is that ASDs are caused not only by common variants of small effect but also by rare highly penetrant variants such as chromosomal deletions and duplications (Kusenda 2008). A substantial proportion of idiopathic autism may be attributable to CNVs. Two recent studies detected *de novo* CNVs in 7–10% of autistic cases from simplex families, 2–3% of cases from multiplex families, and in 1% of controls (Marshall 2008). These results not only implicate CNVs in the aetiology of autism but also indicate that different genetic mechanisms may underlie sporadic, versus familial, autism. Microdeletions and microduplications of chromosome 16p11.2 have been found at varying frequencies among individuals diagnosed with ASDs. Microdeletions are a more common cause of ASDs than the reciprocal microduplication (0.50% vs. 28%, respectively) (Walsh 2011). Microduplications seem instead strongly associated with schizophrenia (McCarthy 2009). Furthermore Walters et al; demonstrated that a 16p11.2 deletion give rise to a strongly-expressed obesity phenotype. Possible explanations include a direct causal relationship between obesity and developmental delay; the involvement of the same or related regulatory pathways; or different outcomes of the same set of behavioural disorders with complex pleiotropic effects and variable ages of onset and expressivities (Walters 2010).

To test whether gene dosage accounting for obesity in carriers of the 16p11.2 deletion may also influence BMI in a converse manner, we assembled and phenotypically analysed cohorts of duplications carriers (result 4.2). The duplication was strongly associated with lower weight and lower BMI. Adults carrying the duplication had a relative risk of being clinically underweight of 8.3. The duplication was also associated with reduced head circumference, 26.7% presenting with

microcephaly, whereas carriers of the reciprocal deletion had an increased head circumference. This suggests that head circumference and BMI may be regulated by a common pathway, or that a causal relationship exists between these two traits in these patients. To evaluate if the phenotypes observed in 16p11.2 deletion and duplication individuals may be due to effects on the expression of genes mapping within or near the rearranged region, we performed an expression assay in lymphoblastoid cell-lines. Expression levels correlated positively with gene dosage for all genes within the CNV, while genes proximal to the rearrangement showed no significant variations. Therefore as in the schizophrenia/autism and microcephaly/macrocephaly dualisms, overweight/underweight could represent opposite pathological manifestations of a common energy-balance mechanism.

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. 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 (Lupski and Stankiewicz 2005). 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.

Stankiewicz et al. recently reported 24 cases with deletions and 17 cases with duplications at 10q11.21q21.1. The only clinical features common to a majority of individuals were ID and DD. Other clinical features identified include failure to thrive, growth retardation, autism spectrum disorders, microcephaly, attention deficit hyperactivity disorder (ADHD). However, a CNV-phenotype association has not been made for the 10q11.22 region and this CNV has not been classified as pathogenic.

We recently reported that a small duplication on 10q11.22 including *GPRIN2* gene, a regulator of neurite outgrowth, and *PPYR1*, a gene involved in energy

homeostasis, is a candidate modifier for Rett syndrome (Artuso 2011). Specifically, duplications were found in the Zappella variant, the Rett variant with recovery of speech, and lacking the typical growth delay, underweighting and autistic features. Since *PPYRI* knockout mice display underweight and reduced white adipose tissue an over-expression of *PPYRI* due to gene duplication may be responsible for the higher body weight characterizing Zappella variant. We concluded that duplication at 10q11.22 may play a role in protecting from both underweighting and autistic features in Rett patients (Artuso 2011). We now report more convincing evidences that this microdeletion is probably clinically relevant and responsible for the ASD phenotype, because significantly enriched in the ASD population when compared to the SID/NSID population ($p=0,001$) (result 4.3). The duplication was observed also in several controls, suggesting that the duplication by itself confers either no phenotype at all or a range of phenotypes of varying severity. Moreover, because genetic variation studies have reinforced the potential influence of *PPYRI* on body weight in humans (Sha 2009), we also demonstrated an increasing BMI value in cases carrying the duplication. The highlighted examples demonstrate how gene dosage effects may influence the development of common disorders often characterized by heterogeneous genetic aetiology.

Other molecular mechanisms by which rearrangements of the genome may convey or alter a disease phenotype result from how the rearrangement on one chromosome affects or is affected by the allele on the other chromosome at that locus. These include the unmasking of either recessive mutations or functional polymorphisms of the remaining allele when a deletion occurs, and potential transvection effects via deletion of regulatory elements required for communication between alleles (Lupski and Stankiewicz 2005). Recessive genes reside within the CNV regions, and the chances of finding a recessive mutation along with a microdeletion are rare (frequency of spontaneous mutation x frequency of the deletion event), but plausible. Profound sensorineural hearing loss has been reported in patients with Smith-Magenis syndrome whose deletions unmask the recessive mutation in the myosin (*MYO15A*) gene located within the 17p11.2 region (Liburd 2001). Functional polymorphisms within *COMT* and *FXII*, unmasked by hemizygous

deletions, have also been reported to result in cognitive decline and psychosis in patients with 22q11.2 deletion and reduced activity of coagulation factor 12 in Sotos syndrome respectively (Gothelf 2005, Kurotaki 2005). Additional example can be represented by the thrombocytopenia absent radius (TAR) syndrome in which one copy of the *RBM8A* gene is not functional, due to a null allele, and the expression of the other copy is reduced, as a result of noncoding SNPs in the 5' UTR or first intron (Albers 2012).

We reported here our experience with three patient showing a 22q11.2 deletion and an atypical phenotype. In order to identify possible recessive alleles we performed targeted sequencing of the 22q11.2 region. In one case we identified a mutation in the *MED15* gene, that is part of the Mediator complex (Blazek 2005). This gene is highly expressed during embryogenesis with high levels in limb bud and neuronal tissue (Berti 2001, Kato 2002). Therefore we hypothesize an involvement of this gene in the polydactyly and severe intellectual deficit showed by our patient. In a second case we identified a mutation in a non-coding RNA. Previous data (Pennacchio 2006) revealed the high propensity of extremely conserved human non-coding sequences to behave as transcriptional enhancer in vivo, and supported both ancient human-fish conservation and human-rodent ultraconservation as highly effective filters to identify such functional elements. From the encode data it appeared that the non-coding RNA *SEPT5-GPIBB* was highly conserved from human to elephant. Therefore, it can represent an enhancer involved in the regulation expression of a distal gene. To confirm this hypothesis we have planned a luciferase assay. In the last case we found no obvious likely pathogenic mutation, but the analysis is still ongoing. In conclusion we demonstrated that targeted sequencing of genes within the pathogenic CNV region using the newly available technologies would be useful to find potential candidate genes.

***6. CONCLUSIONS and FUTURE
PERSPECTIVES***

The conventional wisdom surrounding genomic disorders posits that they fit several criteria: the deletions/duplications are large, highly penetrant, de novo in the majority of individuals, and associated with a uniform constellation of clinical features (Mefford and Eichler, 2009). Smith-Magenis syndrome, Prader-Willi syndrome, and Williams-Beuren syndrome are examples of such “classic” genomic disorders. In contrast to these “classic” genomic disorders, many of the more recently described recurrent genomic lesions identified in large case–control studies demonstrate apparently diverse phenotypes and are frequently inherited while showing reduced penetrance (Klopocki et al., 2007; Mefford et al., 2008; Sharp et al., 2008).

Several explanations have been proposed for the variable expressivity and clinical heterogeneity in some genomic disorders. First, atypical or variable-sized copy number changes may account for the variable phenotypes in some apparently recurrent lesions. A “two-hit” model has also recently been proposed to account for phenotypic variability. One hit may be sufficient to reach a threshold that results in mild neurodevelopmental deficits, whereas a second hit is necessary for the development of a more severe neurological phenotype. Alternatively, the abnormal phenotype in patients with a heterozygous deletion can result from unmasking of a recessive mutation or functional polymorphism of the remaining allele.

It is not clear to what extent such genomic changes are responsible for Mendelian or complex disease traits and common traits, or represent only benign polymorphic variation. Furthermore, some phenotypes caused by genomic rearrangements may not present until late adulthood. This age-dependent penetrance confounds the interpretation of genomic copy-number changes.

We know that rearrangements occur throughout the genome, and therefore it is plausible to assume that such rearrangements or CNVs could be associated with inherited or sporadic disease, susceptibility to disease, complex traits, or common benign traits, or could represent polymorphic variation with no apparent phenotypic consequences, depending on whether or not dosage-sensitive genes are affected by the rearrangement. As demonstrated by this study, some genomic disorders show highly variable penetrance that can make difficult the interpretation of molecular results. The effective identification of such regions will likely require collaborative

efforts by multiple centres, in order to collect a sufficient numbers of patients carrying the same structural variant. A cohort of multiple individuals with a particular pathogenic variant will likely show at least some degree of phenotypic concordance even where penetrance is incomplete, making possible a more defined genotype-phenotype correlation.

For the future we plan to continue the consultation of the literature and the re-evaluation of our cohort paying attention to the CNV regions to find new emerging low penetrance syndromes. We also plan to use Next-Generation Sequencing of selected regions or candidate genes to identified new recessive phenotype.

7. REFERENCES

References

1. Albers CA, et al. Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. *Nat Genet.* 2012 Feb 26;44(4):435-9, S1-2.
2. 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.
3. Artuso R, et al. Investigation of modifier genes within copy number variations in Rett syndrome. *J Hum Genet.* 2011 Jul;56(7):508-15.
4. Bassett AS & Chow EW. Schizophrenia and 22q11.2 deletion syndrome. *Curr Psychiatry Rep.* 2008; 10:148–157.
5. Bassett AS, et al. Clinical features of 78 adults with 22q11 deletion syndrome. *Am J Med Genet Part A.* 2005; 138:307–313.
6. Bassett AS, et al. Clinically detectable copy number variations in a Canadian catchment population of schizophrenia. *J Psychiatr Res.* 2010; 44:1005–1009.
7. Bejjani, B.A. & L.G. Shaffer, Clinical utility of contemporary molecular cytogenetics. *Annu Rev Genomics Hum Genet*, 2008. 9: p. 71-86.
8. Bejjani, B.A. & L.G. Shaffer. 2006. Application of array-based comparative genomic hybridization to clinical diagnostics. *J. Mol. Diagn.* 8: 528–533.
9. Bejjani, B.A. et al. 2005. Use of targeted array-based CGH for the clinical diagnosis of chromosomal imbalance: Is less more? *Am. J. Med. Genet. A* 134: 259–267.
10. Berti L, et al. Isolation and characterization of a novel gene from the DiGeorge chromosomal region that encodes for a mediator subunit. *Genomics.* 2001 Jun 15;74(3):320-32.
11. Bi W, et al. Increased LIS1 expression affects human and mouse brain development. *Nat Genet* 2009;41:168–77.
12. Blazek E, Mittler G, Meisterernst M. The mediator of RNA polymerase II. *Chromosoma.* 2005 Mar;113(8):399-408.
13. Bruno D.L. et al. Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *J Med Genet* 2010;47:299e311.
14. Canales CP & Walz K. Copy number variation and susceptibility to complex traits. *EMBO Mol Med.* 2011 Jan;3(1):1-4.
15. Cardoso C, et al. Refinement of a 400-kb critical region allows genotypic differentiation between isolated lissencephaly, Miller-Dieker syndrome, and other phenotypes secondary to deletions of 17p13.3. *Am J Hum Genet* 2003;72:918–30.
16. Carlson C, et al: Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am J Hum Genet* 61: 620–629 (1997).

17. Cheung S.W. et al. (2005) Development and validation of a CGH microarray for clinical cytogenetic diagnosis. *Genet. Med.* 7, 422–432
18. Colnaghi R, Carpenter G, Volker M, O'Driscoll M. The consequences of structural genomic alterations in humans: genomic disorders, genomic instability and cancer. *Semin Cell Dev Biol.* 2011 Oct;22(8):875-85. Review.
19. Conrad DF, et al. Origins and functional impact of copy number variation in the human genome. *Nature.* 2010 Apr 1;464(7289):704-12.
20. Dobyns WB, Reiner O, Carrozzo R, Ledbetter DH. Lissencephaly. A human brain malformation associated with deletion of the LIS1 gene located at chromosome 17p13. *JAMA* 1993;270:2838–42.
21. Feuk, L., Carson A.R. & Scherer S.W. 2006. Structural variation in the human genome. *Nat. Rev. Genet.* 7: 85–97.
22. Garcia-Minaur S, et al. A novel atypical 22q11.2 distal deletion in father and son. *J Med Genet* 2002;39(10):E62.
23. Girirajan S. & Eichler E.E., Phenotypic variability and genetic susceptibility to genomic disorders. *Human Molecular Genetics*, 2010, Vol. 19, Review.
24. Goodship J, Cross I, LiLing J, Wren C. A population study of chromosome 22q11 deletions in infancy. *Arch Dis Child.* 1998; 79:348–351.
25. Gothelf, D. et al. (2005) COMT genotype predicts longitudinal cognitive decline and psychosis in 22q11.2 deletion syndrome. *Nat. Neurosci.*, 8, 1500–1502.
26. Gu W, Zhang F, Lupski JR. Mechanisms for human genomic rearrangements. *Pathogenetics.* 2008 Nov 3;1(1):4.
27. 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.
28. Iafrate A.J. et al. 2004. Detection of large-scale variation in the human genome. *Nat. Genet.* 36: 949–951.
29. Insel TR. Rethinking schizophrenia. *Nature.* 2010; 468:187–193.
30. Ishkanian A.S. et al. 2004. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat. Genet.* 36: 299–303.
31. Kallioniemi A, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science.* 1992 Oct 30;258(5083):818-21.
32. Kato Y, et al. A component of the ARC/Mediator complex required for TGF beta/Nodal signalling. *Nature.* 2002 Aug 8;418(6898):641-6.
33. Kirchhoff M, et al. Deletions below 10 megabasepairs are detected in comparative genomic hybridization by standard reference intervals. *Genes Chromosomes Cancer.* 1999 Aug;25(4):410-3.
34. Klopocki E, et al. 2007. Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet* 80:232–240.

35. Kobrynski LJ & Sullivan KE. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. *Lancet*. 2007; 370:1443–1452.
36. 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.
37. Kumar RA, et al. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;17:628e38.
38. Kurotaki N., et al. (2005) Phenotypic consequences of genetic variation at hemizygous alleles: Sotos syndrome is a contiguous gene syndrome incorporating coagulation factor twelve (FXII) deficiency. *Genet. Med.*, 7, 479–483.
39. Kusenda M & Sebat J. The role of rare structural variants in the genetics of autism spectrum disorders. *Cytogenet Genome Res* 2008;123:36–43.
40. Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*. 2007 Dec 28;131(7):1235-47.
41. Lee, C., Iafrate A.J., and Brothman A.R., Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nat Genet*, 2007. 39(7 Suppl): p. S48-54.
42. Liburd N., et al. (2001) Novel mutations of MYO15A associated with profound deafness in consanguineous families and moderately severe hearing loss in a patient with Smith–Magenis syndrome. *Hum. Genet.*, 109, 535–541.
43. Lu, X. et al. (2007) Clinical implementation of chromosomal microarray analysis: summary of 2513 postnatal cases. *PLoS ONE* 2, e327
44. Lupski JR & Stankiewicz P. 2005. Genomic disorders: Molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet* 1:e49.
45. Lupski JR. 1998. Genomic disorders: Structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 14:417.
46. Lupski JR. 2009. Genomic disorders ten years on. *Genome Med* 1:42.
47. Lupski JR. Genomic rearrangements and sporadic disease. *Nat Genet* 2007;39: S43e47.
48. Marshall CR, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82:477– 488.
49. McCarroll, S.A. & D.M. Altshuler. 2007. Copy-number variation and association studies of human disease. *Nat. Genet.* 39: S37–42.
50. McCarthy SE, et al. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet* 2009;41:1223–1227.
51. McDonald-McGinn DM, et al. The Philadelphia story: the 22q11.2 deletion: report on 250 patients. *Genet Couns.* 1999;10(1):11-24.

52. Meechan DW, Maynard TM, Tucker ES, Lamantia AS. Three phases of DiGeorge/22q11 deletion syndrome pathogenesis during brain development: patterning, proliferation, and mitochondrial functions of 22q11 genes. *Int J Dev Neurosci*. 2011 May;29(3):283-94.
53. Mefford HC & Eichler EE. 2009. Duplication hotspots, rare genomic disorders, and common disease. *Curr Opin Genet Dev* 19:196–204.
54. Mefford HC, et al. 2008. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med* 359:1685–1699.
55. Muhle R, Trentacoste SV, Rapin I. The genetics of autism. *Pediatrics* 2004;113:e472– e486.
56. 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.
57. O'Donnell H, et al. Detection of an atypical 22q11 deletion that has no overlap with the DiGeorge syndrome critical region. *Am J Hum Genet* 1997;60:1544–1548.
58. Pennacchio LA, et al. In vivo enhancer analysis of human conserved non-coding sequences. *Nature*. 2006 Nov 23;444(7118):499-502.
59. Perry GH, et al. Diet and the evolution of human amylase gene copy number variation. *Nat Genet*. 2007 Oct;39(10):1256-60.
60. Pinkel D. et al. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet*. 20: 207–211.
61. Potocki L, et al. Molecular mechanism for duplication 17p11.2-the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 2000;24:84–7.
62. Rauch A, et al. A novel 22q11.2 microdeletion in DiGeorge syndrome. *Am J Hum Genet* 1999;64:659–666.
63. Redon R. et al. 2006. Global variation in copy number in the human genome. *Nature* 444: 444–454.
64. Roos L, et al. A new microduplication syndrome encompassing the region of the Miller-Dieker (17p13 deletion) syndrome. *J Med Genet* 2009;46:703e10.
65. Ryan AK, et al. Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study. *J Med Genet*. 1997; 34:798–804.
66. Saitta SC, et al. Aberrant interchromosomal exchanges are the predominant cause of the 22q11.2 deletion. *Hum Mol Genet* 2004;13(4):417–428.
67. Sambrook, J. & M.J. Gething, Protein structure. Chaperones, paperones. *Nature*, 1989. 342(6247): p. 224-5.
68. Sebat J. et al. 2004. Large-scale copy number polymorphism in the human genome. *Science* 305: 525–528.

69. Sha BY, 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 Apr;54(4):199-202.
70. Shaffer L.G. et al. 2006. Targeted genomic microarray analysis for identification of chromosome abnormalities in 1500 consecutive clinical cases. *J. Pediatr.*149: 98–102.
71. Shaikh TH, et al. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet.* 2000; 9:489–501.
72. Shaikh, T.H. 2007. Oligonucleotide arrays for highresolution analysis of copy number alteration in mental retardation/multiple congenital anomalies. *Genet. Med.* 9: 617–625.
73. Sharp AJ, et al 2008. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* 40:322–328.
74. Sharp, A.J. et al. 2005. Segmental duplications and copy-number variation in the human genome. *Am. J. Hum. Genet.* 77: 78–88.
75. Shinawi M & Cheung SW. The array CGH and its clinical applications. *Drug Discov Today.* 2008 Sep;13(17-18):760-70.
76. Stankiewicz P, et al. Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am J Hum Genet* 2003;72:1101–16.
77. Stankiewicz P, et al. Recurrent deletions and reciprocal duplications of 10q11.21q11.23 including CHAT and SLC18A3 are likely mediated by complex low-copy repeats. *Hum Mutat.* 2012 Jan;33(1):165-79.
78. Trask BJ: Fluorescence in situ hybridization: applications in cytogenetics and gene mapping. *Trends Genet* 1991; 7: 149-154.
79. Tyson C, et al. Submicroscopic deletions and duplications in individuals with intellectual disability detected by array-CGH. *Am J Med Genet A.* 2005 Dec 15;139(3):173-85.
80. van Bon BW, et al: 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: 511-523.
81. 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.
82. Vissers LE, et al. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet.* 2004 Sep;36(9):955-7.
83. Walsh KM & Bracken MB. Copy number variation in the dosage-sensitive 16p11.2 interval accounts for only a small proportion of autism incidence: a systematic review and meta-analysis. *Genet Med.* 2011 May;13(5):377-84.
84. Walters R. G. et al. A novel highly-penetrant form of obesity due to microdeletions on chromosome 16p11.2. *Nature.* 2010 February 4; 463(7281): 671–675

85. Weiss LA, et al. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009;461:802– 808.
86. Weiss LA, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;358:667e75.
87. Ylstra B. et al. 2006. BAC to the future! Or oligonucleotides: A perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res.* 34: 445–450.
88. Yunis J: High resolution of human chromosomes. *Science* 1976; 191: 1268-1270.
89. Zhang F., Carvalho C.M.B., Lupski J.R. Complex human chromosomal and genomic rearrangements. *Trends Genet*, 25 (2009), pp. 298–307

CURRICULUM VITAE

PERSONAL DATA

Surname(s) / First name(s)	Mucciolo Mafalda
Address	Via Principi Carafa 217, 84949 Castel San Lorenzo, SA
Telephone	Mobile: 3296182430 - 3487787463
E-mail	mucciolo8@unisi.it / mucciolodina@libero.it
Nationality	Italian
Date of birth	27/03/1984
Present position	PhD Student Medical Genetics Department of Molecular Biology University of Siena

WORK EXPERIENCE

November 2008-today	<u>Doctoral School</u> Scholarship for a 4 years Doctoral in Oncology and Genetics Medical Genetics, Department of Molecular Biology, University of Siena
September 2007	<u>Volunteer training</u> One month of unpaid training for Clinical analysis laboratory, "Ospedale di Roccaspide", Salerno

STUDIES

August 2009	<u>Professional Qualification for Biologist</u>
September 2006-September 2008	<u>Master Degree in Molecular Biology.</u> Achieved on September 26, 2008 with marks 110/110 Lode. Degree thesis about "Mutation analysis of RB1 gene in patients with retinoblastoma: genotype-phenotype correlation"
September 2003-November 2006	<u>Bachelor's Degree in Biological Science.</u> Achieved on November 13, 2006 with marks 107/110. Degree thesis about "Molecular mechanisms of lymphangiogenesis in pathological condition"
September 1998-July 2003	<u>School-leaving Certificate</u> Scientific High School "Parmenide", Roccadaspide, SA

RESEARCH ACTIVITY

November 2008-today	<p>Identification of the genetic causes in patients with intellectual disability, autism and multiple congenital anomalies by array CGH analysis (Agilent platform) in order to correlate new copy number variations (CNVs) with a specific phenotype. Data analysis was performed using the main bioinformatics databases (UCSC Genome Browser, Ensemble, GeneCards, Database of Genomic Variant, etc.). Patients data entry was added in the Decipher database (http://decipher.sanger.ac.uk/) and in the Italian Database of Human CNVs (http://gvarianti.homelinux.net/gvarianti/index.php).</p> <p>Array-CGH analysis in two discordant pairs of Rett sisters and four additional discordant pairs of unrelated Rett girls matched by mutation type in order to identify phenotypic modifier genes/regions (J Hum Genet. 2011 Jul).</p> <p>Array-CGH analysis on DNA extracted from patients with Alport syndrome and leiomyomatosis carrying a deletion extending beyond COL4A6 intron 2</p> <p>Identification of copy number variations (six deletions and one duplication) at the 16p11.2 region in our cohort of 632 patient with intellectual disability, multiple congenital anomalies and autism. This data are part of a collaborative study about gene dosage at the chromosome 16p11.2 locus supervised by Medical Genetic Unit of Lousanne (Switzerland) (Nature. 2011 Aug 31).</p> <p>Array-CGH analysis in 12 cases of prenatal diagnosis in which were characterized fetal chromosome</p>
----------------------------	---

	rearrangements, fetal chromosome marker or increased nuchal translucency (>4 mm) with normal karyotype. Specifically this technique was conducted on DNA extracted from amniotic fluid or villi fragments and amplified by whole genome amplification (WGA).
January 2008- October 2008	Mutational screening of RB1 gene using DHPLC, sequencing and MLPA in order to establish the relationship between the type of RB1 gene mutation and the phenotype of Retinoblastoma affected patients.

LABORATORY EXPERIENCES

2012	Three months attendance of the “Département de Médecine Génétique et Développement”, University of Geneva, Switzerland Learned techniques: sequence capture, next generation sequencing and data analysis using Samtools, Pindel and Annovar databases.
2008	Attendance of the “Medical Genetics Laboratory”, University of Siena, from September 2008 till today. Learned techniques: DNA extraction, RNA extraction, Whole Genome Amplification, PCR, RT-PCR, DNA fragments separation on agarose and acrylamide gel, automated sequencing, DHPLC (Denaturing High performance Liquid Chromatography), MLPA (Multiplex Ligation-dependent Probe Amplification), array-CGH (Comparative Genomic Hybridization).
2006	Three months attendance of the “Department of Neuroscience”, University of Siena. Learned techniques: optical microscopy, histochemistry and immunohistochemistry.

DIAGNOSTIC ACTIVITY

2008-today	Intellectual Disability: array CGH analysis in patients affected by intellectual deficit, autism and multiple congenital anomalies. Retinoblastoma: <i>RB1</i> screening in patients affected by
------------	---

	<p>sporadic and familial retinoblastoma.</p> <p>Retinoschisis: <i>XLRS</i> screening in patients affected by sporadic and familial retinoschisis.</p> <p>Medullary thyroid carcinoma: <i>pRET</i> screening in patients affected by sporadic and familial thyroid carcinoma.</p>
--	--

PERSONAL SKILLS AND COMPETENCES

Mother tongue(s)		Italian								
Other language(s)										
Self-assessment	Understanding				Speaking				Writing	
	Listening		Reading		Spoken interaction		Spoken production			
European level (*)										
English	B2	Independent user	B2	Independent user	B2	Independent user	B2	Independent user	B2	Independent user
French	B1	Independent user	B1	Independent user	B1	Independent user	B1	Independent user	B1	Independent user

(*)Common European Framework of Reference (CEF) level

COMPUTER SKILLS AND COMPETENCES

Microsoft Office	good command of Microsoft Office™ tools (Word™, Excel™ and PowerPoint™)
	basic knowledge of graphic design applications (PhotoShop™)

ADDITIONAL INFORMATION

Participation in congress	XIV Congresso Nazionale della Società Italiana di Genetica Umana (SIGU), November 13-16, 2011-Milano
	7 th International Decipher Symposium, Wellcome Trust Genome Campus, May 23-25, 2011 – Hinxton (UK)

	6 th International Decipher Symposium, Wellcome Trust Genome Campus, May 19-21, 2010 – Hinxton (UK)
	XII Congresso Nazionale della Società Italiana di Genetica Umana (SIGU), November 8-10, 2009 – Torino
Author of oral communications at National and International Congress	6 th International Decipher Symposium, Wellcome Trust Genome Campus, May 19-21, 2010 – Hinxton (UK) Copy Number Variations in Autism Spectrum Disorders Mucciolo M , Mari F, Canitano R, Mencarelli MA, Papa FT, Katzaki E, Marozza A, Radice L, Castagnini C, Dosa L, Pollazzon M, Hayek J, Renieri A.
Co-author of oral communications at National and International Congress	XIV Congresso Nazionale della Società Italiana di Genetica Umana (SIGU), November 13-16, 2011-Milano Fenotipi opposti di BMI si associano al dosaggio genico della regione 16p11.2 Disciglio V., Mucciolo M. , Mencarelli M.A., Castagnini C., Pollazzon M., Marozza A., Mari F., Renieri A.
	6 th International meeting on cryptic chromosomal rearrangements and genes in mental retardation and autism, Troina, April 22-24, 2010 Array-CGH analysis in Autism Spectrum Disorders Mari F, Mucciolo M , Canitano R, Mencarelli MA, Papa FT, Katzaki E, Marozza A, Radice L, Castagnini C, Dosa L, Pollazzon M, Hayek J, Renieri A.
	5 th International Meeting on Cryptic Chromosomal Rearrangements and Genes in Mental Retardation and Autism, April 17-18, 2009 - Troina (IT), Array-CGH analysis of two hundred patients with parents investigation and analysis of surrounding genes. Papa FT, Katzaki E, Disciglio V, Mucciolo M , Mencarelli MA, Uliana V, Pollazzon M, Marozza A, Bruccheri MG, Mari F, Renieri A
	5 th International Decipher Symposium, Wellcome Trust Genome Campus, May 20-22, 2009 – Hinxton (UK) Array-CGH analysis in a cohort of 200 patients: the importance of parents investigation and analysis of surrounding genes. Katzaki E, Papa FT, Disciglio V, Mucciolo M , Mencarelli MA, Uliana V, Pollazzon M, Marozza A, Bruccheri MG, Mari F, Renieri A.
	XII Congresso Nazionale SIGU, November 8-11, 2009 –

	<p><u>Torino (IT)</u> Sindrome da microdelezione 21q22.11q22.12: descrizione di tre pazienti. Katzaki E, Morin G, Pollazzon M, Papa FT, Mucciolo M, Buoni S, Hayek G, Andrieux J, Lecerf L, Popovici C, Receveur A, Mathieu-Dramard M, Mari F, Philip N, Renieri A</p>
<p>Poster presentations at National and International Congress</p>	<p><u>European Human Genetics Conference 2012: Nürnberg, Germany, June 23-26, 2012</u> Mirror effects for Autism Spectrum Disorder due to gene dosage at 10q11.22 affecting GPRIN2 gene, a regulator of neurite outgrowth and PPYR1 gene involved in energy homeostasis. V. Disciglio, M. Fichera, R. Ciccone, M. Mucciolo, E. Ndoni, A. Fernández Jaén, O. Galesi, M. Vinci, P. Failla, M.A. Mencarelli, C. Lo Rizzo, F. Mari, O. Zuffardi, C. Romano, A. Renieri.</p>
	<p><u>68° Congresso Nazionale della Società Italiana di Pediatria (SIP), May 9-11, 2012-Roma</u> Microcefalia, dismorfismi, ritardo neuromotorio e del linguaggio in una bambina con duplicazione della regione 19p13.3 Tuccio A., Boddi G., Nanni S., Iantorno L., Tataranno M. L., Marozza A., Mucciolo M., Renieri A., Bartalini G., Balestri P.</p>
	<p><u>7th International Meeting on CNVs and Genes in Intellectual Disability and Autism, Troina, April 13 and 14, 2012</u> EVIDENCE FOR A NEW GENE CONTIGUOUS SYNDROME AT 22q13, NOT INVOLVING SHANK3 GENE Disciglio V.1, Marozza A.1, Mucciolo M., Mencarelli M.A., Fichera M., Romano C., Anderlid B.M., Clayton-Smith J., Metcalfe K., David A., Le Caignec C., Tümer Z., Fryer A., Andrieux J., Novelli A., Pecile V., Mari F., Renieri A.</p>
	<p><u>XIV Congresso Nazionale della Società Italiana di Genetica Umana (SIGU), November 13-16, 2011-Milano</u> Clinical signs, disease categories and CNVs. Mucciolo M, Disciglio V, Mencarelli MA, Marozza A, Castagnini C, Dosa L, Lo Rizzo C, Di Marco C, Mari A, Renieri A.</p>
	<p><u>XIV Congresso Nazionale della Società Italiana di Genetica Umana (SIGU), November 13-16, 2011-Milano</u> Fenotipi opposti di BMI si associano al dosaggio genico</p>

	<p>della regione 16p11.2. V. Disciglio, M. Mucciolo, M.A. Mencarelli, C. Castagnini, M. Pollazzon, A. Marozza, F. Mari, A. Renieri.</p>
	<p><u>XIV Congresso Nazionale della Società Italiana di Genetica Umana (SIGU), November 13-16, 2011-Milano</u> Induced Pluripotent Stem Cells (iPSCs) to study Rett syndrome. De Filippis R, Amenduni M, Disciglio V, Mucciolo M, Epistolato MC, Ariani F, Mari F, Mencarelli MA, Hayek J, Renieri A, Meloni I.</p>
	<p><u>7th International Decipher Symposium, Wellcome Trust Genome Campus, May 23-25, 2011 – Hinxton (UK)</u> Clinical signs, disease categories and CNVs. Mucciolo M, Disciglio V, Mencarelli MA, Pollazzon M, Marozza A, Castagnini C, Dosa L, Lo Rizzo C, Di Marco C, Mari F, Renieri A.</p>
	<p><u>XIII Congresso Nazionale: Società Italiana di Genetica Umana, October 2010-Firenze</u> Microriarrangiamenti in 95 pazienti con diagnosi di spettro autistico. M. Mucciolo, V. Disciglio, F.T. Papa, J. Hayek, L. Dosa, C. Castagnini, A. Marozza, M. Pollazzon, M.A. Mencarelli, R. Canitano, A. Renieri, F. Mari.</p>
	<p><u>American Academy of Child & Adolescent Psychiatry, New York City, NY, October 27-31, 2010</u> Copy Number Variations in Autism Spectrum Disorders Canitano R, Mencarelli MA, Mucciolo M, Papa FT, Katzaki E, Marozza A, Radice L, Castagnini C, Dosa L, Pollazzon M, Hayek J, Renieri A, Mari F,.</p>
	<p><u>ESHG: European Human Genetics Conference, June 2010-Gothenburg</u> Autism Spectrum Disorders: emerging data from Copy Number Variations analysis Mucciolo M, Canitano R, Mencarelli MA, Papa FT, Katzaki E, Marozza A, Radice L, Castagnini C, Dosa L, Pollazzon M, Hayek J, Renieri A, Mari F.</p>
	<p><u>ESHG Annual Meeting, May 23-26, 2009 - Vienna, Austria</u> MS-MLPA to study the contribution of epigenetic silencing in Retinoblastoma. Amenduni M, Mucciolo M, Bruttini M, Sampieri K, Mencarelli MA, Epistolato MC, Toti P, Marozza A, Mari F, Hadjistilianou T, De Francesco S, Acquaviva A, Ariani F, Renieri A.</p>
	<p><u>59th ASHG Annual Meeting, October 20-24, 2009 -</u></p>

	<p><u>Honolulu, Hawaii (USA)</u> Array-CGH analysis to identify novel microdeletion/duplication syndromes and to extend the clinical phenotype associated with susceptibility regions. Papa FT, Katzaki E, Mucciolo M, Mencarelli MA, Uliana V, Pollazzon M, Marozza A, Bruccheri MG, Disciglio V, Ariani F, Meloni L, Mari F, Renieri A</p>
	<p><u>59th ASHG Annual Meeting, October 20-24, 2009 - Honolulu, Hawaii (USA)</u> Dissecting the 13q14 microdeletion syndrome to define the critical region for mental retardation. Amenduni M, Papa FT., Clark RD, Bruttini M, Disciglio V, Mencarelli MA, Epistolato MC, Toti P, Marozza A, Mari F, Hadjistilianou T, De Francesco S., Acquaviva A, Katzaki E, Mucciolo M, Ariani F, Renieri A.</p>
	<p><u>XII Congresso Nazionale SIGU, November 8-11, 2009 - Torino (IT)</u> Sindromi emergenti da microdelezioni. Mencarelli MA, Papa FT, Katzaki E, Mucciolo M, Disciglio V, Uliana V, Pollazzon M, Marozza A, Castagnini C, Dosa L, Bruccheri MG, Mari F, Renieri A.</p>
	<p><u>XII Congresso Nazionale SIGU, November 8-11, 2009 - Torino (IT)</u> Impiego dell'array-CGH nell'analisi di materiale abortivo. Papa FT, Marcocci E, Mucciolo M, Katzaki E, Mencarelli MA, Marozza A, Pollazzon M, Uliana V, Castagnini C, Mari F, Renieri A.</p>
	<p><u>XII Congresso Nazionale SIGU, November 8-11, 2009 - Torino (IT)</u> Identificazione di regioni a bassa penetranza e difficile interpretazione clinica. Mucciolo M, Papa FT, Katzaki E, Mencarelli MA, Marozza A, Pollazzon M, Uliana V, Dosa L, Bruccheri MG, Castagnini C, Buoni S, Canitano R, Radice L, Grosso S, Mostardini R, Balestri P, Hayek G, Mari F, Renieri A.</p>
	<p><u>XII Congresso Nazionale SIGU, November 8-11, 2009 - Torino (IT)</u> Nuova sindrome da microdelezione 8q22.3 in 5 pazienti con ritardo psicomotorio, disturbi comportamentali e caratteristiche faciali peculiari. Pollazzon M, Kuechler A, Papa FT, Mucciolo M, Katzaki E, Bohm D, Buysse K, Clayton-Smith J, Kohlhase J, David A,</p>

	Le Caignec C, Hayek G, Mencarelli MA, Marozza A, Uliana V, Renieri A, Mari F, Mortier G, Passarge E, Wieczorek D.
	<u>XII Congresso Nazionale SIGU, November 8-11, 2009 – Torino (IT)</u> Sindrome emegente da microdelezione 14q32.31-qter: descizione di due nuovi casi e revisione della letteratura. Papa FT, Mencarelli MA, Mucciolo M , Katzaki E, Disciglio V, Bruccheri MG, Hayek J, Weber RG, Mari F, Renieri A.
Publications	Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. <u>Nature.</u> 2011 Aug 31;478(7367):97-102. doi: 10.1038/nature10406.
	Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. <u>Clin Genet.</u> 2011 Jul 29. doi: 10.1111/j.1399-0004.2011.01755.x. Halgren C, Kjaergaard S, Bak M, Hansen C, El-Schich Z, Anderson C, Henriksen K, Hjalgrim H, Kirchoff M, Bijlsma E, Nielsen M, den Hollander N, Ruivenkamp C, Isidor B, Le Caignec C, Zannolli R, Mucciolo M , Renieri A, Mari F, Anderlid BM, Andrieux J, Dieux A, Tommerup N, Bache I.
	Investigation of modifier genes within copy number variations in Rett syndrome. <u>J Hum Genet.</u> 2011 Jul;56(7):508-15. doi: 10.1038/jhg.2011.50. Artuso R, Papa FT, Grillo E, Mucciolo M , Yasui DH, Dunaway KW, Disciglio V, Mencarelli MA, Pollazzon M, Zappella M, Hayek G, Mari F, Renieri A, Lasalle JM, Ariani F.
	Alport syndrome and leiomyomatosis: the first deletion extending beyond COL4A6 intron 2. <u>Pediatr Nephrol.</u> 2011 May;26(5):717-24. Epub 2010 Dec 14. Uliana V, Marcocci E, Mucciolo M , Meloni I, Izzi C, Manno C, Bruttini M, Mari F, Scolari F, Renieri A, Salviati L.
	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 Vera Uliana, Salvatore Grosso, Maddalena Cioni, Francesca Ariani, Filomena T Papa, Silvia Tamburello, Elisa Rossi, Eleni Katzaki, Mafalda Mucciolo , Annabella Marozza,

	Marzia Pollazzon, Maria Antonietta Mencarelli, Francesca Mari, Paolo Balestri, Alessandra Renieri.
	Is HSD17B1 a new sex reversal gene in human? <u>Mol Cell Endocrinol.</u> 2009 Dec 10;313(1-2):70. Katzaki E, Papa FT, Mucciolo M , Uliana V, Renieri A.