Department of Health Sciences

University of Eastern Piedmont

PhD in Molecular Medicine

Coordinator: PROF. EMANUELE ALBANO

Search for genomic micro-rearrangements through array-CGH in patients with Intellectual Disability and Multiple Pituitary Hormone Deficiency, with standard and custom-design array platforms.

PhD candidate

DOTT.ssa ILEANA FUSCO

Supervisor

PROF. MARA GIORDANO

XXVII CYCLE

2011-2014

INDEX

1. INTRODUCTION

1. INTRODUCTION

1.1 Copy Number Variations

The spectrum of human genetic variation ranges from the single base pair to large chromosomal rearrangements. It has become apparent that human genomes differ more as a consequence of structural variation (SV) than of single-base-pair difference. About 10 years ago, scientists began to recognize abundant variation of an intermediate size class known as structural variation. SV was originally defined as insertions, deletions and inversions greater than 1kb in size. [1,2]

Historically, structural variation was primarily assayed cytogenetically in diseased genomes. The first report of disease caused by mega base sized genomic rearrangements was published in the early 1990s [3].

For most of twentieth century geneticists focused on large scale genomic rearrangements because these were the only variants that were visible by early cytogenetic methods. With the sequencing of human genomes now becoming routine, the operational spectrum of structural variants (SVs) and copy number variants (CNVs) has widened to include much smaller events (for example, those >50 bp in length). The challenge now is to discover the full extent of structural variation and to be able to genotype it routinely in order to understand its effects on human disease, complex traits and evolution [1,2].

The recent application of massively parallel sequencing methods has complemented microarray-based methods and has led to an exponential increase in the discovery of smaller structural-variation events. The first studies to assay CNVs primarily used microarrays with large-insert clones (known as bacterial artificial chromosomes (BACs) and oligonucleotide arrays. Subsequent microarrays have included both comparative genomic hybridization (CGH) and SNP-based arrays. With the advent of NGS technologies, new tools can detect CNVs from these data, and the use of these approaches has been particularly widespread during the past few years [1,2].

These alterations may have no phenotypic effect, account for adaptive traits or can underlie a disease.
Clinical geneticists need to discriminate pathogenic or high-risk variants from benign variants in patients. Database of CNVs found in apparently healthy individuals, which can include parents, siblings and population controls represents a valuable tool to the identification of polymorphic benign variants. A major challenge in the field is that there is a growing number of CNVs (known as 'variants of unknown significance') that are suspected to be involved in disease susceptibility but for which additional population-level data are required [2].

Recently, Zarrei M. et al. [2] have published high-quality data on healthy individuals of various ethnicities to construct an updated CNV map of the human genome. Depending on the level of stringency of the map, have been estimated that 4.8–9.5% of the genome contributes to CNV and found approximately 100 genes that can be completely deleted without producing apparent phenotypic consequences. This map will aid the interpretation of new CNV findings for both clinical and research applications [2].

1.1.1 CNV classification

A copy number variant is defined as a segment of DNA that differs in copy number compared with a representative reference genome. The term "CNV" does not imply clinical significance; therefore, a qualifier such as pathogenic or benign CNV is necessary for clear communication of clinical relevance. In addition, the term "CNV" does not imply relative dosage. Copy number loss (deletion) or copy number gain (duplication) must be specified to clarify the nature of a variation.
In many cases, the interpreting geneticist can assess the genomic content of the CNV, correlate with established clinical literature, and provide an interpretation that is unambiguous and consistent with the interpretation derived from multiple laboratories. However, given the presence of benign variants within the genome and the continual discovery of novel CNVs, assessing the clinical significance of CNVs found in a clinical setting can be challenging.
Accordingly, when a variation is extremely rare or limited clinical literature is available, interpretation and reporting practices may vary among laboratories.

Finding category	*Subcategory*	*Definition and/or subclasses*	*Examples*
Pathogenicfor the proband (ie, fetus)	**Causative findings**	Pathogenic finding explaining the phenotype or matching the indication	•22q11 microdeletion in a proband with a tetralogy of Fallot •Trisomy 21 in a fetus referred for cytogenetic testing due to an abnormal first trimester screening (1:20 risk for Down syndrome) • Mosaic terminal duplication 2q31.1q37.3 in a child with ASD (detected on B-Allelic frequency plot, not detected with routine karyotyping and on LogR ratio plot)[22] • UPD (7) in a proband with failure to thrive[22]
	Unexpected diagnoses	Pathogenic findings NOT explaining the phenotype or NOT matching the indication. (a) early-onset treatable diseases (b) early-onset untreatable diseases (c) late-onset treatable diseases (d) late-onset untreatable diseases	(a) a deletion in band 10q11.21 including RET gene associated with multiple endocrine neoplasia IIA (b) DMD deletion in a male fetus referred for prenatal diagnosis because of an abnormal first trimester screening (c) a deletion in CHEK2 associated with a moderately increased risk of breast cancer and risk of other cancers in a proband with severe global developmental delay[13] (d) microdeletion of 17p12 (PMP22 gene) associated with a hereditary neuropathy with liability to pressure palsies (HNPP)
	Susceptibility loci	Variants associated with neurodevelopmental disorders, but of extreme phenotypic heterogeneity and/or variable expressivity.[20]	16p12.1 deletion in a patient with developmental delay[19]
VOUS(variants of unknown clinical significance)	—	(a) potentially pathogenic (without 'enough' evidence)[1, 20] (b) truly VOUS (unknown significance)[1] (c) likely benign (without 'enough' evidence for benign)[1]	
Benign findings	—	(a) benign (found in many healthy individuals)[1] (b) polymorphic (found in >1% of the general population)[1]	
Status for recessive diseases	—	Comprehensive reporting of heterozygous recessive mutations is not recommended. However, carrier status in case of a well-characterized recessive disorder with a reasonably high population frequency and/or with clinical features consistent with the patient's reason for referral, may be considered for disclosure.[21]	Deletion of CFTR gene
Incidental findings	—	Abnormalities found by chance, unintentionally, in parents of probands	Mosaic Turner syndrome discovered on B-allelic frequency plot during quality control of the array profile of a pregnant woman referred for prenatal diagnosis due to foetal ultrasound abnormalities

Table 1. A proposal for a generic classification of array findings [5].

A CNV classification is already proposed and published by several authors. The CNV is documented as clinically significant in multiple peer-reviewed publications, even if penetrance and expressivity of the CNV are known to be variable [4].
In 2014 Srebniak MI et al. proposed a generic classification of array findings, defining several subcategories (Table 1) [5].
However an international classification and terminology for array findings are indispensable in order to avoid miscommunication, to facilitate comparing cohorts studied by different researchers and to optimize pre-test counseling.

1.1.2 Mechanisms of formation of Copy Number Variations

Large-scale genomic changes, such as copy number variations (CNVs), are being increasingly discovered as a genetic source of human diversity and the pathogenic factors of diseases. Such clinical conditions resulting from human genome architecture are termed 'genomic disorders'. Although CNVs are found throughout the genome, some regions are more prone to rearrangements than others. Recent experimental findings have explained the links between different genome architectures and CNV mutagenesis [6,7].

CNVs can be divided into two classes: recurrent CNVs, which occur in a few genomic positions with breakpoints within low copy repeats (LCRs) that are clustered together in unrelated individuals, and nonrecurrent CNVs, which are distributed throughout the genome and have unique breakpoints [6,7].

The genomic repeats (representing DNA primary structures) can be divided into LCR/or segmental duplication (SD) and high-copy repeats, which play a significant role in human evolution and have a strong association with genomic CNVs.

Stankiewicz and Lupski [8] defined region-specific low-copy repeats (LCRs) as paralogous genomic segments spanning 10–400 kb of genomic DNA and sharing ≥95%– 97% sequence identity. Compared to LCR/SD, high-copy repeats constitute a great portion of the human genome.

The most common type of high-copy repeats are the interspersed repeats, which cover over 44% of the human genome. One of the major classes of interspersed repeats is the retrotransposon, including short interspersed elements (SINEs), long interspersed elements (LINEs), and endogenous retroviruses (ERVs).

The origin of recurrent CNVs has been attributed mainly to non-allelic homologous recombination (NAHR) or defective crossing over during meiosis, when misalignment occurs due to the high degree of sequence similarity between LCRs (Fig. 1). Other studies demonstrated that also high-copy repeats, in particular Alu sequence (the most common SINEs), can induce CNV formation via classical DNA recombination based mechanisms, such as NAHR. [7,8].

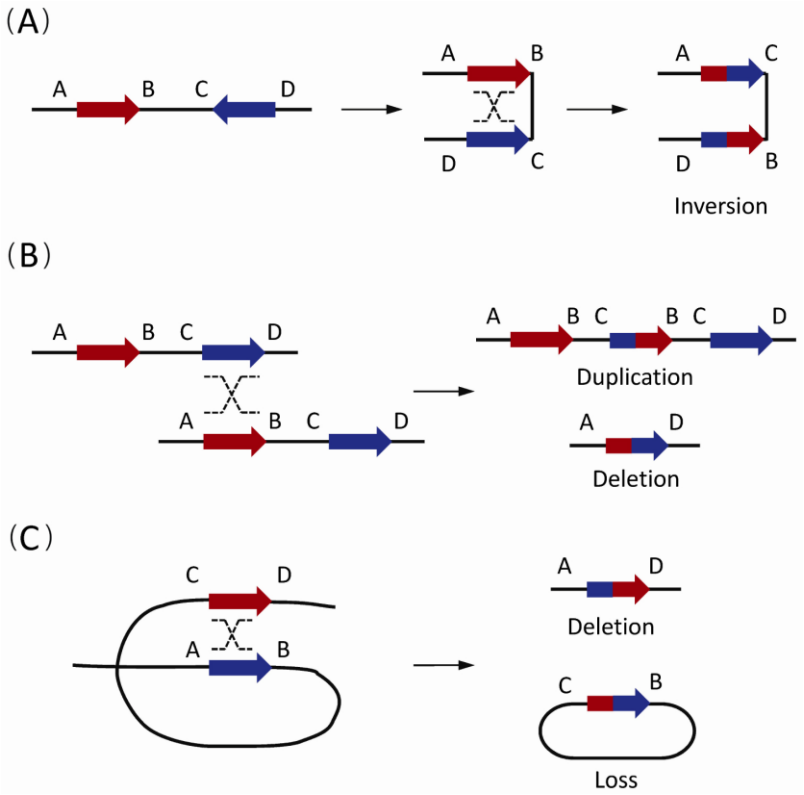

Fig. 1. The non-allelic homologous recombination (NAHR) events between paired low-copy repeats (LCRs)/segmental duplications (SDs). (A) The NAHR event between reversely oriented LCRs/SDs can cause inversion, a copy-neutral structural variation. (B) The inter-chromatid NAHR events between directly oriented LCRs/SDs result in deletions and duplications. (C) The intra-chromatid NAHR events between directly oriented LCRs/SDs can generate deletions and ring-shaped DNA segments that will be lost in subsequent cell divisions [6].

Nonrecurrent CNVs have scattered breakpoints and the boundaries share limited or no nucleotide homology. The origin of nonrecurrent CNVs is less clear. This categories of CNVs are of variable size and may arise via mechanisms like nonhomologous end joining (NHEJ) and replication-based mechanisms described by the fork stalling and template switching (FoSTeS) and microhomology-mediated break-induced replication (MMBIR) models.

It is becoming clear that most disease causing CNVs are nonrecurrent and generally arise via replication-based mechanisms. Patients carried nonrecurrent CNVs may share a critical region whose copy number changes results in common clinical features [6].

1.2 History of cytogenetics

In 1879 the Human cytogenetics research began with the observations of the German pathologist Arnold, who examined carcinoma and sarcoma cells because the voluminous nuclei of these cells facilitated analysis. Later, Flemming and Hansemann were the first to examine human mitotic chromosomes.

In 1888, Waldeyer coined the term "chromosome" after staining techniques had been developed to make them more discernible (*chromos* = Greek for colour; *soma* = Greek for body) [9].

In 1937 Blakeslee and Avery discovered the induction of chromosome doubling by colchicine treatment in plants [10]. After that it has been investigated the cytological mechanism causing this doubling and in the 1930s the use of colchicine for chromosome preparations has been implemented in plant cytogenetics [11].

Until Eagle developed specific culture media in 1955, the cytogenetic analysis of chromosomes depended on spontaneously dividing cells. Tjio and Levan (1956), using cultured embryonic cells, were the first researchers to report the diploid number of human chromosomes as 46 [9].

This enabled the detection of numerical chromosome aberrations like trisomy 21 in Down syndrome, 45,X in Turner syndrome, 47,XXY in Klinefelter syndrome, trisomy 13, trisomy 18, and Philadelphia chromosome in a patient with chronic myeloid leukaemia [12].

Later, many more chromosomal abnormalities, such as other trisomies, translocations, inversions, insertions, deletions, duplications and complex chromosome rearrangements, have been described.

Novel methods for investigating the mechanisms underlying copy number changes, characterizing gene interactions and analyzing genes within copy number variations have been developed [9].

Moorhead *et al.* (1960) established an in vitro culture method for the accumulation of dividing cells using colchicine to arrest cells at metaphase [13]. Subsequent technical improvements in cytogenetics included the use of phytohemagglutinin (a substance that stimulates the division of T lymphocytes in vitro) and the introduction of banding techniques at the end of the 1960s.

Banding techniques use chemical treatments to produce differentially stained regions on chromosomes. The banding pattern is highly characteristic for each chromosome and facilitates the complete identification of the human karyotype. Because the development of newer techniques has continued to increase the resolution of chromosomes, human cytogenetics has evolved from a more basic science into a valuable strategy for diagnosing prenatal, postnatal and acquired chromosomal abnormalities [1].

Conventional G-banded karyotyping remains the gold standard, but it is time-consuming (about 10-14 days is required to obtain the result), labor-intensive and the limitations to detect imbalances larger than 5–10 Mb [14]. The considerable gap between the limited resolution for observing chromosome structure through banding techniques at the light microscopy and gene levels was bridged after the introduction and application of several molecular cytogenetic approaches [9].

In 1986, Pinkel et al. [15-16] developed a method to visualize chromosomes using fluorescent-labelled probes called fluorescent in situ hybridization **(**FISH). FISH technology permits the detection of specific nucleic acid sequences in morphologically preserved chromosomes, cells, and tissues and the detection of specific microdeletion syndromes [12].

There are currently a number of commercially available FISH probes for the most common disorders and this method is still predominantly used when the clinical phenotype is suggestive of a particular disorder [17].

Importantly, FISH permits karyotype analysis of nuclei in non-dividing cells and has proven to be advantageous in many ways, but it is time consuming because preparations must be hybridised and then microscopically analysed. These problems led to the development of a variation of FISH called Comparative Genomic Hybridization (CGH) [12,18].

The CGH technique is an efficient approach to genome-wide screening for chromosomal copy number changes (gains/duplications and losses/deletions) within a single experiment. This technique was originally introduced to study chromosomal abnormalities that occur in solid tumors and other malignancies, which can be difficult to culture and is based on using tumor DNA extracted directly from either fresh or archival tissue [9].

CGH uses two genomes, a test and a control, which are differentially labeled and competitively hybridized to metaphase spreads from chromosomally normal individuals (Fig. 2). The fluorescent signal intensity of the labeled test DNA relative to that of the reference DNA can then be linearly plotted across each chromosome, allowing the identification of copy number changes [18,19].

Although chromosomal CGH has increased the potential for identifying new chromosomal abnormalities, this technique is time consuming and in addition to a poor reproducibility, does not significantly improve resolution (> 3 Mb) compared with routine G-banding chromosome analysis [9].

Later, the development of array-based CGH (array-CGH) approaches involving the substitution of metaphase chromosomes with DNA sequences adhered onto glass slides (Fig. 2), has increased the resolution for detecting copy number changes in the human genome [20,21].

It provides better quantification of copy number and more precise information on the breakpoints of segments that are lost or gained using conventional CGH [22], with a resolution that is dependent on the spacing of the probes.

For clinical testing the resolution generally involves a spacing of 50 Kb to 1 Mb between adjacent probes on the array, often with additional coverage at the subtelomeric regions. [23-24].

Because both genome-wide array CGH analysis and conventional karyotyping, based on staining chromosomes, aim to identify chromosomal aberrations by screening the genome, Vermeesch et al. proposed to call this novel technology molecular karyotyping. [25]

1.3 Array-CGH

Many human genetic disorders result from unbalanced chromosomal abnormalities, in which there is net gain or loss of genetic material. Traditionally, cytologists have detected such abnormalities by generating a karyotype of a person's chromosomes. In recent years, however, researchers have increasingly turned to newer cytogenetic techniques: firstly FISH, then comparative genomic hybridization and recently has been developed a newer method that combines the principles of CGH with the use of microarrays [26].

1.3.1 Array CGH Methodologies

The principle of array CGH was first described in 1997 by Solinas Toldo [20]. The first applications were published by Pinkel, Snidjers and Buckley in the following years [22].

Array-CGH is the equivalent of conducting thousands of FISH experiments at once, and it provides better quantification of copy number and more precise information on the breakpoints of segments that are lost or gained with conventional CGH [12,22].

The principle is very similar to that employed for traditional CGH, where total genomic DNA obtained from control cells and test samples are differentially labeled using green (fluorescein isothiocyanate, FITC) and red (Texas red) fluorescent dyes, denatured, co-precipitated in the presence of blocking DNA to suppress repetitive sequences and subsequently co-hybridized to normal metaphase chromosomes (Fig. 2) [9,22]. A copy number difference between test DNA and reference DNA, will alter the fluorescence ratio of the two dyes along the targets, so a chromosomal imbalances across the genome can thus be quantified and positionally defined.

In the array-CGH, instead of metaphase spreads, the hybridization targets are thousands of spots of reference DNA sequences, immobilized in a precisely gridded manner on the slide (Fig. 2).

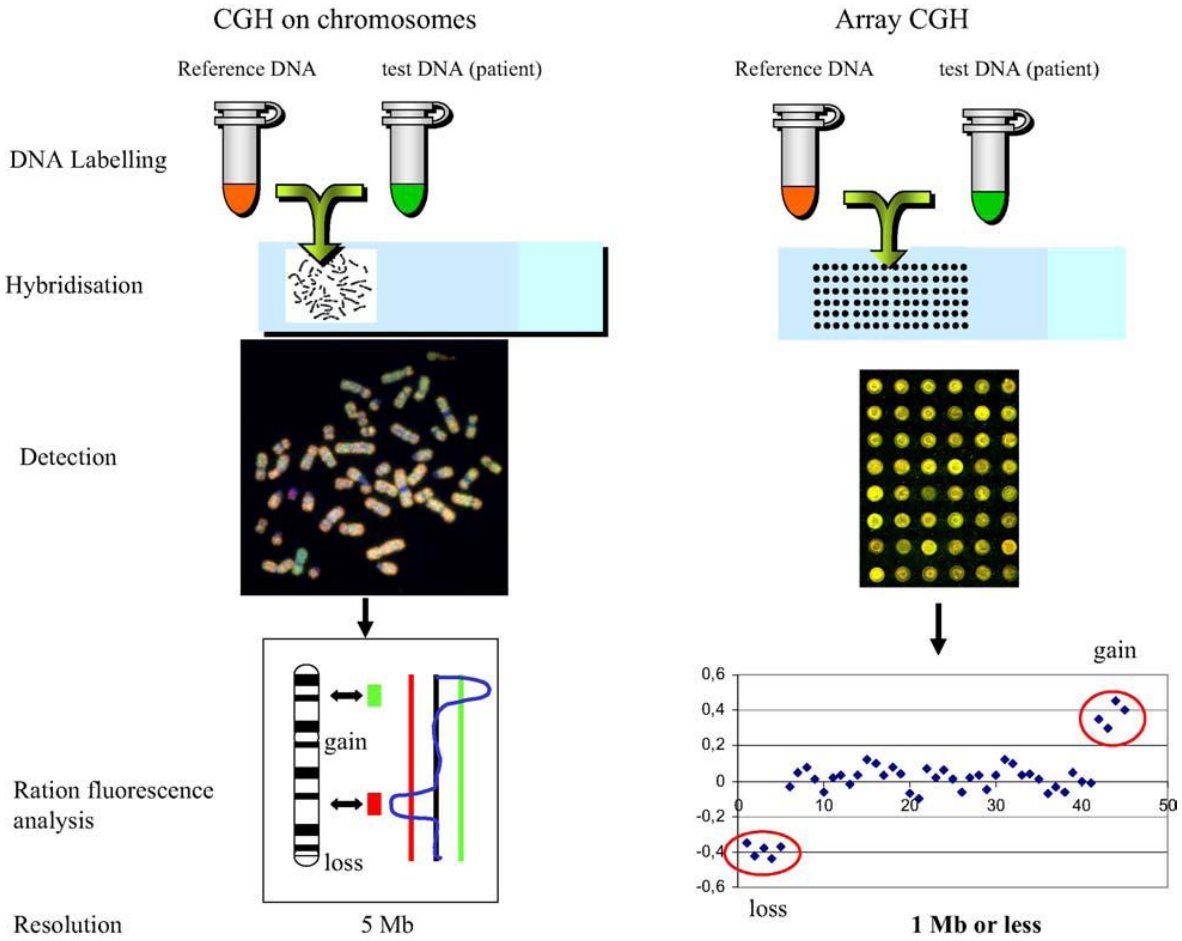

Fig. 2. Comparison between CGH on chromosomes on the left and array CGH on the right. For both techniques genomic DNA of a patient (test DNA) are labeled by a fluorochrome and genomic DNA of a normal subject (reference DNA) labeled by another fluorochrome. These two DNA are cohybridized on chromosomes or genomic DNA fragments spotted on a glass. The fluorescence is captured using a camera (CGH on chromosomes), or a scanner (array CGH). The fluorescence ratio is measured by a software and presented in graphic form.[22]

Data processing plays a central role in interpretation. Various softwares have been developed allowing to process the data, facilitate the interpretation and to generate a graph of the results automatically [22]. The resolution of array CGH depends on the size of the genomic fragments as well as on their density [22,25].

The initial arrayed DNA segments could be larger (~150 kb) human DNA segments inserted into a BAC clones or bacterial/P1-derived artificial chromosomes (PAC clones). As the resolution of the array yields improved, shorter sequences have been used as targets, including smaller cDNA fragments, PCR products and oligonucleotides [9].

Thus, soon after the first BAC arrays were introduced, various companies developed arrays containing single-stranded oligonucleotide probes. Unlike BAC arrays, oligonucleotide arrays do not have the same underlying design concept. Variations include manufacturing method, sample throughput capacity, probe length (25–80 bases), and presence or absence of single nucleotide polymorphisms (SNPs) within the probes.

All available oligonucleotide platforms are whole-genome arrays, but they can be customized either to function as targeted arrays (similar to available targeted BAC arrays) or to substantially increase resolution in a specific genomic region of interest. This versatility will eventually prove useful as array-CGH technology becomes widespread [27].

Furthermore, array-CGH provided resolution at the nucleotide level. SNP arrays have the highest resolution (5-10 kb) of all of the available array-based platforms. The co-hybridization of the test and reference DNAs is not required because the test DNA can hybridize directly to the SNP array. In addition to CNVs, the genotype information obtained from SNP arrays enables the detection of stretches of homozygosity and thus the identification of recessive disease genes, mosaic aneuploidy or uniparental disomy (UPD) [9].

While only SNP arrays enable the detection of copy number-neutral regions in the absence of heterozygosity (AOH), these arrays have limited ability to detect single-exon copy CNVs due to the distribution of SNPs across the genome. Combining both array-CGH and SNP genotyping in a single platform optimizes the clinical diagnostic capability, offering the simultaneous detection of copy number neutral and small intragenic copy number changes [9].

The current limitations of the technology include the inability to detect balanced chromosome rearrangements and the equivocal nature of copy number alterations of unknown significance that may be identified. Nevertheless, it was used routinely in the clinical setting for example with a normal chromosome result in cases of intellectual disability and/or multiple congenital anomalies (MCA); as a result the diagnostic yield in this patient group has increased considerably [28].

1.3.2 Array platforms

As already discussed before, it has become apparent that copy number variants are quite common in the human genome and can have dramatic phenotypic effects as a result of altering gene dosage, disrupting coding sequences, or perturbing long-range gene regulation. These DNA anomalies are associated with many genetic diseases including congenital anomalies, developmental delay, and mental retardation. As a result, many arrays have been designed to diagnose these DNA alterations as well as to detect gains and losses. Identifying the specific segmental genomic alterations and the genes they contain will yield molecular targets for diagnostics and therapy. Recent studies characterized the reliability and accuracy of various CGH technologies to detect chromosomal aberrations.

In the 2007 Association of Biomolecular Resource Facilities Microarray Research Group project analyzed HL-60 DNA with five platforms, a test genome, and analysis software that would facilitate comparison of the resolution of each platform. At the time of the study, these platforms represented the state of the art for detecting chromosomal aberrations: Agilent CGH 44B Microarray, Illumina HumanHap 550 BeadChip, Affymetrix GeneChip Human Mapping 500K Array Set, a human BAC19K array developed by Roswell Park Cancer Institute, and the Affymetrix Human Genome U133 Plus 2.0 gene expression array.
Each platform was assessed on its repeatability between replicates and on detection of the reported gains and losses. It has been found that all five of the arrayCGH platforms detected 100% of the eight previously reported CNVs in almost all replicates. Thus, at this level of resolution, the selection of a platform may depend more on practical considerations such as price than on a substantial difference in technical performance [29].

Recently, Rodríguez-Revenga L et al. investigated the practical performance of three different oligonucleotide-based microarray platforms for their implementation in the molecular diagnosis of patients with idiopathic developmental delay/intellectual disability (DD/ID) [30]
The results obtained using a custom microarray (KaryoArray®) and two different commercial medium- and high-resolution whole-genome oligonucleotide microarrays (Agilent 4x44K and Agilent 244K) have been compared.

The first combines targeted and whole genome oligonucleotides based on the clinical and laboratory expertise.

An overall diagnostic yield of around 15% has been obtained. Their study indicated that all three platforms were suitable for detecting causal CNVs and the frequency of VOUS was similar (among 4–6% of cases).

As expected, whole custom genome-targeted array revealed less common CNVs than did the commercial platforms, which has a clear advantage in a clinical setting. This difference might be explained due to the fact that with custom array have been analyzed patients with clinical selection criteria as well the molecular characterization stricter than the others studied with the commercial one [30].

1.3.3 Customized array

The recent emergence of array comparative genomic hybridization has revolutionized the ability to identify micro-rearrangements associated with various diseases.

In recent years, CGH method using custom-designed high-density oligonucleotide-based arrays allowed the development of a powerful tool for detection of alterations at the scale of exons and made it possible to provide flexibility through the possibility of increasing the density of probes loaded on chips. This allowed the emergence of high-density (HD) chips arrays thereby increasing the number of genes tested.

Several study have been demonstrated that the contribution of this technique is undeniable compared with conventional techniques. It appeared reproducible, reliable and sensitive enough to detect heterozygous single-exon deletions or duplications, complex rearrangements and somatic mosaicism. The possibility of simultaneous analysis of several genes and its scalability make it a valuable tool for a new diagnostic approach of CNVs and should facilitate the molecular diagnosis of heterogeneous groups of diseases such as muscular dystrophies16 or mental retardation.

In addition, it improves reliability of variation detection and allows determination of boundaries precisely enough to direct targeted sequencing of breakpoints. Possibility of high scale sequencing of breakpoints will bring real progress in understanding molecular mechanisms of rearrangements, searching for genotype–phenotype correlations and to guide certain therapeutic strategies [31].

Another remarkable aspect of Customized High Resolution CGH-Array was the identification of additional cryptic submicroscopic imbalances in known pathogenic samples previously established by karyotype and Multiplex Ligation-dependent Probe Amplification (MLPA)/FISH, at least when phenotype does not match completely with genotype.

These findings may have additional implications on the interpretation of results regarding the complexity of the extra rearrangement and for genetic counseling. It also suggests that trisomies or submicroscopic deletions/duplications may co-exist with additional rearrangements, which in some way might contribute to the abnormal phenotypes.

Finally, it has been shown that a well design customized aCGH tool (such as KaryoArray v3.0; Agilent-based 8x60 K), may provide a good balance among prize/resolution, an excellent sensitivity and specificity for introducing Molecular Karyotyping in clinical practice [32].

2. AIM OF THE STUDY

2. AIM OF THE STUDY

The aim of the present study was the search for genomic micro-rearrangements through a standard array-CGH platform in a cohort of patients with intellectual disability and through a customized platform in a subset of short stature patients with pituitary anomalies.

The objective of the first part of my research was to analyze 53 patients for the presence of microrearrangements in the whole genome for diagnostic and research purposes and to dissect the phenotype in order to identify genotype-phenotype correlation and assign a possible pathogenic role to genes included in the rearranged region.

The second part of the research was the identification of genes on a candidate region on chromosome X, involved in short stature consequent to growth Hormone (GH) deficiency. This part of the project stems from the observation of an excess of males among the GHD patients (on average the male:female ratio is 2:1) and from the identification of genomic rearrangements on the Xq13-q21 region, identified with standard cytogenetic methods in males with severe short stature and other anomalies. These data indicate the presence of X-linked genetic factors strongly involved in GHD and that could explain a part, or even a large part, of the at present unexplained genetic component of the disease in males. We planned to investigate this region on the X chromosome for the presence of micro-rearrangements in a cohort of male patients affected by IGHD (Isolated Growth Hormone Deficiency) and CPHD (Combined Growth Hormone Deficiency), in which the presence of mutations in genes already known to be involved in GHD has been excluded. We used a customized array platform (Agilent) containing high oligomers coverage for the region Xq13.q21, strongly candidate for containing genes or cis-regulatory elements involved in GHD.

3. MATERIALS AND METHODS

3. MATERIALS AND METHODS

3.1 Samples preparation: gDNA Extraction

Use reagents from the ReliaPrepTM Blood gDNA Miniprep System (Promega)

- Equilibrate a thermomixer and heat block or water bath to 56°C.
- Thoroughly mix the blood sample for at least 10 minutes in a rotisserie shaker at room temperature. If the blood has been frozen, thaw completely before mixing for 10 minutes.
- Dispense 20µl of Proteinase K (PK) Solution into a 1.5ml microcentrifuge tube.
- Add 200µl of blood to the tube containing the Proteinase K (PK) Solution, and briefly mix.
- Add 200µl of Cell Lysis Buffer (CLD) to the tube. Cap and mix by vortexing for at least 10 seconds. This vortexing step is essential for obtaining good yields.
- Incubate at 56°C for 10 minutes.
- While the blood sample is incubating, place a ReliaPrep™ Binding Column into an empty Collection Tube.
- Remove the tube from the heating block. Add 250µl of Binding Buffer (BBA), cap the tube, and mix by vortexing for 10 seconds with a vortex mixer.

  Note: The lysate should be dark green at this point. This vortexing step is essential for obtaining good yields.
- Add the contents of the tube to the ReliaPrep™ Binding Column, cap it and place it in a microcentrifuge.
- Centrifuge for 1 minute at maximum speed. Check the binding column to make sure the lysate has completely passed through the membrane. If lysate is still visible on top of the membrane, centrifuge the column for another minute.

  Note: The sample can be centrifuged at lower speed, if desired. Increase the centrifugation time accordingly to ensure the lysate has completely passed through the membrane.
- Remove the collection tube containing flowthrough, and discard the liquid as hazardous waste.

- Place the binding column into a fresh collection tube. Add 500µl of Column Wash Solution (CWD) to the column, and centrifuge for 3 minutes at maximum speed. Discard the flowthrough.

  Note: If any of the wash solution remains on the membrane, centrifuge the column for another minute.
- Repeat the previous wash-Step twice for a total of three washes.
- Place the column in a clean 1.5ml microcentrifuge tube. 14. Add 50µl of Nuclease-Free Water to the column. Centrifuge for 1 minute at maximum speed.

  Note: Eluting in 50µl significantly increases the concentration of the DNA but reduces yield by 25–30%. 15.
- Discard the ReliaPrep™ Binding Column, and save eluate. Do not reuse binding columns or collection tubes.
-

3.2 gDNA Quantitation and Quality Analysis

Use the NanoDrop ND-1000 UV-VIS Spectrophotometer (or equivalent) to assess gDNA concentration and purity.

UV-VIS Spectrophotometry

- In the Nanodrop program menu, select Nucleic Acid Measurement, and then select Sample Type to be DNA-50.
- Use 1.5 µL of Nuclease-Free Water to blank the instrument.
- Use 1.5 µL of each gDNA sample to measure DNA concentration. Record the gDNA concentration (ng/µL) for each sample.
- Record the A260/A280 and A260/A230 ratios. High-quality gDNA samples have an A260/A280 ratio of 1.8 to 2.0, which indicates the absence of contaminating proteins. Scanning the absorbance from 220-320 nm will show whether contaminants exist that affect absorbance at 260 nm. Check the absorbance scans for a peak at 260 nm and an overall smooth shape. The ideal 260/230 ratio for pure DNA is of 2 to 2.4.

3.3 Array-based CGH

Array based CGH analysis was performed using commercially available oligonucleotide microarrays containing about 41,000 60-mer probes (SurePrint G3 Human CGH Microarray Kit 8×60K AGILENT).

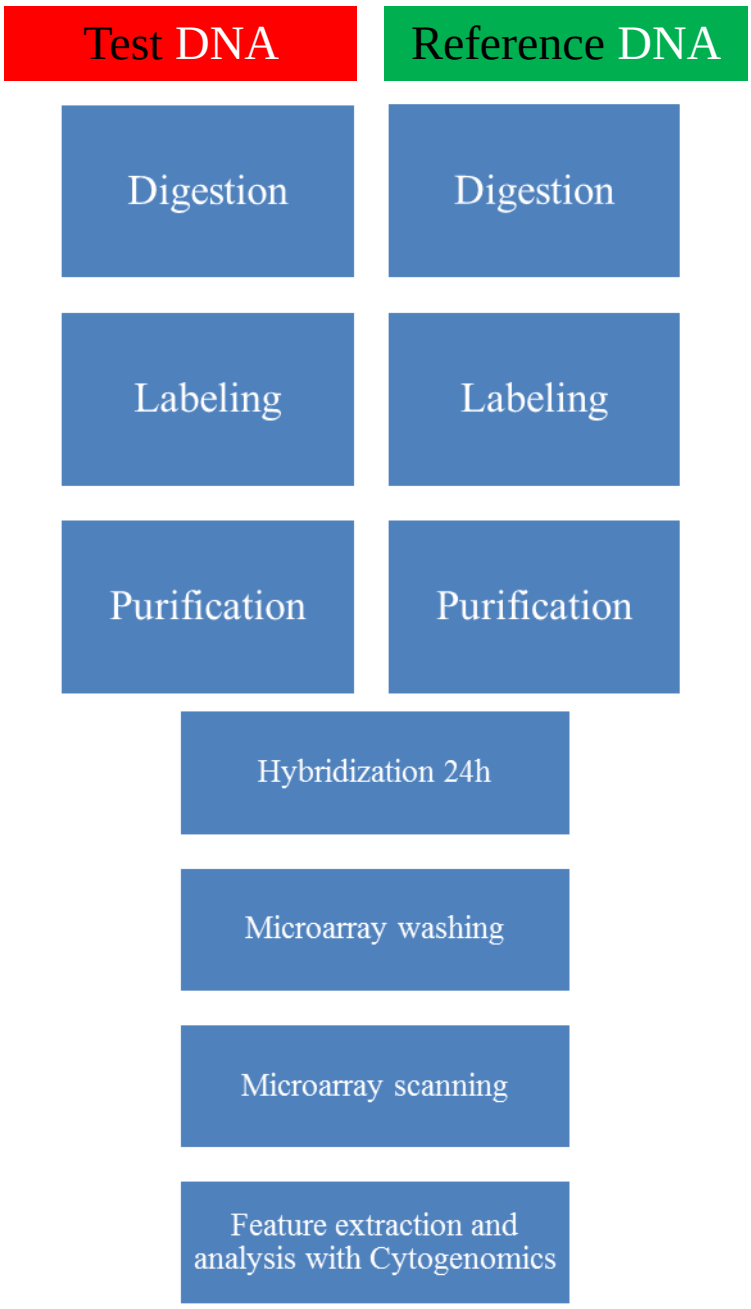

Figure 3 : Direct workflow for sample preparation and microarray processing.

A) Samples preparation

The samples preparation was executed essentially according to the Agilent protocol (Oligonucleotide Array-Based CGH for Genomic DNA Analysis 7.3v, Fig 3).

The patient/reference DNAs (400ng) were mixed with MilliQ water to a total volume of 13 µl.

B) Digestion of gDNA

- The patient/reference DNAs were mixed with 2.5 µl Random primer solution.
- Transfer sample tubes to a thermal cycler:

Step	Temperature	Time
Step 1	98°C	10 minutes
Step 2	4°C	hold

Table 2: DNA denaturation and fragmentation using a thermal cycler.

C) Fluorescent Labeling of gDNA

- Each sample was added with 5 µl of 5x Reaction Buffer, 2.5 µl of 10x dNTPs mix, 1.5 µl of Cy5-dUTP (test sample) or 1.5 µl of Cy3-dUTP (reference sample) and with 0.5 µl of Exo-Klenow.
- Transfer sample tubes to a thermal cycler:

Step	Temperature	Time
Step 1	37°C	2 hours
Step 2	65°C	10 minutes
Step 3	4°C	hold

Table 3: DNA labeling using a thermal cycler.

D) Purification of labeled gDNA

- Spin the labeled gDNA samples in a centrifuge for 1 minute at 6,000 × g to drive the contents off the walls and lid.
- Add 430 µL of 1×TE (pH 8.0) to each reaction tube.
- For each gDNA sample to be purified, place a column into a 2-mL collection tube and label the column appropriately. Load each labeled gDNA onto a column.
- Cover the column with a cap and spin for 10 minutes at 14,000 × g in a microcentrifuge at room temperature. Discard the flow-through and place the column back in the 2-mL collection tube.
- Add 480 µL of 1×TE (pH 8.0) to each column. Spin for 10 minutes at 14,000 × g in a microcentrifuge at room temperature. Discard the flow-through.
- Invert the column into a fresh 2-mL collection tube that has been appropriately labeled.
- Spin for 1 minute at 1,000 × g in a microcentrifuge at room temperature to collect purified sample. The volume per sample will be approximately 20 to 32 µL.
- Add 1×TE (pH 8.0), or use a concentrator to bring the sample volume to 9.5 µL. Do not excessively dry the gDNA because the pellets will become difficult to resuspend.
- Take 1.5 µL of each sample to determine yield and specific activity.
- In a fresh 200 µL Thin-Wall Tube, combine test and reference sample using the appropriate cyanine-5-labeled sample and cyanine-3-labeled sample for a total mixture volume of 16 µL.

E) Hybridization

- Mix the components to prepare the Hybridization Master Mix.

Component	Volume (µL) per hybridization	Volume (µl) per 10 samples
Cot-1 DNA (mg/ml)	2	20
10x aCGH Blocking Agent	4.5	45
2x HI-RPM Hybridization Buffer	22.5	225
Final volume of Hybridization Master Mix	**29**	**290**

Table 4 : Hybridization Master Mix.

- Add the appropriate volume of the Hybridization Master Mix to the 200 µL Thin-Wall Tube that contains the labeled gDNA to make the total volume of 45 µL
- Mix the sample by pipetting up and down, then quickly spin in a centrifuge to drive contents to the bottom of the reaction tube.
- Transfer sample tubes to a thermal cycler.

Step	Temperature	Time
Step 1	98°C	3 minutes
Step 2	37°C	30 minutes
Step 3	37°C	hold

Table 5 : Thermal cycler program

- Remove sample tubes from the thermal cycler.
- Spin 1 minute at 6000 × g in a centrifuge to collect the sample at the bottom of the tube.
- The samples are ready to be hybridized
- Load a clean gasket slide into the Agilent SureHyb chamber base with the gasket label facing up and aligned with the rectangular section of the chamber base.
- Slowly dispense hybridization sample mixture onto the gasket well
- Put a microarray slide “active side” down onto the gasket slide, so the numeric barcode side is facing up and the “Agilent”-labeled barcode is facing down. Assess that the sandwich-pair is properly aligned.
- Put the SureHyb chamber cover onto the sandwiched slides and slide the clamp assembly onto both pieces.
- Hand-tighten the clamp firmly onto the chamber.
- Vertically rotate the assembled chamber to wet the slides and assess the mobility of the bubbles. Tap the assembly on a hard surface if necessary to move stationary bubbles.
- Load each assembled chamber into the oven rotator rack.
- Set your hybridization rotator to rotate at 20 rpm.
- Hybridize at 67°C for 24 hours.

F) Microarray Wash

- Remove the hybridization chamber from the incubator.
- With gloved fingers, remove the microarray-gasket sandwich from the chamber base by lifting one end and then grasping in the middle of the long sides. Keep the microarray slide numeric barcode facing up as you quickly transfer the sandwich to slide-staining dish #1 (Table 6).
- Gently twist the forceps to separate the slides.
- Let the gasket slide drop to the bottom of the staining dish.
- Remove the microarray slide and quickly put into slide rack in the slide-staining dish #2 containing Agilent Olig aCGH/ChIP-on-Chip Wash Buffer 1 at room temperature (Table 6). Wash microarray slide for 5 minutes
- Transfer slide rack to slide-staining dish #3, which contains Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 2 at 37°C and wash microarray slide for 1 minute.
- Slowly remove the slide rack trying to minimize droplets on the slides.
- Scan slides immediately to minimize the impact of environmental oxidants on signal intensities.

	Dish	Wash Buffer	Temperature	Time
Disassembly	#1	Agilent Oligo aCGH/Wash Buffer 1	Room T	
1st wash	#2	Agilent Oligo aCGH Wash Buffer 1	Room T	5 min
2nd wash	#3	Agilent Oligo aCGH Wash Buffer 2	37°C	1 min

Table 6 : Wash conditions.

G) Microarray Scanning and Analysis

Put slides in a slide holder:

- Carefully place the end of the slide without the barcode label onto the slide ledge.
- Gently lower the microarray slide into the slide holder. Make sure that the active microarray surface faces up, toward the slide cover.
- Close the plastic slide cover, pushing on the tab end until you hear it click.

Scan the microarray slides: Agilent SureScan Microarray Scanner.

- Put assembled slide holders into the scanner cassette.
- Select Protocol AgilentG3_CGH for G3 microarrays. Select Protocol AgilentHD_CGH for HD microarrays.
- Verify that the Scanner status in the main window says Scanner Ready.
- Click Start Scan.

H) Microarray Analysis

- After scanning is completed, extract features and analyze.
- Feature extraction is the process by which data is extracted from the scanned microarray image (.tif) and translated into log ratios, allowing researchers to identify aberrations in their samples through Agilent CytoGenomics.

3.4 Web-based databases

Many useful databases have been developed to help in the interpretation of array-CGH results.

We used UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway), the Database of Genomic Variants (DGV, http://dgv.tcag.ca/dgv/app/home), the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER, https://decipher.sanger.ac.uk/) and the International Standards for Cytogenomic Arrays consortium (ISCA, http://clinicalgenome.org/).

3.5 Customized array CGH

A customized array platform (Agilent) was designed through the Agilent e-array tool (https://earray.chem.agilent.com/earray/). This platform includes a high coverage of the region Xq13q21 (at least one probe every 5 kb), a lower density coverage for all the rest of the chromosome X (at least one probe every 10 kb) and a coverage of other genomic regions already associated with GH-deficiency and pituitary development (Table 7).

X chromosome	about one probe every 10 kb along entire X chromosome
Xq13q21	chrX:72,300,000-78,500,000 (about one probe every 5 kb)
GH1 cluster	chr17:59,000,000-60,000,000
GHRHR	chr7:30,400,000-31,600,000
GHRH	chr20:35,000,000-35,800,000
FGF8	chr10:103,000,000-104,000,000
LHX4	chr1:178,000,000-179,000,000
HESX1	chr3:56,800,000-57,800,000
PROP1	chr5:177,000,000-177,700,000
LHX3	chr9:137,800,000-138,800,000
POU1F1	chr3:86,800,000-87,800,000
SOX3	chrX:139,000,000-140,000,000
SOX4	chr6:21,000,000-22,000,000
PITX1	chr5:133,800,000-134,800,000
BMP4	chr14:53,000,000-54,000,000
OTX2	chr14:55,800,000-56,800,000
15q24	chr15:72,000,000-76,000,000

Table 7: Genomic regions included in customized array platform. For each GHD gene a surrounding interval of about 1 Mb has been considered to include regulatory elements outside the coding region.

3.6 MLPA and PCR analysis

MLPA analysis

For each deletion/duplication a synthetic MLPA probes have been specifically designed, according to the provider's protocol ("Designing synthetic MLPA probes", MRC-Holland b.v. Version 14, Last update 19-12-2014) to cover the region detected by aCGH. The protocol analysis and the reference probes have been taken by the commercial SALSA MLPA probemix P200-A1 Human DNA Reference-1 (MRC-Holland, Amsterdam, Netherlands; http://www.mrc-holland.com).

PCR analysis

The deletion detected on X-chromosome through the custom aCGH was validated with a PCR experiment. Two couple of primers were designed inside and outside the deleted region.

4. PART A

aCGH analysis in Intellectual Disability patients

Part A

4.1 Intellectual Disability

The World Health Organization defines Intellectual disability (ID) as a condition of arrested or incomplete development of the brain with the onset occurring before 18 years of age, especially characterized by impairment of skills that are manifested during the developmental period and which contribute to the overall level of intelligence (i.e. cognitive, language, motor, and social abilities).

ID is a common neurodevelopmental disorder that is characterized by an intelligence quotient (IQ) of 70 or below, and deficits in at least two behaviors related to adaptive functioning.

It is estimated that ID affects 2−3% of the paediatric population and is present in every social class and culture.

ID is a condition that involves motor function, cognitive ability, language or combinations and may become evident during infancy or early childhood as developmental delay, but it is more precisely diagnosed during school years. ID is a common problem in child health and a frequent reason for referral to a neuropsychiatrist [33,34].

ID is divided into 5 categories based on IQ: mild, moderate, severe, profound and unable to classify (DSMIV). However, epidemiological studies often use a simplified classification, dividing their subjects into mild ID (IQ50-70) and severe ID (IQ<50).

While the prevalence of severe ID is relatively stable, the prevalence of mild ID is variable and often depends heavily on external environmental factors (i.e. the level of maternal education, access to education/opportunity and access to healthcare). Study design, age of subjects, and the catchment population for the studies may also contribute to the variability seen across mild ID prevalence studies.

In addition to categorize based on severity/IQ level, ID can also be grouped into syndromic intellectual disability (SID) and non-syndromic intellectual disability (NS-ID).

In SID, patients present with one or multiple clinical features or co-morbidities in addition to ID. While S-ID has a clear definition, there is controversy over the classification of NS-ID. Traditionally, NS-ID has been defined by the presence of intellectual disability as the sole clinical feature. However, it has been a challenge to rule out the presence of more subtle neurological anomalies and psychiatric disorders in these patients, as they may be less apparent, or difficult to diagnose due to the cognitive impairment. Thus the distinction between S-ID and NS-ID is often unclear [35].

Causes of ID

ID can be caused by environmental and/or genetic factors. However, for up to 60% of cases, there is no identifiable cause. Environmental exposure to certain teratogens, viruses or radiation can cause ID, as can severe head trauma or injury causing lack of oxygen to the brain. While these factors explain some cases of NS-ID, it is also important to consider genetic etiology. Genetic causes of ID are thought to be present in 25–50% of cases, although this number increases proportionally with severity.

Chromosomal abnormalities have been reported in ID, with a broad range of prevalence, and many different types of aberrations have been identified. Autosomal trisomies that are compatible with human viability and aneuploidies of the X-chromosome almost always result in some degree of ID as part of a syndrome, as illustrated by trisomy 21, or Down Syndrome—the most common genetic form of ID. Additionally, pathogenic CNV have been found to be associated with ID in a large number of studies, and will likely contribute to the discovery of many ID causing genes in the future.

Over the past 15 years many single gene causes of NSID have been identified. Many of these NS-ID genes may also cause S-ID, autism or other neurodevelopmental phenotypes, making it likely that other genetic modifiers or environmental factors may be involved in disease etiology. It also important the genotype/phenotype comparisons, which are often difficult to elucidate. There is also a possibility that some instances of NS-ID are multifactorial, with more than one gene contributing to disease in an individual, however this has not been well studied [35].

However, there is considerable discrepancy in estimates as to the percentage of cases in which an etiologic diagnosis can be established. The report of a National Institutes of Health (NIH) Consensus Conference held in 1995 concluded "that a diagnosis or cause of the mental retardation can be identified in 40–60% of all patients undergoing evaluation", although one study reported a specific genetic or syndrome diagnosis in only 19.9% of cases [33,34].

4.1.1 The clinical utility of aCGH in Intellectual Disability

The traditional cytogenetic analysis of ID with banding techniques detects large-scale chromosomal abnormalities and has been an important diagnostic tool in the past for the identification of some causes of ID. However, conventional cytogenetic analysis cannot reliably detect rearrangements of genomic segments smaller than 3−5 million base pairs (Mb). The introduction of CGH array allowed the detection of smaller micro-deletions or duplications, thus increasing the number of ID cases in which a molecular diagnosis has been identified.

Microarray-based genomic copy-number analysis is now a commonly requested clinical genetic test for this patient population and is offered under various names, such as ''chromosomal microarray'' (CMA) and ''molecular karyotyping". A specific genetic diagnosis facilitates comprehensive medical care and accurate recurrence risk counseling for the family. A retrospective analysis of 36,325 patients with ID estimated that a pathogenic abnormality could be detected in ~19% of unselected ID patients via genome-wide array-based assays with a 30–70 kb median probe spacing [36].

Array CGH enables a 100-1000 fold higher resolution with respect to standard karyotype (that detect imbalances larger than 5–10 Mb), but still fails to reveal any clinically relevant CNV in the majority of patients tested. Thus, a delineation of clinical criteria allowing for separation of patients with high probability of significant CNV would be desirable [37].

4.2 Recruitment of the patients

For this study a total of 53 pediatric patients (29 males and 24 females) have been recruited among those attending the Unit of Pediatrics, Neurology and Child Neuropsychiatry of the Hospital of Novara. All of them have been evaluated by neuropsychiatrists, pediatricians and genetists and a specific clinically recognizable syndrome was excluded in all. Patients were selected among those having intellectual disability without known etiology, in association with one or more major congenital anomalies, or dysmorphisms or both. All the patients had a normal karyotype on G banding analysis.

4.3 Results and discussion

Among 53 individuals, 16 chromosomal rearrangements not reported in the DGV as common CNV, were detected in 16 individuals (30.2%) with a male:female ratio 2.2:1. In table 8 is reported a summary of the phenotypes and the imbalances detected in these patients. Among them, seven deletions (44%) and nine duplications (56%) have been found. The estimated maximum sizes of the genomic losses and gains ranged from 94 Kb to about 9 Mb.

For each deletion/duplication a probe was designed within a gene included in the region. The analysis was performed on the patient to confirm the presence of the rearrangements. In all the selected patients the MLPA confirmed the imbalances detected by aCGH. Thus, the analysis was extended to the available parents to determine if it was inherited or arose de novo. For 9 of the 16 patients, we collected the parent's DNA. One patient (# 7851) was adopted and for the other 6 we are still waiting for the parent's blood samples. Two unbalances arose *de novo*.

Among these 16 imbalances we excluded the involvement of the detected rearrangement in two patients as it was inherited from a healthy parent, smaller than 300Kb and doesn't contain candidate genes. It is likely that these two imbalances are rare benign rearrangements (# 7280, # 7591, Table 8). Thus, 14 of 53 patients (26%) carry potentially clinically relevant chromosomal imbalances.

I here report six cases among the most relevant for the research chosen on the basis of the following criteria:

1) patients with rearrangements that have not been previously described in literature; in some cases the rearranged region includes genes strongly correlated to the phenotype

2) rearrangements in deleted/duplicated regions already described but that narrow the "minimal critical region" associated to phenotype

Case	sex M=0; F=1	Phenotype[a]	Array-CGH results[b]	Size (bp)	Principal genes included in the rearrangement	Inheritance
# 7277	0	Moderate ID, psychomotor retardation, facial dysmorphisms	Dup. 1q21.1 - q21.2	1,221,964	*GJA5, GJA8*	n. d.
# 7285	1	ID, facial dysmorphisms	Dup. 12q13.13 - q14.1	9,273,199	*SUOX, KIF5A, CDK4, DDIT3, RDH5, AAAS, ERBB3,TAC3,MIP, CYP27B1, AMHR2, MYO1A, TSFM, ITGA7*	n. d.
# 7323	0	ID, congenital hypothyroidism, abnormal isoleucine metabolism	Del. 15q13.2 - q13.3	1,899,633	*CHRNA7, TRPM1*	Maternal
# 7283	1	facial dysmorphisms, hirsutism, mild ID, clumsiness, an epileptic seizure	Dup. 6q27	257,123	*PDE10A, C6orf118*	*De novo*
# 7280	0	Cryptogenetic focal epilepsy	Dup. 19p13.11 - p12	273,297	-	Paternal
# 7281	0	language delay, facial dysmorphisms, ID	Dup. Xq13.1 - q13.2	288,428	*HDAC8, PHKA1*	Maternal
# 7278	0	ID, Epilepsy	Del. Xq24	94,499	*SLC25A43, SLC25A5*	Maternal
# 7616	1	focal epilepsy,ID	Dup. 16p11.2	523,388	*ALDOA*	*De novo*
# 7591	1	cranial and dysmorphisms	Del. 20p12.1	98,891	-	Paternal
# 7789	0	Epilepsy, ID	Dup. 8p23.2	517,494	CSMD1	n. d.
# 7791	0	Epilepsy, ID, Dermatitis	Dup. 13q12.3	982,261	KATNAL1, *HMGB1 ALOX5AP*	*De novo*
# 7731	0	Autism, mild dysmorphisms	Del. 5q21.1	1,832,254	ST8SIA4, SLCO4C1, SLCO6A1	n. d.
# 7786	0	ID, autism spectrum disorder	Del. Xq27.2 - q27.3	519,633	SPANXN4	n. d.
# 7850	0	Hypotonia, psychomotor retardation, facial dysmorphisms, SS,GHD	Del. Xq21.1	5,766,170	*TBX22, BRW3, POU3F4, HMGN5, SH3BGRL*	Maternal
# 7851	0	chronic fatigue, persistent pain, hypotension, behavioural alterations	Del. 2q11.2	1,487,431	*ZAP70, CNGA3*	n. d.
# 7852	1	SS,GHD,scoliosis,uterus didelphys	Dup. 15q25.2	1,598,682	CPEB1, AP3B2 LOC338963	n. d.

[a]ID= Intellectual Disabilit, SS=short stature, GHD= growth hormone deficiency
[b]del=deletion; dup= duplication

Table 8: Phenotypes description and chromosomal imbalances detected in 16 patients through a-CGH. The most relevant cases are highlighted in yellow.

Patient # 7323

aCGH results:

arr15q13.2-q13.3 (30.322.138x2, 31.014.508-32.914.140x1, 32.928.004x2)

At the time of diagnosis the patient was a 13-year-old boy, born from consanguineous healthy parents of Italian origin. He presented intellectual disability, motor delay, congenital hypothyroidism, abnormal isoleucine metabolism and an IQ of 42 (WISC-III).

The pregnancy was normal until the IV month, then the mother was treated with drugs for hypertension. At birth the weight was 3150g and Apgar scores were 7/5. The standard cytogenetic analysis revealed a normal male karyotype (**46**,XY). The patient was negative for fragile-X syndrome analysis.

The brother of the proband at the time of neuropsychiatric evaluation was a 16 years old. He suffered from nocturnal epilepsy and he showed an IQ of 71 (WAIS-R)

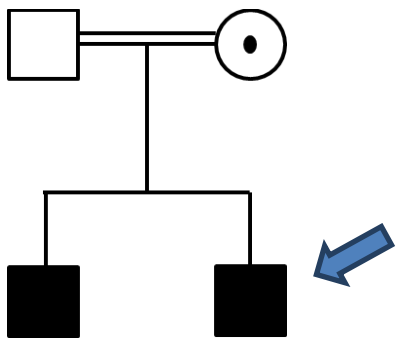

Fig. 4: Pedigree of the family; the proband is indicated by an arrow.

Array CGH revealed the presence of a deletion of 1,9 Mb located on 15q13.2-q13.3 chromosome (Fig. 6). The same deletion was present in the brother and was inherited from their unaffected mother (Fig. 4).

This region contains 12 genes: *TRPM1, KLF13, OTUD7A, CHRNA7, ARHGAP11A, FAN1, MTMR10, MIR211, FAM7A3, FAM7A1, FAM7A2, LOC100288615*.

The 15q13.3 microdeletion syndrome (OMIM #612001) is a recently described disorder characterized by a high degree of variable expressivity. Individuals with this deletion often manifest a variety of phenotypic features including intellectual disability, seizures, autism, behavioural issues, and psychiatric conditions such as schizophrenia or bipolar disorder. This deletion is inherited in approximately 75% of cases and has also been found in apparently normal individuals, consistent with incomplete penetrance.

Most patients with 15q13.3 microdeletion carry an approximately 2 Mb deletion, including 1.5 Mb of unique sequence with the remainder consisting of segmental duplications. The proximal breakpoint of the largest deletion is contiguous with breakpoint 3 (BP3) of the Prader–Willi/Angelman Syndrome critical region extending 3.95 Mb distally to BP5 (Fig. 5).

The smaller 1.5-Mb deletion has a proximal breakpoint within the larger deletion (BP4) and shares the same distal BP5. This recurrent 1.5-Mb deletion contains six genes:*MRMR15, MTMR10, TRPM1, KLF13, OTUD7A*, and *CHRNA7* (Fig. 6) [38,39].

Among these, *CHRNA7* have been previously reported as the strongest candidate gene responsible for the clinical findings associated with 15q13.3 microdeletion syndrome [38,39].

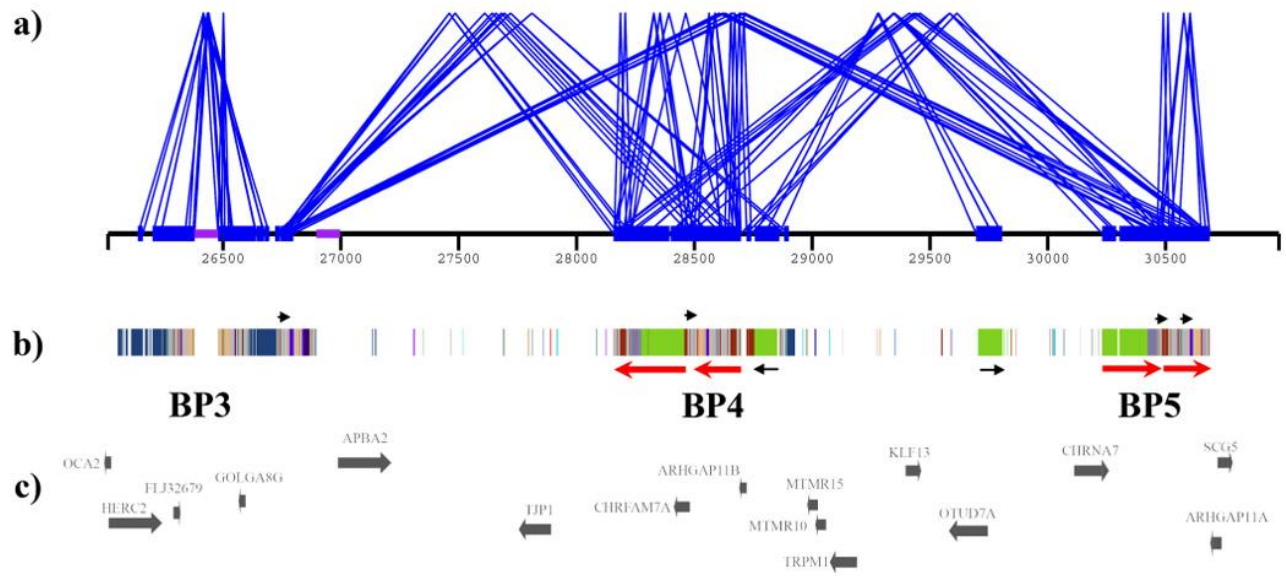

Fig. 5: Duplication architecture of 15q13 breakpoint regions [38].

The search in the Decipher and ISCA databases for similar micro-rearrangements at 15q13.3 led to the identification of nine deletions: 5 in patients with syndromic intellectual disability and developmental delay (DECIPHER 259348, 272774, 288876, 288071, 288865 Fig. 6) and 4 in cases with no further phenotype details (DECIPHER 283651, 285668,283927 and ISCA nssv577499, Fig. 6).

Recently a limited number of smaller deletions encompassing only *CHRNA7* in patients with features of 15q13.3 microdeletion syndrome have been published [40,41], pointing to *CHRNA7* as the gene mainly responsible for the phenotype of these patients (Fig 6).

Fig.6: A schematic overview of the 15q13.2-q13.3 genomic region. The deletions are represented by red bars.

The ***CHRNA7*** gene (cholinergic receptor, neuronal nicotinic, alpha polypeptide 7, and OMIM #118511) contains 10 exons and spans approximately 75 kb on chromosome 15q13.3. It is a member of the nicotinic acetylcholine receptor superfamily of ligand-gated ion channels that mediate signal transmission at synapses. The CHRNA7 protein forms a homopentameric synaptic ion channel that is highly expressed in the brain.

The here reported family show an extremely variable expressivity and incomplete penetrance of the 15q13.3 microdeletion syndrome. The proband is severely affected with an IQ of 42 and dismorphism. The brother is mild affected and the mother is healthy.

It is interesting that the brother of the proband manifests a phenotypic features that has never been described in the 15q13.3 microdeletion syndrome. Thus, the role of *CHRNA7* and possibly of other genes in the region has to be investigated.

Patient # 7283

aCGH results:

arr6q27 (165.443.883x2, 165.562.275-165.819.397x3, 165.886.755x2)

The patient was born after an uneventful pregnancy by cesarean section. The birth weight was 3140g and Apgar scores were 9/10. She is now 7 years old and present facial dysmorphisms, hirsutism, mild intellectual disability, clumsiness and a previous epileptic seizure. The mother is affected by epilepsy and in the maternal lineage the parents reported three cases not better specified of intellectual disability (II3,III2,IV1, Fig. 7).

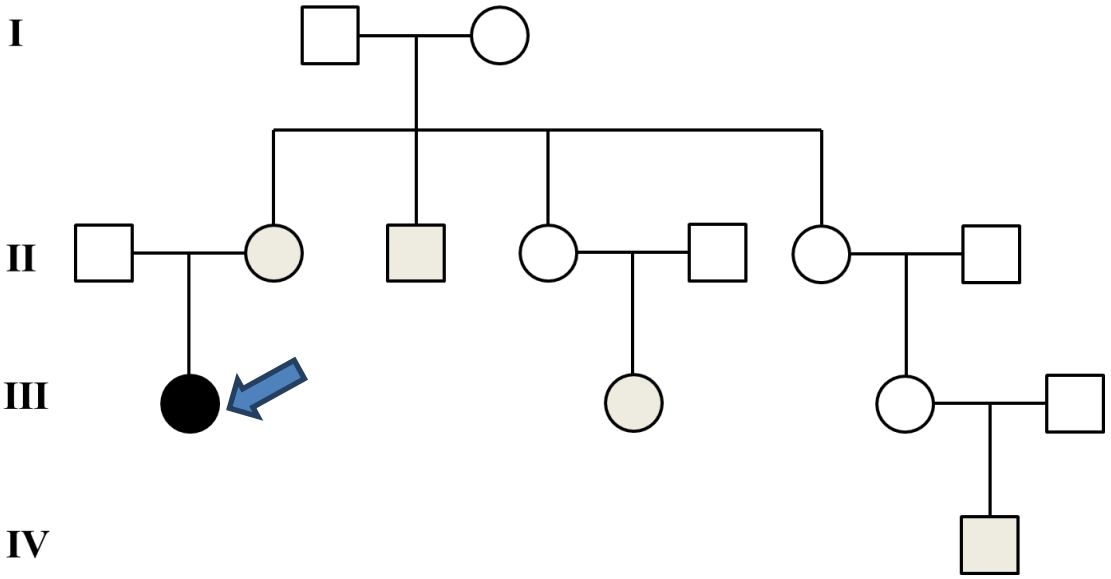

Fig.7: Pedigree of the family, the proband is indicated by an arrow.

Array CGH revealed the presence of a duplication of 257,123 bp located on 6q27 chromosome (Fig. 8). The pedigree was suggestive of a dominant heritance with incomplete penetrance (Fig. 7). However the analysis of the parents shows that the duplication was *de novo* as it was absent in both the parents.

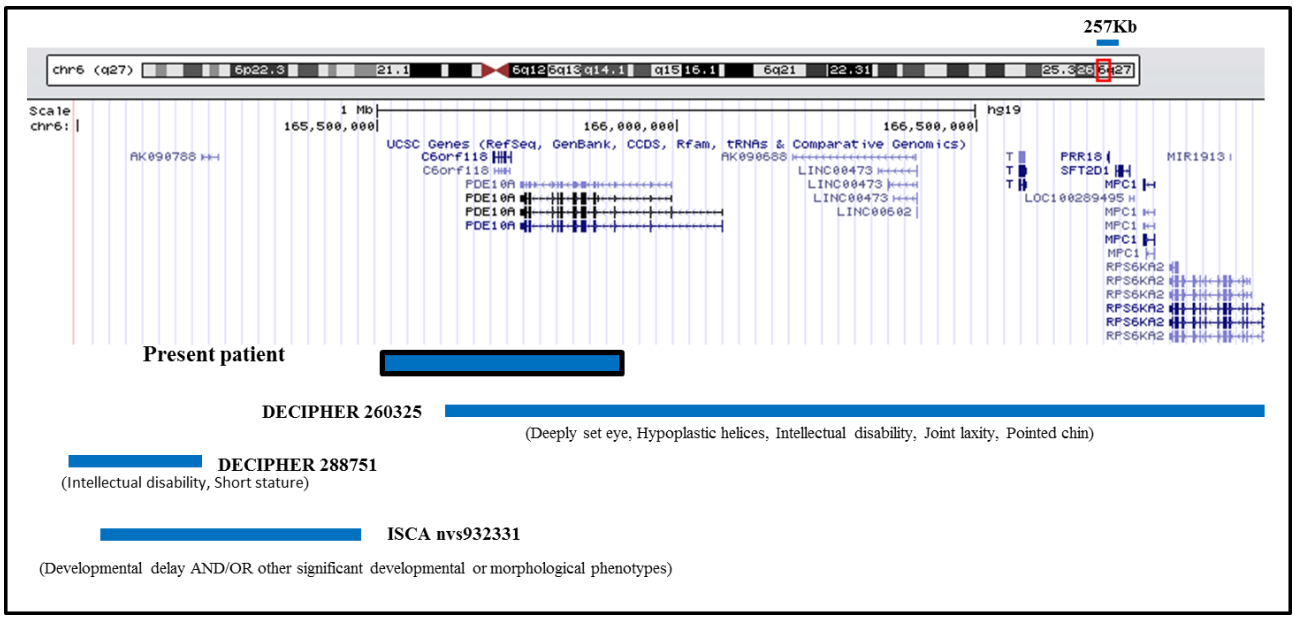

Fig.8: A schematic overview of the 6q27 genomic region. The duplications are represented by blue bars.

This region contains 2 genes: *PDE10A* and *C6orf118* (Fig. 8). To our knowledge no clear genotype-phenotype correlation has been previously described in literature.

The ***PDE10A*** gene (phosphodiesterase 10A) contains 24 exons and spans approximately 200 kb on chromosome 6q27. The 89-kDA protein encoded by this gene belongs to the cyclic nucleotide phosphodiesterase family. It plays a role in signal transduction by regulating the intracellular concentration of cyclic nucleotides.

This protein can hydrolyze both cAMP and cGMP to the corresponding nucleoside 5' monophosphate, but has higher affinity for cAMP, and is more efficient with cAMP as substrate. There are 11 different families of PDEs that vary in their substrate specificity, kinetic properties, mode of regulation, intracellular localization, and tissue expression patterns.

The most striking characteristic of *PDE10A* is its restricted distribution in body: is highly expressed in the brain and has limited expression in peripheral tissues. In the brain, *PDE10A* is highly concentrated in the basal ganglia including MSNs of the striatum. Thus, influence of *PDE10A* on dopamine signaling has been extensively studied [42,43,44].

There is a growing amount of evidence to suggest that PDEs play a critical role in modulating dopamine signaling and selective inhibitors of these enzymes are currently being explored as novel therapeutics to treat schizophrenia, Huntington's disease, cognitive disorders and Parkinson's disease [44].

The ***C6orf118*** gene (chromosome 6 open reading frame 118) contains 9 exons and spans approximately 30 kb on chromosome 6q27. The protein encoded by this gene is uncharacterized. It is highly conserved among species, but there are no information on its function.

In conclusion the family showed a typical dominant with incomplete penetrance inheritance. CGH revealed a duplication encompassing a gene highly expressed in the brain, that represents a good candidate for the disease in the family. However, this is a *de novo* duplication suggesting a causative role in the proband.

Patient # 7281

aCGH results:

arrXq13.1-q13.2 (71.523.738x1, 71.550.297-71.838.724x2, 71.872.520x1)

The patient (II1), a 10-year-old boy, was the second child of non-consanguineous healthy parents. He started walking at the age of 15 months and he manifested language delay. He presented facial dysmorphisms, and intellectual disability. Karyotype and DNA analysis for Fragile X Syndrome were normal.

The brother of the proband (II2) presents a global developmental delay, motor delay and facial dysmorphisms. The two sister (II3,4) are healthy (Fig. 9).

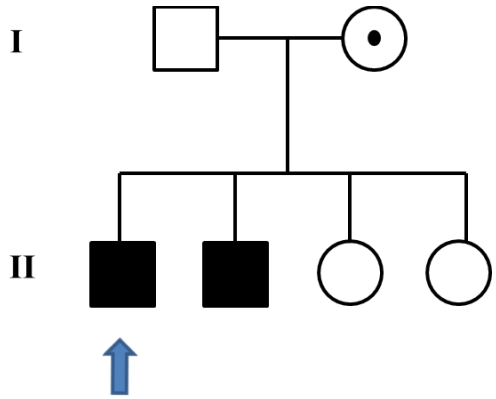

Fig 9: Pedigree of the family, the proband is indicated by an arrow.

Array CGH revealed the presence of a duplication of 290 kb located on Xq13.1-q13.2 chromosome (Fig. 10). The same duplication was present in the brother and was inherited from the heterozygous mother (Fig. 9). This region contains 2 genes: *HDAC8* and *PHKA1*.

The *HDAC8* gene (Histone Deacetylase 8) contains 8 exons and spans approximately 238,5 kb on chromosome Xq13.1-q13.2.

The protein encoded by this gene belongs to class I of the histone deacetylase family. It catalyzes the deacetylation of lysine residues in the histone N-terminal tails and represses transcription in large multiprotein complexes with transcriptional co-repressors.

HDACs are evolutionary conserved enzymes that remove acetyl groups from lysine residues of histones and other important cellular non-histone proteins.

They play a major role in epigenetic gene silencing during development, for example, X-chromosome inactivation, cell differentiation and morphogenesis. Epigenetic chromatin remodelling and DNA modifications represent the central mechanisms for regulation of gene expression during brain development.
Thus, based on the function of HDACs, increasing numbers of HDAC inhibitors are presently being developed for treatment of neuropsychiatric diseases. HDACs were already associated with disorders involving human brain development [45].

Patients with mutations in *HDAC8* have features suggestive of Cornelia de Lange syndrome (CdLS). CdLS in most of the cases is a congenital autosomal dominant caused by mutations in *NIPBL, SMC3* and *RAD21* genes. There are also cases of X-linked dominant form of the disease caused by mutations in *SMC1A* and *HDAC8* genes, characterized by facial dysmorphism, pre- and post-natal growth retardation, developmental delay, intellectual disability, and multiorgan involvement.

Mutations in *HDAC8* cause a small number of CdLS cases, accounting for ~ 4% of mutations in patients. Reported mutations include missense (the most frequent type), chromosomal microdeletions or intragenic microduplications, nonsense and splice site.
A recent study reported a cohort of 38 individuals with an emerging spectrum of features caused by *HDAC8* mutations. For several individuals, the diagnosis of CdLS was not considered prior to genomic testing. Many cases are heterozygous females, each with marked skewing of X-inactivation in peripheral blood DNA. We also identified eight hemizygous males who are more severely affected. The craniofacial appearance caused by *HDAC8* mutations overlaps that of typical CdLS but often displays delayed anterior fontanelle closure, ocular hypertelorism, hooding of the eyelids, a broader nose and dental anomalies, which may be useful discriminating features [45,46,47].

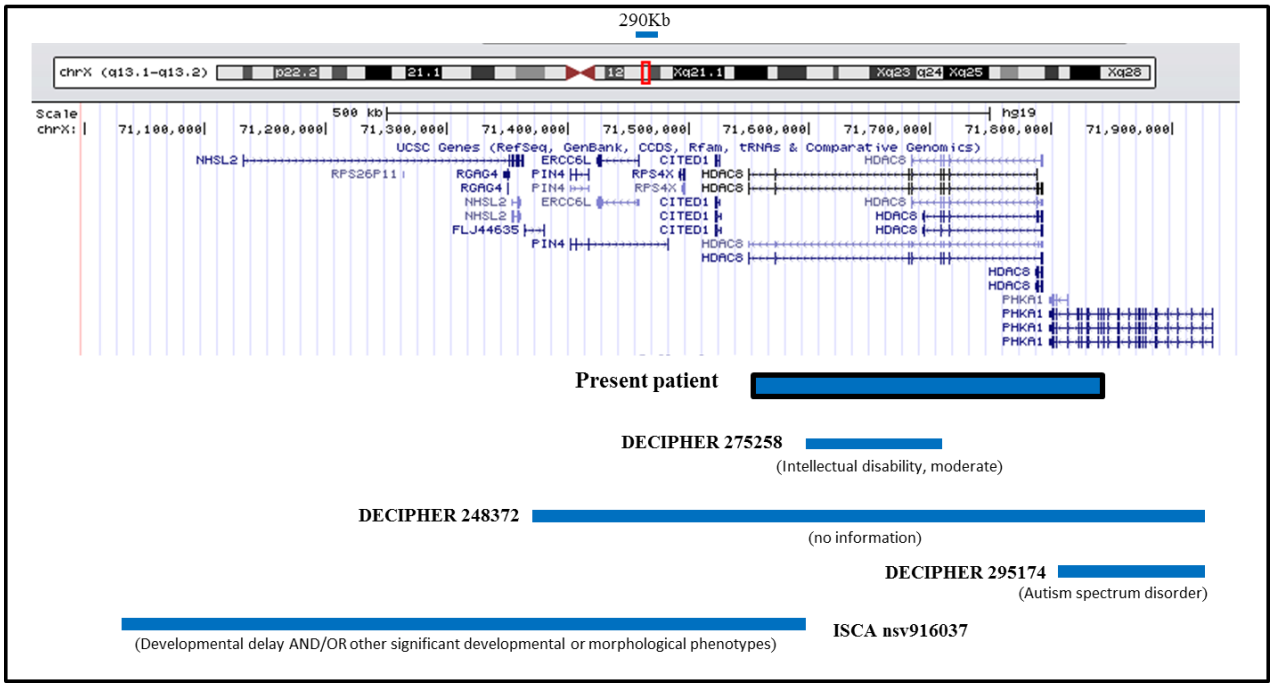

Fig 10: A schematic overview of the Xq13.1-q13.2 genomic region. The duplications are represented by blue bars.

The ***PHKA1*** gene (phosphorylase kinase, alpha 1 muscle) contains 32 exons and spans approximately 135 kb on chromosome Xq13.1-q13.2.

Phosphorylase kinase is a polymer of 16 subunits, four each of α, β, γ and δ. The α subunit includes the skeletal muscle and hepatic isoforms, and the skeletal muscle isoform is encoded by this gene. Mutations in this gene cause Glycogenosis type VIII, also known as muscle Phosphorylase b Kinase (PHK) deficiency.

Interestingly it has been reported a patient with a new mutation in this gene who presented with PHK deficiency, cognitive impairment, but no overt myopathy. This report supports the concept that PHK deficiency is a mild metabolic myopathy and suggests that PHK mutations may interfere with normal brain function [48].

CGH revealed in this patient a duplication encompassing entirely *HDAC8,* and a part of the *PHKA1* gene. Both the genes could be good candidates for causing the proband phenotype. However patients with *HDAC8* mutations show a phenotype quite different from that of our 2 patients.

PHKA1 might be more related to the clinical features of the patients. It is difficult to predict the effect of a partial duplication of the gene. Theoretically if the duplication is a perfect tandem duplication, patients carry at least one intact copy of the gene. However it might be inverted or not perfect in tandem and consequently it might interrupt the gene. A molecular study is needed to characterize the duplicated region and to asses a role to this rearrangement in the family.

Patient # 7278

aCGH results:

arrXq24 (118.540.385x1, 118.579.289-118.673.787x0, 118.691.162x1)

The patient, a 16-year-old boy, was the second child of healthy non-consanguineous parents of Italian origin. He was born after an uneventful pregnancy by normal delivery at term. His sister was healthy. In the family there was the brother of the mother with intellectual disability and epilepsy (II3) interpreted as an encephalitis complication.

The first neurologic evaluation reported clumsiness, awkward gait, difficulties with heel walking. At follow-up he showed poor narrative skills and limited vocabulary, sociability disorder, impulsiveness, repetitiveness and fatigue. He was referred to our clinic for clumsiness and learning disability by the child neurologist. Karyotype and DNA analysis for Fragile X Syndrome were normal.

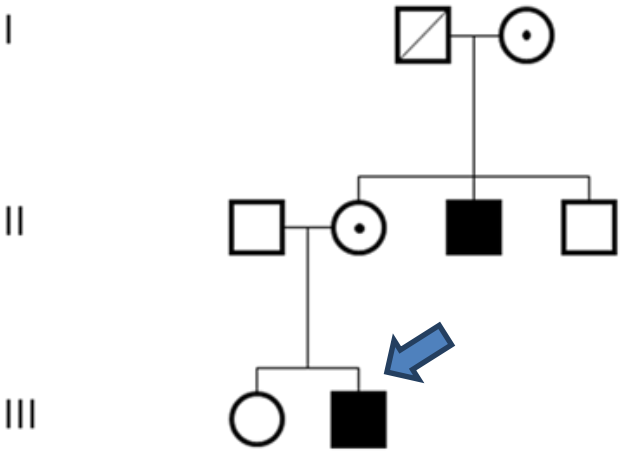

Fig 11: Pedigree of the family, the proband is indicated by an arrow.

Array CGH revealed the presence of a deletion of approximately 95 kb was located on Xq24 chromosome (Fig. 12). The same deletion was present in the mother, maternal uncle and grandmother (Fig. 11). This region contains 3 genes: *SLC25A43, SLC25A5, CXorf56*.

SLC25A5

SLC25A5 (solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), 5) is a member of the mitochondrial carrier subfamily of solute carrier (SLC) protein genes. SLC are membrane proteins that control the exchange of various molecules including amino acids and drugs. The product of *SLC25A5* is ubiquitously expressed and facilitates exchange of ADP from the cytoplasm into the mitochondrial matrix and ATP from the mitochondrial matrix into the cytoplasm.

The protein forms a homodimer embedded in the inner mitochondria membrane. Suppressed expression of this gene has been shown to induce apoptosis and inhibit tumor growth. Diseases associated with this gene include non-syndromic intellectual disability, and syndromic intellectual disability [49].

Microdeletions at Xq24 reported in males always have been associated with a syndromic form of ID due to the loss of *UBE2A* gene [50,51].

Recently, Vandewalle J et al. reported overlapping microdeletions at Xq24 that do not include *UBE2A*, in the affected males of a non-syndromic X chromosome-linked intellectual disability (XLID) family (MRX70) as well as in two sporadic patients, all presenting with non-syndromic ID that could include other minor features. The smallest region of overlap contains two members of the mitochondrial SLC25 family, *SLC25A43* and *SLC25A5* (Fig. 9).

However, the subsequent identification of an intragenic microdeletion including *SLC25A43* but not *SLC25A5* in a healthy boy excluded a role for the first one in cognition, leaving the second as the only candidate gene for the ID. Since this gene is highly expressed in brain and has been reported to be part of the post-synaptic density (PSD), *SLC25A5* is proposed as a novel XLID gene and mitochondrial nucleotide exchange as a novel ID-related pathway [49].

In this study we found the smallest deletion encompassing *SLC25A5* detected in ID patients, confirming this gene as a good candidate for non-syndromic ID (Fig. 12). The causative role of this deletion is confirmed by the segregation in the family, as it has been detected in the affected maternal uncle and it was absent in the health uncle (Fig.11)

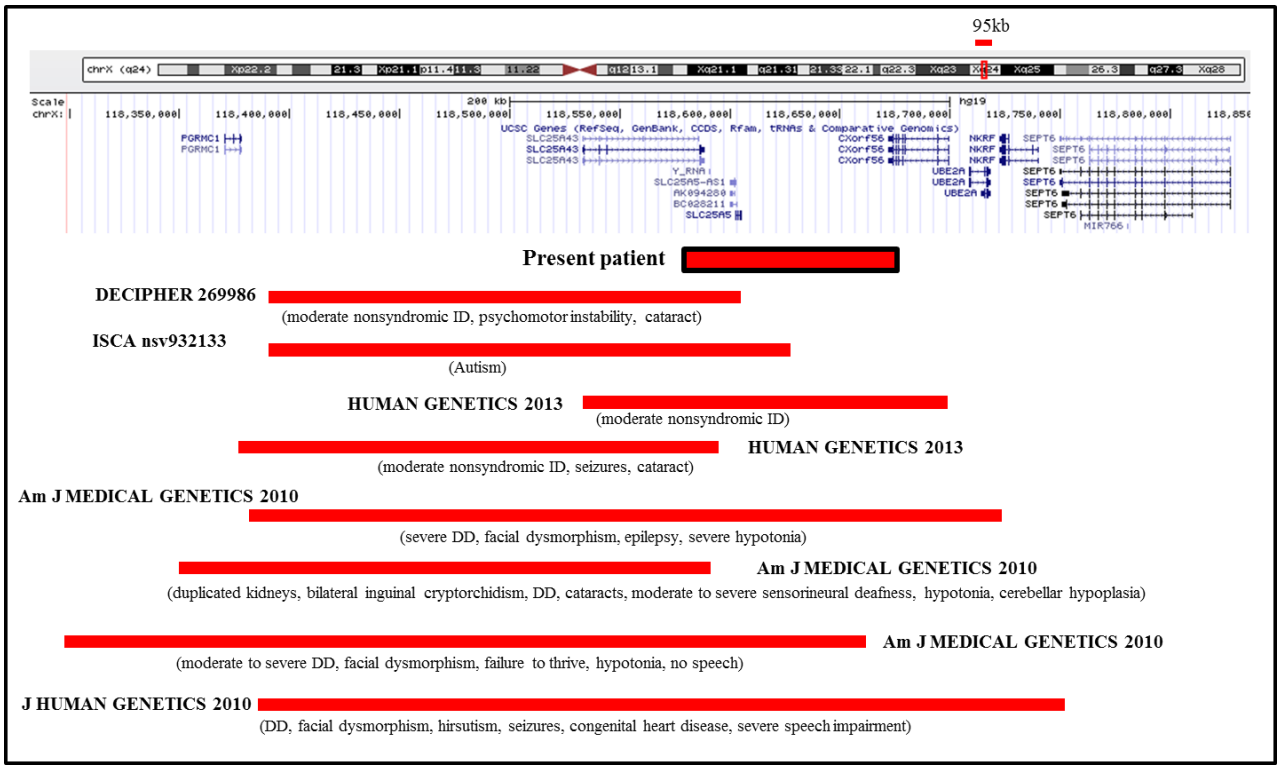

Fig. 12: A schematic overview of the Xq24 genomic region. The deletions are represented by red bars.

Patient # 7791

aCGH results:

arr13q12.3 (30.467.858x2, 30.517.291-31.499.551x3, 31.543.071x2)

The patient was born after an uneventful pregnancy by cesarean section. He started walking at the age of 24 months and is reported psychomotor retardation, speech delay, possible autism, dysmorphisms, dermatitis. He presented partial symptomatic epilepsy.

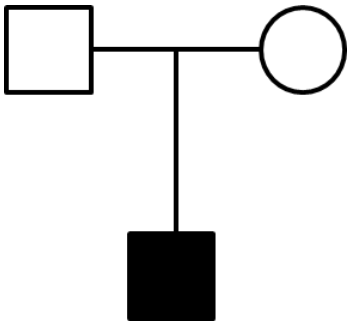

Fig 13: Pedigree of the family

Array CGH revealed the presence of a duplication of approximately 982 kb located on 13q12.3 chromosome (Fig. 14). The duplication was not present in the parents, thus it arose *de novo* (Fig. 13).

This region contains 7 genes: *HMGB1, ALOX5AP, LOC440131, KATNAL1, USPL1, LOC100188949* and *C13orf33*.

KATNAL1

KATNAL1 (KATANIN p60 subunit A-like 1) codes for a recently identified ATPase with high sequence similarity to KATANIN p60 (*KATNA1*), a member of the microtubule-severing enzyme family. Microtubules are dynamic cytoskeletal polymers that play important roles in cell division, morphogenesis, mobility, and signaling. The remodeling of microtubules is mediated by microtubule-severing enzymes**.** The KATANIN protein exists as a heterodimer of a 60-kDa catalytic subunit (p60, KATNA1) and a 80-kDa targeting and regulatory subunit (p80, KATNB1).

Katanin and Katanin-like proteins are highly expressed in the nervous system of different organisms. A neuronal function for katanin was first demonstrated in cultured rat neurons where it was shown to stimulate axonal outgrowth and establish neuronal polarity.

Knowledge of the role of the katanin like proteins is still limited. Recently, however, de novo mutations in *KATNAL2* (KATANIN p60 subunit A-like 2), a close family member of KATNAL1 have been identified in patients with autism spectrum disorders by exome sequencing [52].

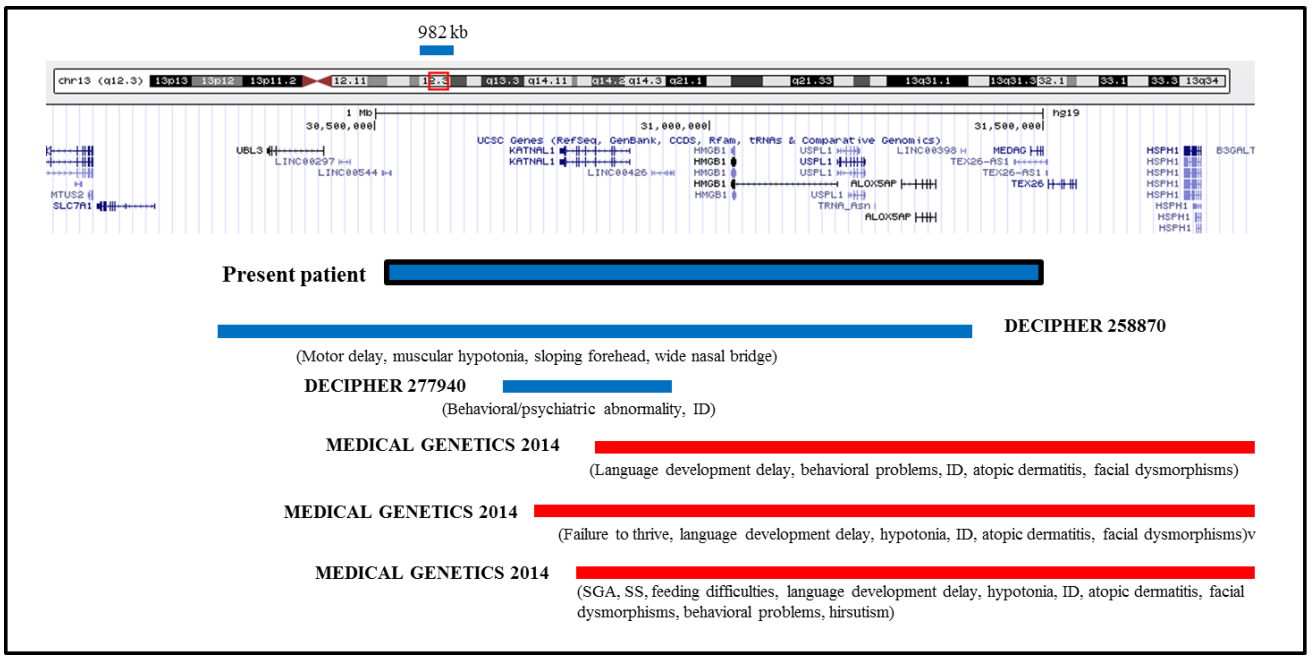

Fig. 14: A schematic overview of the 13q12.3 genomic region. The duplications are represented by blue bars, instead the deletions by red bars.

HMGB1

HMGB1 (High-Mobility Group Box 1) belongs to the HMG-box family. The HMGB proteins are nuclear proteins, comprised of two DNA-binding domains (HMG boxes A and B) and an acidic tail, which seems to have a regulatory function in DNA-binding.

The HMGB1 protein is widely expressed in all tissues of vertebrates and is the most abundant non-histone protein in the nucleus. It is engaged in numerous nuclear processes such as transcription, replication, recombination, and repair.

The importance of HMGB1 is reflected by the fact that homozygous knock-out mice die shortly after birth, very likely due to inefficient activation of glucocorticoid receptor responsive genes.

Apart from its role as a nuclear chromatin-associated protein, HMGB1 is released in the extracellular environment inducing secretion of proinflammatory cytokines. Accordingly, it plays an important role in various conditions, such as sepsis, rheumatoid arthritis, hemorrhagic shock, ischemia-reperfusion injury, and invasive tumor growth.

With regard to the central nervous system, HMGB1 displays a complex temporal and spatial expression pattern during early brain development. It promotes neurite outgrowth and cell migration in rat brain and is critical for forebrain development. In adulthood, its expression is down regulated in the majority of neurons [52].

ALOX5AP

ALOX5AP (Arachidonate 5-Lipoxygenase-Activating Protein) encodes a protein which, with 5-lipoxygenase, is required for leukotriene synthesis. Leukotrienes are arachidonic acid metabolites which have been implicated in various types of inflammatory responses, including asthma, arthritis and psoriasis. This protein localizes to the plasma membrane. Inhibitors of its function impede translocation of 5-lipoxygenase from the cytoplasm to the cell membrane and inhibit 5-lipoxygenase activation.

Our patient showed psychomotor retardation, language deficits, possible autism, dysmorphisms, dermatitis and partial symptomatic epilepsy. The search in the Decipher database for similar micro-rearrangements at 13q12.3 led to the identification of 2 deletions: the first patient with motor delay, muscular hypotonia, sloping forehead, wide nasal bridge (DECIPHER 258870), the second one with behavioural/psychiatric abnormality and ID (DECIPHER 277940, Fig. 14).

A 13q12.3 microdeletion syndrome has recently been described involving a ~300 kb critical region spanning only three genes, namely, *KATNAL1, HMGB1*, and the noncoding *RNA LINC00426*.
The described patients present moderate demonstrated or apparent intellectual disability, postnatal microcephaly, and eczema/atopic dermatitis as the predominant symptoms. In addition, they had pronounced feeding difficulties in early infancy. They displayed similar facial features such as malar flattening, a prominent nose with underdeveloped alae nasi, a smooth philtrum, and a thin vermillion of the upper lip [52].

These three unrelated patients with heterozygous deletions encompassing the region of 13q12.3 chromosome [52], have some characteristics similar to the phenotype of our patient. It is relevant that our patient as the patients with deletions show dermatitis.

The *HMGB1* gene is a good candidate, since the control of gene transcription through epigenetic modification of the chromatin structure in neurons has been shown to play an important role in cognitive processes and in the etiology of ID. Furthermore, *HMGB1* and closely related family members of *KATNAL1* have been implicated in regulation of neurite outgrowth, which is consistent with the observation that many ID-genes are involved in synaptic plasticity. The role of *KATNAL1* in causing ID is also supported by the recent observation that mutations in its family member *KATNAL2* are a cause of nonsyndromic ID/autistic spectrum disorders.

In conclusion our patient, the Decipher cases and the three unrelated patients just described present some similar phenotype characteristics and the same critical region encompassing only three genes. Current knowledge qualifies two of them, the microtubule-severing enzyme KATNAL1 and the chromatin- associated protein HMGB1 as candidate genes for autosomal dominant ID.

The future identification of novel patients carrying small microdeletion/duplication would confirm this hypothesis and help increase our knowledge of the complex molecular pathways causing ID [52].

Patient # 7850

aCGH results:

arrXq21.1 (77.456.818x1,77.489.632-83.255.802x0,83.287.869x1)

The patient was born at the 41st week of gestation by caesarean section due to fetal macrosomia. The birth weight was 4,530 g (>97° percentile), birth length 54 cm (90-97° percentile), and head circumference 40 cm (>97° percentile). Apgar scores were 9 and 10 at 1 and 5 minutes, respectively. He was the only child of non-consanguineous Italian parents (Fig. 15).

At birth, the patient showed cleft palate (plastic surgery of the hard palate was performed at the 5th month of life), hip dysplasia (corrected with retractor), right blepharoptosis (surgically corrected at 7 years of age) and dysmorphic face. At the fourth month of age, the parents noted the presence of severe head and limb hypotonia and a delayed psychomotor development.

At 18 months, bilateral sensorineural hearing loss was diagnosed and a bilateral hearing aid was implanted. The standard cytogenetic analysis revealed a normal male karyotype (46,XY).

The patient proved negative for chromosome 22.11.2 deletion syndrome and fragile-X syndrome analysis.

The family history was negative for intellectual disability. At 3 years of age, because of short stature, a growth hormone (GH) stimulation test with arginine (GH peak: 5 ng/ml; n.v.>10 ng/ml) and with glucagon (GH peak: 2.1 ng/ml; n.v.>10 ng/ml) were performed and led to a diagnosis of complete growth hormone deficiency (GHD). Neuropsychiatric test (Wisc IV test) showed an intellectual disability whit an IQ < 70.

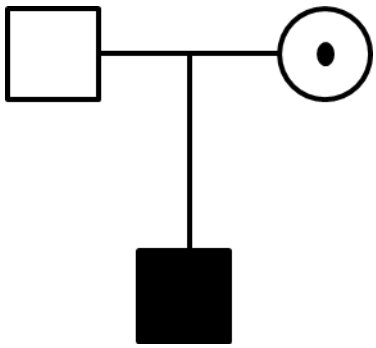

Fig 15: Pedigree of the family

At the age of 18 months, the patient was tested for the presence of causative 150 mutations in the gene encoding connexin 26 (*GJB2*), as he manifested bilateral neurosensorial hearing loss. No mutation associated with hearing impairment was identified in this gene.

The array-CGH analysis, performed at the age of 13 years, revealed the presence of an interstitial deletion at Xq21.1 of approximately 5.8Mb (Fig. 16). The same deletion was present at the heterozygous state in the mother (Fig. 15).

The Xq21.1 deleted region was examined using the Human resource websites (http://www.ncbi.nlm.nih.gov/projects/genome/guide/human) at NCBI and the Archive EnsEMBL (http://www.ensembl.org/info/website/archives/index). The deleted region includes 12 OMIM genes: *CYSLTR1, ZCCHC5, LPAR4, P2RY10, GPR174, ITM2A, TBX22, BRWD3, HMGN5, SH3BGRL, POU3F4, CYLC1* (Fig 13).

Among these, 3 genes (*TBX22*, *BRW3* and *POU3F4*) have been previously reported mutated in well-defined disorders, thus facilitating the phenotype-genotype correlation in our patient. However, none of the genes included in the deletion have been previously correlated to the pituitary hormone deficiencies observed in our patient. The PROP1 gene was thus sequenced in order to detect a gene defect elsewhere that could explain the GH and gonadotropin deficiency, but no mutation was identified.

TBX22

TBX22 (T-box 22) is a member of a phylogenetically conserved family of genes that share a common DNA-binding domain, the T-box. T-box genes encode transcription factors involved in the regulation of developmental processes. Mutations in this gene have been associated with the inherited X-linked disorder, cleft palate with ankyloglossia, and it is believed to play a major role in human palatogenesis.

BRW3

BRW3 (bromodomain and WD repeat domain containing 3) encode for a protein that contains a bromodomain and several WD repeats. It is thought to have a chromatin-modifying function, and may thus play a role in transcription. It also has a role in the regulation of cell morphology and cytoskeletal organization and is required in the control of cell shape. Mutations in this gene cause mental retardation X-linked type 93, which is also referred to as mental retardation X-linked with macrocephaly. This gene is also associated with translocations in patients with B-cell chronic lymphocytic leukemia.

POU3F4

POU3F4 (POU class 3 homeobox 4) encodes a member of the POU-III class of neural transcription factors. This family member plays a role in inner ear development. The protein is thought to be involved in the mediation of epigenetic signals which induce striatal neuron-precursor differentiation. Mutations in this gene cause X chromosome-linked nonsyndromic mixed deafness.

We report a maternally inherited deletion of approximately 5.8 Mb at Xq21.1, in a male subject with developmental delay, dysmorphic facial features, cleft palate, intellectual disability, short stature and hearing loss. The deleted region includes 14 OMIM genes. Among these, *TBX22*, *BRWD3* and *POU3F4* well explain some of the clinical features of this patient.

Mutations in the *TBX22* gene are a well-established cause of X-linked cleft palate with ankyloglossia as well as contributing to the prevalence of isolated cleft palate [53].

The phenotypic spectrum of subjects bearing *TBX22* mutations can vary, even within the same family, from asymptomatic females to males or females with a bifid uvula, a cleft of the soft palate, or a complete cleft of the hard and soft secondary palate, along with ankyloglossia [54].

The deafness causative gene is undoubtedly *POU3F4*, which causes the X-linked neurosensorial deafness DFN3. This defect is associated with either *POU3F4* point mutations or small deletions in a region located at 900 Kb upstream of the gene and disturbing a regulatory element [55].

Male patients show a well characterized phenotype with both stapes fixation and progressive mixed hearing loss [56,57]. Hearing loss was also observed in the mother of our patient as well as in about 40% of the *POU3F4* mutation carrier females. Carrier females usually exhibit a postlingual onset of the hearing impairment that progresses over time [58] with a variable expressivity attributable to variations in the degree of skewing of X inactivation.

Loss of function mutations affecting *BRWD3* are associated with a phenotype including mild to moderate intellectual disability, macrocephaly, dysmorphic facial features, skeletal signs and behavioral disturbance [59,60].

Among these alterations, a partial deletion encompassing 74 Kb and including the 30 last exons of the *BRWD3* gene was reported in a male with high forehead, deep-set eyes, hypertelorism, short palpebral fissures, anteverted nares, downturned 200 corners of the mouth, pointed chin, and skeletal anomalies [61]. This phenotype is similar to that of the patient presented here: macrocephaly, dysmorphic facial features including prominent forehead and abnormal ears, behavioral disturbance, skeletal symptoms like pes planus, and cubitus valgus.

These findings strongly suggest that the lack of *BRWD3* causes intellectual disability, macrocephaly and possibly the skeletal symptoms (pes planus and cubitus valgus) observed in our patient. The search in the Decipher and ISCA databases for similar size deletions at Xq21.1 led to the identification of six deletions in males with syndromic intellectual disability and developmental delay (DECIPHER 287069,253735, 257570,292180 and ISCA nsv533989, nsv530231, Fig. 16) and in one patient with no further phenotype details (DECIPHER 283100, Fig. 16) but none of these exactly corresponds to that of our patient.

However, among those observed in female patients, there are two deletions (DECIPHER 262726 and 265126, Fig. 16) that completely cover that of our patient. The milder associated phenotype includes broad nasal tip, skeletal defects, speech delay, and cleft palate, which are also features of the present patient.

Other remarkable clinical signs of our patient are short stature and multiple pituitary hormone deficiency. The presence of mutations in *PROP1*, the most common genetic cause of combined pituitary hormone deficiency (CPHD) associated with short stature [62], was excluded.

It is interesting that the deletion partially overlaps with regions duplicated in patients with multiple congenital anomalies, developmental delay, GH deficiency and short stature [63,64,65]. This suggests that Xq21.1 might contain either a protein coding gene or a regulatory element involved in pituitary hormone secretion. Within the 5.8 Mb region there is no evidence for a protein coding gene directly involved in pituitary functioning.

However, *ITM2A*, included in the deletion, might be related to the severe short stature as it encodes an integral transmembrane protein involved in early cartilage development [66]. It has been suggested that the expression of *ITM2A* influences the chondrogenic differentiation potential of mesenchymal stem cells in vitro [67]. It might be hypothesized that the absence of *ITM2A* greatly influences the cartilage development with a possible impact on postnatal growth.

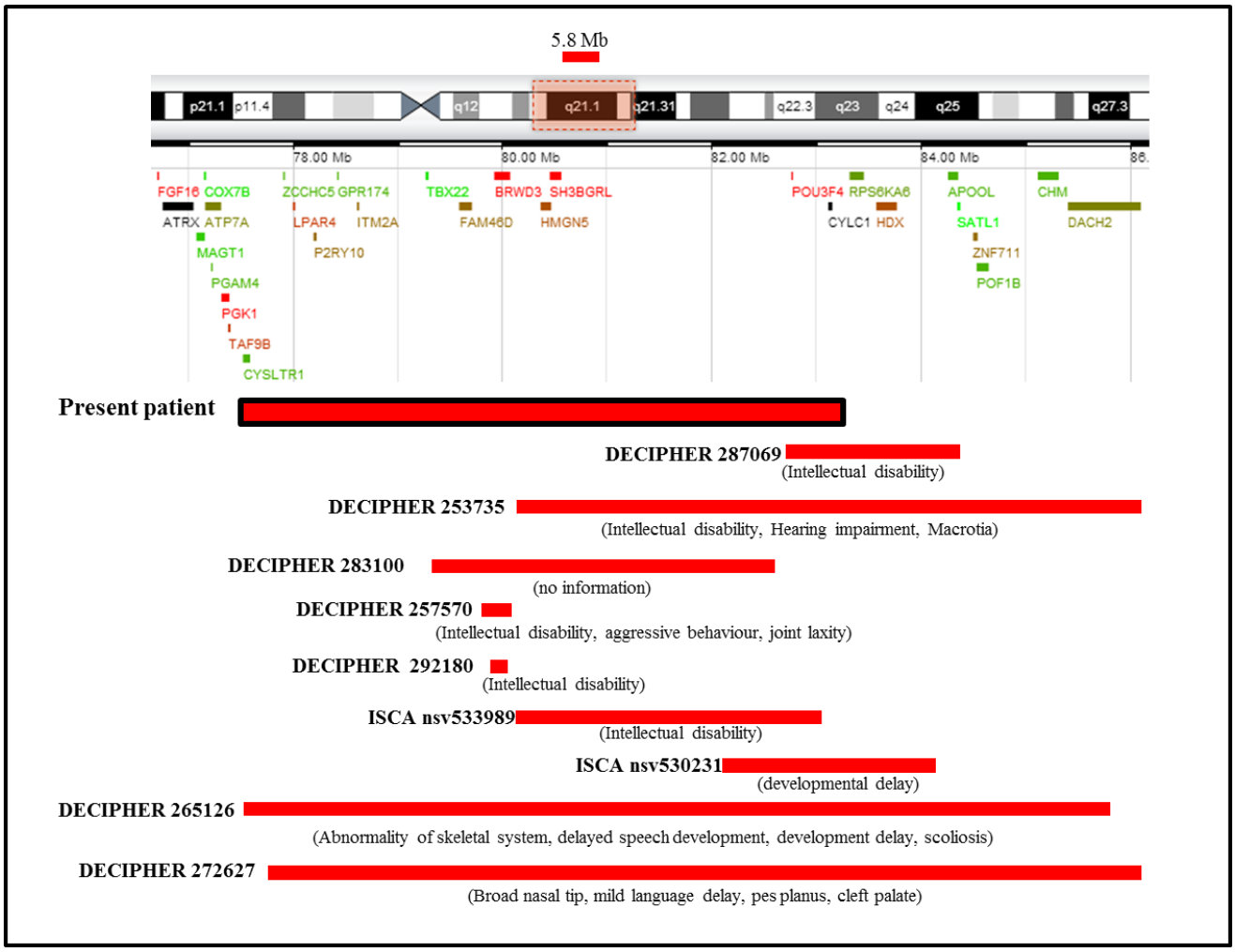

Fig. 16: A schematic overview of the Xq21.1 genomic region. The deletions are represented by red bars.

4.4 Conclusion

Microarray-based comparative genomic hybridization has revolutionized clinical cytogenetic, as it provides a relatively quick method to scan the genome for gains and losses at a resolution higher 100-1000 fold than standard karyotype. It significantly increased the detection resolution of chromosome imbalance from 5–10 Mb (standard karyotype) to 5-50 Kb (41Kb in our study).

Numerous different aCGH platforms have been designed and have been used successfully in the diagnostic setting. In the past few years, these new methodologies led to the identification of novel genomic disorders in patients with ID and/or multiple congenital anomalies.

Among 53 patients, we found 37 individuals (70%) negative for the presence of chromosomal imbalance or carrying common CNVs. In 16 patients (30%) we detected del/dup not reported in DGV as common CNV. Fourteen of them (26%) are potentially clinically relevant chromosomal imbalances as estimate in accordance with the literature.

If we will definitively confirm all these potentially relevant rearrangements, our detection rate is slightly higher (26%) than the average of literature (17-20%) [33,34,35]. Among the chromosomal rearrangements detected, we found a higher number of male patients (69%) than female as the 25% of the total rearrangements encompassed the X-chromosome in accordance to what reported in literature [35].

Obviously the detection rate depends from the patients selection. In our cohort many of the patients were syndromic and underwent previous diagnostic procedures that already excluded a wide range of other causes of ID. All children had routine blood tests and metabolic screening tests, such as amino acids in plasma and organic acids in urine. Analysis of cerebrospinal fluid was performed in 80% of children, and all had conventional karyotyping and fragile-X syndrome analysis before array CGH analysis [37].

5. PART B

Custom aCGH analysis in Multiple Pituitary Hormone Deficiency patients

Part B

5.1 The Xq13q21 region

Growth hormone deficiency (GHD) may occur in isolation, or combined with other anterior and posterior pituitary hormone deficiencies, namely thyrotropin (TSH) prolactin (PRL), adrenocorticotropin (ACTH), luteinizing hormone (LH), follicle-stimulating hormone (FSH), with or without extrapituitary features such as optic nerve hypoplasia and midline forebrain defects. The incidence is between 1 in 3500 and 1 in 10000 live births, with the majority of cases being idiopathic in origin. Familial cases account for 5–30% of cases, suggesting a genetic aetiology for the condition.

However, mutations in signalling molecules and pituitary transcription factors have been identified in a modest proportion of patients with IGHD/CPHD and each gene accounts only for a small percentage of cases. In most cases the genetic causes of hypopituitarism remain to be identified. Their characterization would enable a greater understanding of the disorder, in addition to unravelling the processes inherent in the normal development of the pituitary gland.

A distortion in sex ratio in GHD subjects has been reported in different studies [68,69,70] with a male predominance ranging from 1.5:1 to 3:1. Although some authors suggest a referral bias (i.e., greater concern for short stature in boys), this distorted sex ratio can be explained by the involvement in GHD of a gene or genes on the X chromosome.

This hypothesis is supported by data from the literature describing genomic rearrangements involving chromosome X in patients with short stature. Thirteen independent duplications involving proximal Xq region, most of which characterized by conventional cytogenetics, were reported in males in all cases but two exhibited severe short stature (Table 9).

In general, the abnormal phenotype associated with duplications is attributed to increased gene dosage due to functional disomy for the duplicated region. As a consequence of skewed X-chromosome inactivation resulting in inactivation of the dup(X), most dup(Xq) females are phenotypically normal or may only manifest minor features. In contrast, dup(Xq) males are almost always more severely affected.

Most of the dupXq short stature subjects reported in table 9 were not tested for GH serum levels. Among the four patients investigated for endocrine functions (Table 9), GHD was found in the three short stature subjects [71;72;73] and not in the normal stature patient [74]. MRI analysis in one of the GHD patients [71] showed pituitary morphological abnormalities (empty sella) usually associated with growth hormone deficiency and other impaired endocrinological functions. Thus, the duplication of a gene roughly localized by conventional cytogenetics at Xq13q21 seems to be responsible of impaired growth hormone production.

These data indicate the presence of X-linked genetic factors strongly involved in GHD and that could explain a part, or even a large part, of the at present unexplained genetic component of the disease. The identification of this hypothetical gene/s on the X-chromosome and particularly at the Xq13q21 region could also explain the distorted sex ratio thus representing a major gene in GHD. This might be the first gene causing GHD with a major relevance in the disease as mutations in the genes already known to be involved in GHD explain only a minority of the cases.

Paper	Sparkes 1977	Steinbach 1980	Vejerslev 1985	Schwartz 1986	Cremers 1987	Thode 1988	Muscatelli 1992	Yokoyama 1992	Hou 2003	Cheng 2005	Gabbett 2008	Lugtenberg 2009	Shapira 1997
Cytoband	Xq13 q24	Xq13q 22	Xq13.1q2 1.2	Xq21 -q24	Xq21q22	Xq13.1 q21.2	Xq13q21	Xq13. 3q21. 2	Xq13.2 q21.2	Xq21.3 2q25	Xq21.1q 21.3	Xq13.2q2 1.1	Xq13.2q21. 3
Inherited	na	YES	YES	YES	NO	YES	NO	YES	YES	na	YES (**)	NO	YES (**)
Short stature	YES	YES	YES	na	YES	YES	YES	YES	YES	YES	NO	YES	YES
GH deficiency	na	na	na	na	na	no	na	YES	YES	na	no	na	YES
IGF-1 deficiency	na	na	na	na	na	YES	na	YES	YES	na	na	na	na
Hypoplastic Genitalia	YES	YES	YES	YES	YES	YES	na	YES	YES	YES	YES	YES	YES
Hypothyroidism	na	na	na	na	na	na	na	na	na	YES	na	na	YES
IUGR	YES	na	na	na	na	na	na	YES	na	YES	na	na	na
Hypotonia	na	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
MR/Psycho motor delay	na	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
MRI	na	na	na	na	Severe disorder of myelinization	na	na	empty sella	corpus callosum dysgenesis	abnormal myelination	na	na	na

n.a = not assessed

(**) The carrier mother was a mosaic for the dup(Xq)

IUGR Intra-Uterin Growth Retardation

Table 9: Rearrangements detected in the region Xq13q21.

5.2 Recruitment of the patients

A total of 366 GHD subjects (342 sporadic and 24 familial cases) have been recruited by several collaborating Italian centres during the last five years. Of these 221 are Isolated Growth Hormone Deficiency (IGHD) with a male:female ratio 2:1 and 145 are Combined Pituitary Hormone Deficiency (CPHD) with a male:female ratio of 1.5:1. In the first step of our work 24 sporadic CPHD male patients with severe short stature (SDS <-2) that have been already tested for known mutations in GH associated genes and with a GH deficiency yet unexplained, have been analysed with a customized array platform (Agilent).

5.3 Results and discussion

Among the 24 selected patients, three did not passed the quality control and could not be analyzed through the customized array platform. In 12 patients suspected polymorphic CNVs have been detected, without genotype-phenotype correlations and already reported in the Database of Genomic Variants.

In four cases we identified 4 rearrangements candidate as pathological, not described in the DGV. A summary of the aCGH results obtained with the analysis of the 4 samples are presented in Table 10.

Other techniques (PCR/MLPA) were used for the confirmation of genomic rearrangements detected in the four patients (n=1 and n= 4 respectively). In all the selected patients the PCR and the MLPA confirmed the imbalances identified through aCGH.

The region Xq13q21 has not been found in any patients. However, a CNV has been detected in the X chromosome distal region of one subjects (#2149, Table10). Interestingly, it is a small Xq27 deletion, located in a non-coding regions 340kb downstream the *SOX3* gene (Fig. 17). *SOX3* mutations and CNVs have already been described in IGHD and CPHD patients with a variable phenotype, sometimes associated with mental retardation.

Moreover CNVs located in non-coding regulatory regions in proximity of *SOX3* and *SOX9* genes (strictly connected to *SOX3*) have been demonstrated to have a critical role in the deregulation of gene expression [75,76,77,78] and in the determination of a clinical phenotype. Our hypothesis is a misregulation of SOX3 in patients carrying this deletion.

Then a 51kb deletion upstream the POU1F1 gene has been detected (#2156, Table 10). It seems like the most interesting rearrangement identified out of X chromosome, near the candidate genes. Also in that case is possible a long-range dysregulation of gene expression.

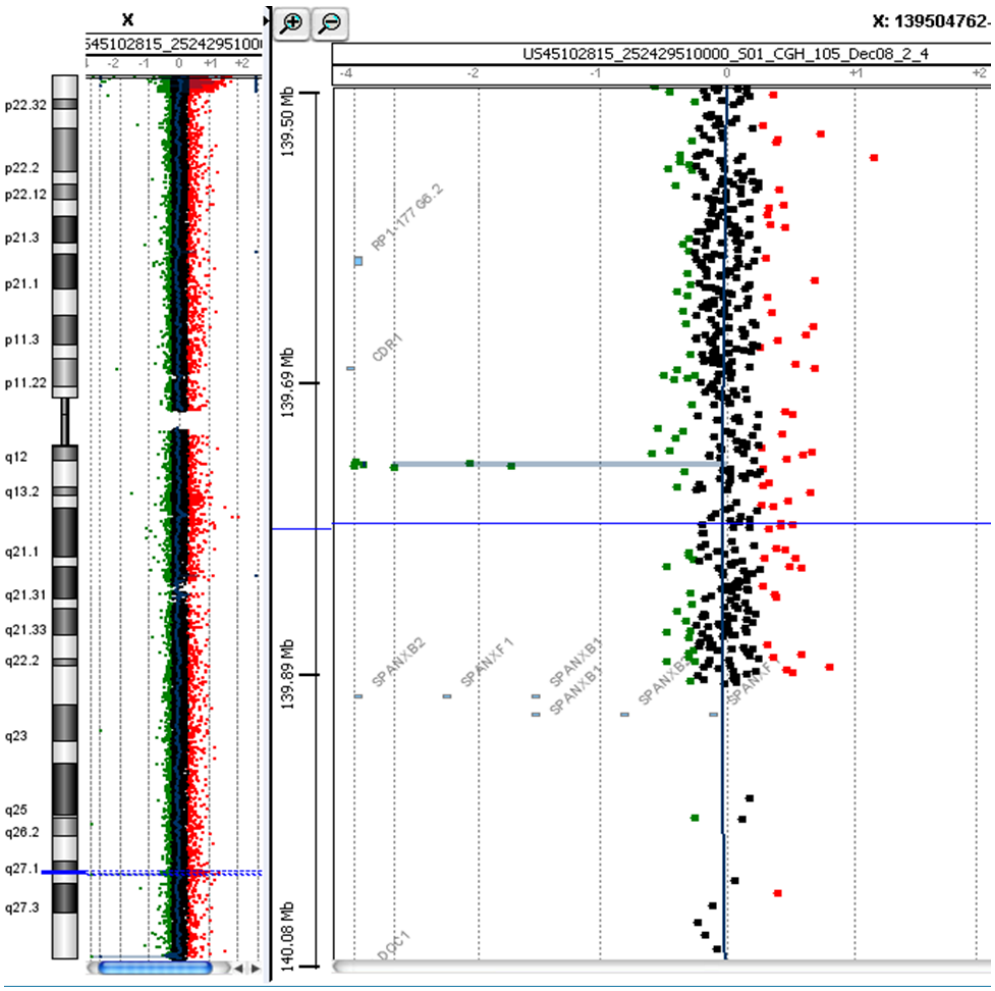

Fig. 17. Identification of delXq27 downstream the SOX3 gene, with aCGH: case #2149.

All the confirmed variants have been tested in a group of patients and healthy controls, for a total of 100 subjects. None deletions/duplications have been found in CPHD (around 60) and IGHD (3 cases tested up to now) patients (Table 11).

Case	Phenotype	Array-CGH results	Chr. Position (start-stop bp)	Size (bp)	Position from the nearest candidate genes.
# 3160	Deficit GH, Gn, pituitary hypoplasia, hypogonadism	Dup 2p	chr2: 40929997-40930056 chr2: 41298981-41299040	369 kb	-
# 2166	Deficit ACTH, TSH, GH, Gn, hypogonadism, idiopathic partial empty sella	Del 18q22.1	chr18: 64307551-64307610 chr18: 64373214-64373273	65 kb	-
# 2149	Deficit ACTH, TSH, GH, Gn, hypogonadism, hypothalamus-pituitary disconnection with ectopic neurohypophysis	Del Xq27	chrX: 139756871-139756930 chrX: 139760128-139760187	3,2 kb	340 kb downstream SOX3
# 2156	Deficit ACTH, TSH, GH, Gn, hypogonadism, partial empty sella	Del 3p11.2	chr3: 87337651-87337710 chr3: 87340551-87340610	2,9 Kb	51 kb upstream POU1F1

Table 10: Chromosomal imbalances detected in 4 patients through custom aCGH.

Case*	Array-CGH results	Positive CPHD patients **/ total patients	Positive IGHD patients/ total patients	Positive controls/ total. controls
#3160	dup 2p	1/60	0/3	0/40
#2166	del 18q22.1	1/54	0/3	0/40
#2149	del Xq27	1/60	-	0/45
#2156	del 3p11.2	1/62	0/3	0/42
* Patient with the rearrangement detected through aCGH ** Patient with the identified CGH variant has been included				

Table 11: Search for the identified variations, in a group of patients and controls.

We planned to investigate the X chromosome and particularly the Xq proximal region for the presence of micro-rearrangements in a large cohort of male patients affected by IGHD and CPHD. In all the patients selected, the presence of point mutations in genes known to be involved in IGHD or CPHD deficiency has already been excluded. In collaboration with University of Pavia, we designed and used a high resolution customized array platform (Agilent) that contains a high oligomer coverage for the region Xq13.q21 along with a lower density coverage for the rest of the X-chromosome.

In the first step of this study we analyzed 21 male patients and we identified 4 (19%) potentially clinically relevant chromosomal imbalances, of which one is on X-chromosome, that could be the cause of patients disease. For all 4 rearrangements we are extending the number of GHD patients and controls to test the CNVs, in order to evaluate the possibility to perform functional studies.

In the course of the project we will expand the number of patients with the same clinical features, to increase the number of cases that could be analyzed with this customized array platform.

By this approach we will identify genomic regions deleted or duplicated on the X chromosome and particularly at the Xq13q21 region that is strongly candidate to contain genes or cis-regulatory elements involved in GHD. The identification of this hypothetical gene/s on the X-chromosome could explain the distorted sex ratio and might be the first gene causing GHD with a major relevance in the disease as mutations in the genes already known to be involved explain only a minority of the cases.

The identification of genes causing short stature associated to pituitary hormone deficiency could aid in the management of individual patients with hypopituitarism.

For the future we are planning to design a new high resolution customized array platform for X-chromosome, containing probes for all the X-linked gene. This platform should greatly increase the detection of rearrangements in males.

6. REFERENCES

6. REFERENCES

1. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. Nat Rev Genet. 2011 May;12(5):363-76.
2. Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human genome. Nat Rev Genet. 2015 Mar;16(3):172-83.
3. Lupski JR, de Oca-Luna RM, Slaugenhaupt S, Pentao L, Guzzetta V, Trask BJ, Saucedo-Cardenas O, Barker DF, Killian JM, Garcia CA, Chakravarti A, Patel PI. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell.* 1991 Jul 26;66(2):219-32.
4. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST: Working Group of the American College of Medical Genetics Laboratory Quality Assurance C: American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. Genet Med 2011; 13:680–685.
5. Srebniak MI, Diderich KE, Govaerts LC, Joosten M, Riedijk S, Galjaard RJ, Van Opstal D. Types of array findings detectable in cytogenetic diagnosis: a proposal for a generic classification. Eur J Hum Genet. 2014 Jul;22(7):856-8.
6. Chen L, Zhou W, Zhang L, Zhang F. Genome architecture and its roles in human copy number variation. Genomics Inform. 2014 Dec;12(4):136-44.
7. Ottaviani D, LeCain M, Sheer D. The role of microhomology in genomic structural variation. Trends Genet. 2014 Mar;30(3):85-94.
8. Stankiewicz P, Lupski JR. Genome architecture, rearrangements and genomic disorders. Trends Genet 2002;18:74-82.
9. Riegel M. Human molecular cytogenetics: From cells to nucleotides. Genet Mol Biol. 2014 Mar;37(1 Suppl):194-209.
10. Blakeslee AF and Avery AG. Methods of inducing doubling of chromosomes in plants. J Hered 1937; 28:392-411.
11. Levan A. The effect of colchicine on root mitosis in Allium. Heredity 1938; 24:471-486.
12. Kannan TP, Zilfalil BA. Cytogenetics: past, present and future. Malays J Med Sci. 2009 Apr;16(2):4-9.
13. Moorhead PS, Nowell PC, Mellman WJ, Battips DM and Hungerford DA.Chromosome preparations of leukocytes cultured from human peripheral blood. Exp Cell Res 1960; 20:613-616.

14. Faas, B.H., et al., *Detection of cryptic subtelomeric imbalances in fetuses with ultrasound abnormalities.* Eur J Med Genet, 2008. **51**(6): p. 511-9.
15. Pinkel D, Gray JW, Trask B, van den Engh G, Fuscoe J, van Dekken H. Cytogenetic analysis by in situ hybridization with fluorescently labeled nucleic acid probes. *Cold Spring Harb Symp Quant Biol.* 1986; **51:**151157.
16. Pinkel D, Straume T, Gray JW. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci USA.* 1986;**83:**29342938.
17. Trask, B.J., *Fluorescence in situ hybridization: applications in cytogenetics and gene mapping.* Trends Genet, 1991. **7**(5): p. 149-54.
18. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science.* 1992;**258:**818821.
19. Kallioniemi A (2008) CGH microarrays and cancer. Curr Opin Biotechnol 19:36-40.
20. Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Döhner H, Cremer T and Lichter P (1997) Matrixbased comparative genomic hybridization: Biochips to screen for genomic imbalances. Genes Chromosomes Cancer 20:399-407.
21. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, *et al.* (1998) High resolution analysis of DNA copy number variations using comparative genomic hybridization to microarrays. Nat Genet 20:207-211.
22. Sanlaville D, Lapierre JM, Turleau C, Coquin A, Borck G, Colleaux L, Vekemans M, Romana SP. Molecular karyotyping in human constitutional cytogenetics. *Eur J Med Genet.* 2005;**48:**214231.
23. Veltman, J.A. and B.B. de Vries. Diagnostic genome profiling: unbiased whole genome or targeted analysis? J Mol Diagn, 2006. **8**(5): p. 534-7; discussion 537-9.
24. Toruner GA, Streck DL, Schwalb MN, Dermody JJ. An oligonucleotide based array-CGH system for detection of genome wide copy number changes including subtelomeric regions for genetic evaluation of mental retardation. Am J Med Genet A. 2007 Apr 15;143A(8):824-9.
25. J.R. Vermeesch, C. Melotte, G. Froyen, S. Van Vooren, B. Dutta, N. Maas, S. Vermeulen, B. Menten, F. Speleman, B. De Moor, P.Van Hummelen, P. Marynen, J.P. Fryns, K. Devriendt, Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis, J. Histochem. Cytochem. 53 (2005) 413–422.
26. Theisen A. Microarray-based Comparative Genomic Hybridization (aCGH). Nature Education 2008; 1(1):45.

27. Aradhya S, Cherry AM. Array-based comparative genomic hybridization: clinical contexts for targeted and whole-genome designs. Genet Med. 2007 Sep;9(9):553-9.
28. Edelmann, L. and K. Hirschhorn, Clinical utility of array CGH for the detection of chromosomal imbalances associated with mental retardation and multiple congenital anomalies. Ann N Y Acad Sci, 2009. 1151: p. 157-66.
29. Hester SD, Reid L, Nowak N, Jones WD, Parker JS, Knudtson K, Ward W, Tiesman J, Denslow ND. Comparison of comparative genomic hybridization technologies across microarray platforms. J Biomol Tech. 2009 Apr;20(2):135-51.
30. Rodríguez-Revenga L1, Vallespín E, Madrigal I, Palomares M, Mur A, García-Miñaur S, Santos F, Mori MÁ, Lapunzina P, Mila M, Nevado J. A parallel study of different array-CGH platforms in a set of Spanish patients with developmental delay and intellectual disability. Gene. 2013 May 25;521(1):82-6.
31. Vasson A1, Leroux C, Orhant L, Boimard M, Toussaint A, Leroy C, Commere V, Ghiotti T, Deburgrave N, Saillour Y, Atlan I, Fouveaut C, Beldjord C, Valleix S, Leturcq F, Dodé C, Bienvenu T, Chelly J, Cossée M. Custom oligonucleotide array-based CGH: a reliable diagnostic tool for detection of exonic copy-number changes in multiple targeted genes. Eur J Hum Genet. 2013 Sep;21(9):977-87.
32. Vallespín E1, Palomares Bralo M, Mori MÁ, Martín R, García-Miñaúr S, Fernández L, de Torres ML, García-Santiago F, Mansilla E, Santos F, M-Montaño VE, Crespo MC, Martín S, Martínez-Glez V, Delicado A, Lapunzina P, Nevado J. Customized high resolution CGH-array for clinical diagnosis reveals additional genomic imbalances in previous well-defined pathological samples. Am J Med Genet A. 2013 Aug;161A(8):1950-60.
33. Liang JS, Shimojima K, Yamamoto T. Application of array-based comparative genome hybridization in children with developmental delay or mental retardation. Pediatr Neonatol. 2008 Dec;49(6):213-7.
34. Stankiewicz P, Beaudet AL. Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic mental retardation. Curr Opin Genet Dev. 2007 Jun;17(3):182-92. Epub 2007 Apr 30.
35. Shapiro BK, Batshaw ML. Intellectual disability. In: Kliegman RM, Behrman RE, Jenson HB, Stanton BF, eds. Nelson Textbook of Pediatrics . 19th ed. Philadelphia, Pa: Elsevier Saunders; 2011:chap 33.
36. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer

SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DHConsensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. . Am J Hum Genet. 2010 May 14;86(5):749-64.

37. Shoukier M1, Klein N, Auber B, Wickert J, Schröder J, Zoll B, Burfeind P, Bartels I, Alsat EA, Lingen M, Grzmil P, Schulze S, Keyser J, Weise D, Borchers M, Hobbiebrunken E, Röbl M, Gärtner J, Brockmann K, Zirn B. Array CGH in patients with developmental delay or intellectual disability: are there phenotypic clues to pathogenic copy number variants? Clin Genet. 2013 Jan;83(1):53-65.
38. Sharp AJ, Mefford HC, Li K, Baker C, Skinner C, Stevenson RE, Schroer RJ, Novara F, De Gregori M, Ciccone R, Broomer A, Casuga I, Wang Y, Xiao C, Barbacioru C, Gimelli G, Bernardina BD, Torniero C, Giorda R, Regan R, Murday V, Mansour S, Fichera M, Castiglia L, Failla P, Ventura M, Jiang Z, Cooper GM, Knight SJ, Romano C, Zuffardi O, Chen C, Schwartz CE, Eichler EE. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. Nat Genet. 2008 Mar;40(3):322-8.
39. Hoppman-Chaney N, Wain K, Seger PR, Superneau DW, Hodge JC. Identification of single gene deletions at 15q13.3: further evidence that CHRNA7 causes the 15q13.3 microdeletion syndrome phenotype. Clin Genet. 2013 Apr;83(4):345-51.
40. Masurel-Paulet A, Andrieux J, Callier P et al. Delineation of 15q13.3 microdeletions. Clin Genet 2010: 78: 149–161.
41. Mikhail FM, Lose EJ, Robin NH et al. Clinically relevant single gene or intragenic deletions encompassing critical neurodevelopmental genes in patients with developmental delay, mental retardation, and/or autism spectrum disorders. Am J Med Genet A 2011: 155A: 2386–2396.
42. Tam GW, van de Lagemaat LN, Redon R, Strathdee KE, Croning MD, Malloy MP, Muir WJ, Pickard BS, Deary IJ, Blackwood DH, Carter NP, Grant SG. Confirmed rare copy number variants implicate novel genes in schizophrenia. Biochem Soc Trans. 2010 Apr;38(2):445-51.
43. McDonald ML, MacMullen C, Liu DJ, Leal SM, Davis RL. Genetic association of cyclic AMP signaling genes with bipolar disorder. Transl Psychiatry. 2012 Oct 2;2:e169.
44. Ramirez AD, Smith SM. Regulation of dopamine signaling in the striatum by phosphodiesterase inhibitors: novel therapeutics to treat neurological and psychiatric disorders. Cent Nerv Syst Agents Med Chem. 2014;14(2):72-82.

45. Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y, Clark D, Cole KE, De Baere E, Decroos C, Di Donato N, Ernst S, Francey LJ, Gyftodimou Y, Hirashima K, Hullings M, Ishikawa Y, Jaulin C, Kaur M, Kiyono T, Lombardi PM, Magnaghi-Jaulin L, Mortier GR, Nozaki N, Petersen MB, Seimiya H, Siu VM, Suzuki Y, Takagaki K, Wilde JJ, Willems PJ, Prigent C, Gillessen-Kaesbach G, Christianson DW, Kaiser FJ, Jackson LG, Hirota T, Krantz ID, Shirahige K. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature. 2012 Sep 13;489(7415):313-7.
46. Ramos FJ, Puisac B, Baquero-Montoya C, Gil-Rodríguez MC, Bueno I, Deardorff MA, Hennekam RC, Kaiser FJ, Krantz ID, Musio A, Selicorni A, FitzPatrick DR, Pié J. Clinical utility gene card for: Cornelia de Lange syndrome. Eur J Hum Genet. 2014 Dec 24.
47. Kaiser FJ, Ansari M, Braunholz D, Concepción Gil-Rodríguez M, Decroos C, Wilde JJ et al. Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance. Hum Mol Genet. 2014 Jun 1;23(11):2888-900.
48. Echaniz-Laguna A, Akman HO, Mohr M, Tranchant C, Talmant-Verbist V, Rolland MO, Dimauro S. Muscle phosphorylase b kinase deficiency revisited. Neuromuscul Disord. 2010 Feb;20(2):125-7.
49. Vandewalle J, Bauters M, Van Esch H, Belet S, Verbeeck J, Fieremans N, Holvoet M, Vento J, Spreiz A, Kotzot D, Haberlandt E, Rosenfeld J, Andrieux J, Delobel B, Dehouck MB, Devriendt K, Fryns JP, Marynen P, Goldstein A, Froyen G. The mitochondrial solute carrier SLC25A5 at Xq24 is a novel candidate gene for non-syndromic intellectual disability. Hum Genet. 2013 Oct;132(10):1177-85.
50. de Leeuw N, Bulk S, Green A, Jaeckle-Santos L, Baker LA, Zinn AR, Kleefstra T, van der Smagt JJ, Vianne Morgante AM, de Vries BB, van Bokhoven H, de Brouwer AP. UBE2A deficiency syndrome: Mild to severe intellectual disability accompanied by seizures, absent speech, urogenital, and skin anomalies in male patients. Am J Med Genet A. 2010 Dec;152A(12):3084-90.
51. Honda S1, Orii KO, Kobayashi J, Hayashi S, Imamura A, Imoto I, Nakagawa E, Goto Y, Inazawa J. Novel deletion at Xq24 including the UBE2A gene in a patient with X-linked mental retardation. J Hum Genet. 2010 Apr;55(4):244-7.

52. Bartholdi D, Stray-Pedersen A, Azzarello-Burri S, Kibaek M, Kirchhoff M, Oneda B, Rødningen O, Schmitt-Mechelke T, Rauch A, Kjaergaard S. A newly recognized 13q12.3 microdeletion syndrome characterized by intellectual disability, microcephaly, and eczema/atopic dermatitis encompassing the HMGB1 and KATNAL1 genes. Am J Med Genet A. 2014 May;164A(5):1277-83.
53. Andreou AM, Pauws E, Jones MC, Singh MK, Bussen M, Doudney K, Moore GE, Kispert A, Brosens JJ, Stanier P: TBX22 missense mutations found in patients with X-linked cleft palate affect DNA binding, sumoylation, and transcriptional repression. Am J Hum Genet 2007, 81: 700-712.
54. Pauws E, Peskett E, Boissin C, Hoshino A, Mengrelis K, Carta E, Abruzzo MA, Lees M, Moore GE, Erickson RP, Stanier P: X-linked CHARGE-Like Abruzzo-Erickson syndrome and classic cleft palate with ankyloglossia result from TBX22 splicing mutations. Clin Genet 2013, 83:352-358.
55. de Kok YJ , Vossenaar ER, Cremers CW, Dahl N, Laporte J, Hu LJ, Lacombe D, Fischel-Ghodsian N, Friedman RA, Parnes LS, Thorpe P, Bitner-Glindzicz M, Pander HJ, Heilbronner H, Graveline J, den Dunnen JT, Brunner HG, Ropers HH, Cremers FP. Identification of a hot spot for microdeletions in patients with X-linked deafness type 3 (DFN3) 900 kb proximal to the DFN3 gene POU3F4. Hum Mol Genet 1996, 5:1229-1235.
56. Rush ET, Schaefer GB: Identification of an X-linked deletion syndrome through comparative genomic hybridization microarray. Semin Pediatr Neurol 2010, 17:51-53.
57. Vore AP, Chang EH, Hoppe JE, Butler MG, Forrester S, Schneider MC, Smith LL, Burke DW, Campbell CA, Smith RJ: Deletion of and novel missense mutation in POU3F4 in 2 families segregating X-linked nonsyndromic deafness. Arch Otolaryngol Head Neck Surg 2005, 131:1057–1063.
58. Marlin S, Moizard MP, David A, Chaissang N, Raynaud M, Jonard L, Feldmann D, Loundon N, Denoyelle F, Toutain A: Phenotype and genotype in females with POU3F4 mutations. Clin Genet 2009, 76: 558-563.
59. Field M, Tarpey PS, Smith R, Edkins S, O'Meara S, Stevens C, Tofts C, Teague J, Butler A, Dicks E, Barthorpe S, Buck G, Cole J, Gray K, Halliday K, Hills K, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Varian J, West S, Widaa S, Mallya U, Wooster R, Moon J, Luo Y, Hughes H, Shaw M, Friend KL, Corbett M, Turner G, Partington M, Mulley J, Bobrow M, Schwartz C, Stevenson R, Gecz J, Stratton MR, Futreal PA, Raymond FL. Mutations in the BRWD3 gene cause X-linked mental retardation associated with macrocephaly Am J Hum Genet 2007 81:367-74.

60. Grotto S, Drouin-Garraud V, Ounap K, Puusepp-Benazzouz H, Schuurs-Hoeijmakers J, Le Meur N, Chambon P, Fehrenbach S, van Bokhoven H, Frébourg T, de Brouwer AP, Saugier-Veber P: Clinical assessment of five patients with 349 BRWD3 mutation ad Xq21.1 gives further evidence for mild to moderate intellectual disability and macrocephaly. Eur J Med Genet 2014, 57:200-206.
61. Ozturk N, VanVickle-Chavez SJ, Akileswaran L, Van Gelder RN, Sancar A: Ramshackle (Brwd3) promotes light-induced ubiquitylation of Drosophila Cryptochrome by DDB1-CUL4-ROC1 E3 ligase complex. Proc Natl Acad Sci U S A 2013, 110 :4980-4985.
62. Pfäffle R, Klammt J: Pituitary transcription factors in the aetiology of combined pituitary hormone deficiency. Best Pract Res Clin Endocrinol Metab 2011, 25:43-60.
63. Yokoyama Y, Narahara K, Tsuji K, Moriwake T, Kanzaki S, Murakami M, Namba H, Ninomiya S, Higuchi J, Seino T: Growth hormone deficiency and empty sella syndrome in a boy with dup(X)(q13.3-q21.2). Am J Med Genet 1992, 42:660-664.
64. Shapira M, Dar H, Bar-El H, Bar-Nitzan N, Even L, Borochowitz Z: 373 Inherited inverted duplication of X chromosome in a male: report of a patient and review of the literature. Am J Med Genet 1997, 72:409-414.
65. Sismani C, Donoghue J, Alexandrou A, Karkaletsi M, Christopoulou S, Konstantinidou AE, Livanos P, Patsalis PC, Velissariou V: A prenatally ascertained, maternally inherited 14.8 Mb duplication of chromosomal bands Xp13.23-q21.31 associated with multiple congenital abnormalities in a male fetus. Gene 2013, 530:138-142.
66. Van den Plas D, Merregaert J: In vitro studies on Itm2a reveal its involvement in early stages of the chondrogenic differentiation pathway. Biol Cell 2004, 96:463-470.
67. Boeuf S, Börger M, Hennig T, Winter A, Kasten P, Richter W: Enhanced ITM2A expression inhibits chondrogenic differentiation of mesenchymal stem cells. Differentiation 2009, 78:108-115.
68. Frindik JP, Baptista J. Adult height in growth hormone deficiency: historical perspective and examples from the national cooperative growth study. Pediatrics. 1999;104(4 Pt 2):1000-4.
69. Lindsay R, Feldkamp M, Harris D, Robertson J, Rallison M. Utah growth study: growth standards and the prevalence of growth hormone deficiency. J Pediatr. 1994;125:29–35.
70. Root AW, Kemp SF, Rundle AC. Effect of long-term recombinant growth hormone therapy in children--the National Cooperative Growth Study. J Pediatr Endocrinol Metab. 1998;11:403-12.

71. Yokoyama Y, Narahara K, Tsuji K, Moriwake T, Kanzaki S, Murakami M, Namba H, Ninomiya S, Higuchi J, Seino Y. Growth hormone deficiency and empty sella syndrome in a boy with dup(X) (q13.3-q21.2). Am J Med Genet. 1992;42: 660-4.
72. Hou JW; Chang Gung Med J. Inherited tandem duplication of the X chromosome: dup(X)(q13.2-q21.2) in a family. 2004;27:685-90.
73. Shapira M, Dar H, Bar-El H, Bar-Nitzan N, Even L, Borochowitz Z. Inherited inverted duplication of X chromosome in a male: report of a patient and review of the literature. Am J Med Genet. 1997;72:409-14.
74. Gabbett MT, Peters GB, Carmichael JM, Darmanian AP, Collins FA. Prader-Willi syndrome phenocopy due to duplication of Xq21.1-q21.31, with array CGH of the critical region. Clin Genet. 2008;73:353-9.
75. Sutton E, Hughes J, White S, Sekido R, Tan J, Arboleda V, Rogers N, Knower K, Rowley L, Eyre H, Rizzoti K, McAninch D, Goncalves J, Slee J, Turbitt E, Bruno D, Bengtsson H, Harley V, Vilain E, Sinclair A, Lovell-Badge R, Thomas P. Identification of SOX3 as an XX male sex reversal gene in mice and humans. J Clin Invest. 2011 4;121:328-41
76. Benko S, Gordon CT, Mallet D, Sreenivasan R, Thauvin-Robinet C, Brendehaug A, Thomas S, Bruland O, David M, Nicolino M, Labalme A, Sanlaville D, Callier P, Malan V, Huet F, Molven A, Dijoud F, Munnich A, Faivre L, Amiel J, Harley V, Houge G, Morel Y, Lyonnet S. Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. J Med Genet. 2011 48:825-30.
77. Vetro A, Ciccone R, Giorda R, Patricelli MG, Della Mina E, Forlino A, Zuffardi O. XX males SRY negative: a confirmed cause of infertility. J Med Genet. 2011 ;48:710-2.
78. Kurth I, Klopocki E, Stricker S, van Oosterwijk J, Vanek S, Altmann J, Santos HG, van Harssel JJ, de Ravel T, Wilkie AO, Gal A, Mundlos S. Duplications of noncoding elements 5' of SOX9 are associated with brachydactyly-anonychia. Nat Genet. 2009 41:862-3.

Acknowledgements

Un grande e profondo ringraziamento ad Antonio, ai miei genitori, alle mie sorelle e a Margot.

Ringrazio Mara, Deepak, Simona, Lucia, Alessandra, Nadia, Giulia e tutto il laboratorio.

Infine un ringraziamento a Matilde che sta per arrivare!

Dedicata a Salvatore…

Publications

Novel Mutations in the GH Gene (*GH1*) Uncover Putative Splicing Regulatory Elements

Deepak Babu, Simona Mellone, Ileana Fusco, Antonella Petri, Gillian E. Walker, Simonetta Bellone, Flavia Prodam, Patricia Momigliano-Richiardi, Gianni Bona, and Mara Giordano

Laboratory of Genetics (D.B., S.M., I.F., P.M.-R., M.G.), Department of Health Sciences, University of Eastern Piedmont and Interdisciplinary Research Center of Autoimmune Diseases, and Unit of Pediatrics (A.P., G.E.W., S.B., F.P., G.B.), Department of Health Sciences, University of Eastern Piedmont, 28100 Novara, Italy

Mutations affecting exon 3 splicing are the main cause of autosomal dominant Isolated GH Deficiency II (IGHDII) by increasing the level of exon 3-skipped mRNA encoding the functionally inactive dominant-negative 17.5-kDa isoform. The exons and introns of the gene encoding GH (*GH1*) were screened for the presence of mutations in 103 sporadic isolated GH deficiency cases. Four different variations within exon 3 were identified in 3 patients. One carried c.261C>T (p.Pro87Pro) and c.272A>T (p.Glu91Val), the second c.255G>A (p.Pro85Pro) and c.261 C>T, and the third c.246G>C (p.Glu82Asp). All the variants were likely generated by gene conversion from an homologous gene in the *GH1* cluster. In silico analysis predicted that positions c.255 and c.272 were included within 2 putative novel exon splicing enhancers (ESEs). Their effect on splicing was confirmed in vitro. Constructs bearing these 2 variants induced consistently higher levels both of transcript and protein corresponding to the 17.5-kDa isoform. When c.255 and c.272 were combined in *cis* with the c.261 variant, as in our patients, their effect was weaker. In conclusion, we identified 2 variations, c.255G>A and c.272A>T, located in 2 novel putative exon splicing enhancers and affecting *GH1* splicing in vitro by increasing the production of alternatively spliced isoforms. The amount of aberrant isoforms is further regulated by the presence in *cis* of the c.261 variant. Thus, our results evidenced novel putative splicing regulatory elements within exon 3, confirming the crucial role of this exon in mRNA processing. **(*Endocrinology* 155: 1786–1792, 2014)**

The *GH1* gene is located on chromosome 17q23 within a cluster of 5 highly homologous genes, all consisting of 5 exons and 4 introns, including the placentally expressed *GH2*, 2 chorionic somatomammotropin genes *CSH1* and *CSH2*, and a pseudogene *CSHP1* (1). When correctly spliced, *GH1* produces the 22-kDa isoform that includes all the 5 exons with the complete biological activity of GH (2, 3). Despite the correct processing, even under normal conditions, a small percentage of alternatively spliced isoforms are produced. The presence of an in-frame cryptic splice site within exon 3 gives rise to a transcript lacking the first 45 bp of exon 3 and encodes a shorter active isoform of 20 kDa, representing 5%–10% of GH transcripts (4). A 17.5-kDa isoform (representing 0.1%–5% of GH transcripts) is produced by the complete skipping of exon 3, thus lacking the entire loop connecting helix 1 and helix 2 in the tertiary structure of GH and generating a GH isoform with no biological activity (5). Trace amounts of the severely truncated isoforms of 11.3 and 7.4 kDa, which are biologically inactive, have also been identified being generated by the skipping of exons 3 and 4 or 2 to 4, respectively (6). Multiple mechanisms have evolved to maintain the small amounts of these aberrantly spliced isoforms, especially that encoding the 17.5-kDa protein. Because *GH1* has weak canonical splice sites, multiple *cis*-acting splicing regulatory elements (splicing

ISSN Print 0013-7227 ISSN Online 1945-7170
Printed in U.S.A.

Received December 18, 2013. Accepted March 4, 2014.
First Published Online March 17, 2014

Abbreviations: ASF/SF2, alternative splicing factor 1/pre-mRNA splicing factor SF2; ESE, exon splicing enhancer; IGHD, isolated GH deficiency; SR proteins, serine/arginine-rich proteins; SRp40, serine/arginine-rich splicing factor 5; TW2, Tanner-Whitehouse 2nd Edition.

doi: 10.1210/en.2013-2146

enhancers) are essential to maintain the correct exon 3 definition through the activation of the canonical intron 2 and 3 splice sites and silencing of the cryptic sites. Two exon splicing enhancers (ESEs), ESE1, encompassing the first 7 bases of exon 3 (from c.172 to c.178), and ESE2 (from c.190 to c.204), located 12 nt upstream of the cryptic splice site in exon 3 (Figure 1A), and an intron splicing enhancer (ISE) within intron 3, have been well characterized (7–11).

Several mutations leading to aberrant splicing have been reported in isolated GH deficiency (IGHD) patients within these enhancer motifs (7, 8, 10, 11). The increased amount of the 17.5-kDa isoform exhibits a dominant negative effect both in tissue culture and transgenic mice experiments by disrupting the secretory pathway and trafficking of normal GH and other hormones, including ACTH (12, 13).

We here report the identification of variations within *GH1* exon 3 in sporadic IGHD patients and absent in a group of 205 normal stature controls and in the public databases. In silico analysis suggested that 2 of these variations affect exon 3 splicing, because they are located within putative ESEs. In vitro mRNA and Western blot analysis confirmed the deleterious effect of the single variations on splicing, suggesting the presence of further splicing regulatory elements within *GH1* exon 3.

Subjects and Methods

Subjects

A total of 103 sporadic patients with IGHD and 205 normal stature individuals, all belonging to the Italian population, were included in the genetic analysis. The subjects were referred to the clinical centers because they had a height less than or equal to −2 SDS or a height less than or equal to −1.8 SDS in combination with a height velocity over 1 year less than −1.5 SDS using the criteria of Tanner-Whitehouse (14). Patients with a known postnatal cause of acquired hypopituitarism were excluded. Skeletal maturation was estimated as bone age (radius, ulna, and short bone) with the TW2 (Tanner-Whitehouse 2nd Edition) method by a pediatric endocrinologist (15). They were all evaluated for GH serum levels after 2 provocative tests (with arginine or clonidine or insulin or glucagon or with GHRH + arginine (16). Traditionally, a diagnosis of GHD is supported by GH peaks less than 10 ng/mL both after the 2 different stimuli (17), or less than 20 ng/mL after the double provocative test with GHRH + arginine. The GHD patients fulfill these criteria and had a mean (±SD) secretion peak of 4.4 ± 2.5 ng/mL after the classical stimuli (n = 78) or 9.4 ± 5.8 ng/mL after the test with GHRH + arginine (n = 25). None of the GHD patients was deficient for

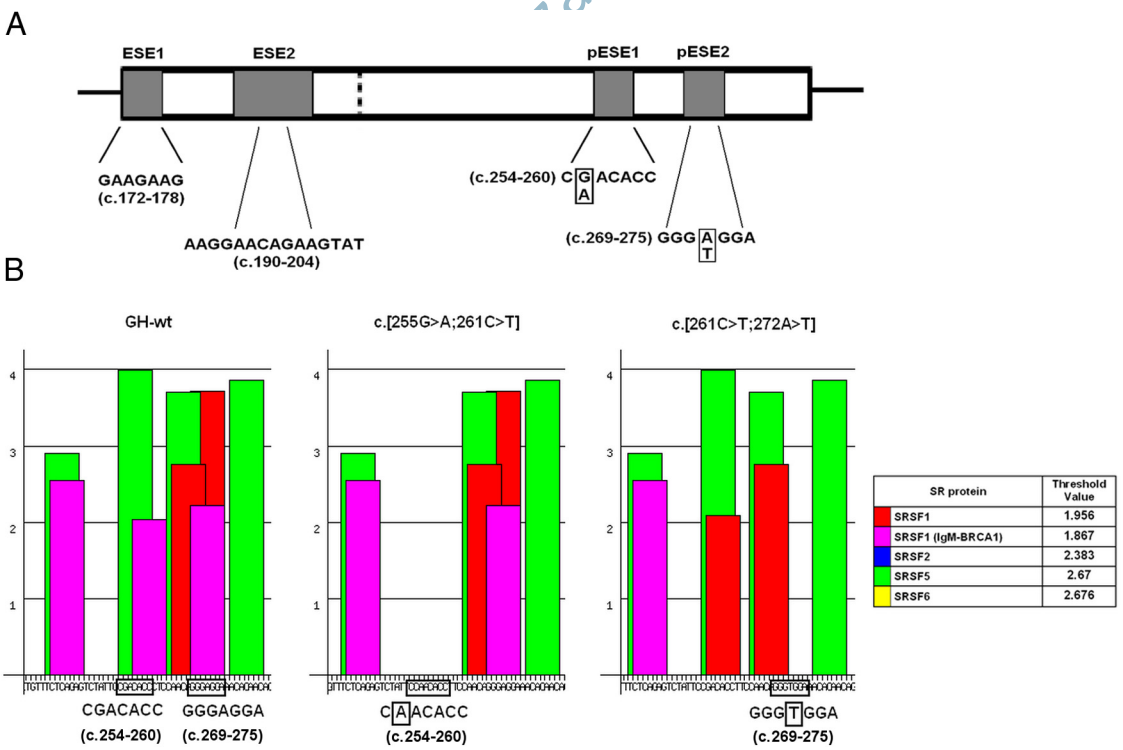

Figure 1. A, ESEs in exon 3. The 2 known ESEs (ESE1 and ESE2) are indicated. The cryptic splice site is at c.216 and is indicated with a dotted line. The newly identified putative ESEs are indicated as pESE1 and pESE2. The mutations identified in patient 1 and patient 2 in the pESEs are boxed. B, ESE finder analysis performed on the wild-type exon 3 and on the different exon 3 alleles identified in patients. The analysis using wild-type sequence revealed all the putative exon splicing enhancer sites within exon 3. ESE motifs with scores above the threshold for each SR protein are indicated in differently colored bars. The bar heights reflect the score of the motifs. Threshold values and color code for each of the different motifs are indicated in Table 1. The introduction of c.255G>A and c.272A>T variations caused the abolition of the 2 putative ESE elements at positions c.254–260 and c.269–275, respectively.

other pituitary hormones, and none had a documented family history of the disease or consanguineous parents. All the patients have been also screened for mutations in *GHRHR* (data not shown). Patients carrying mutations in this gene were not included in the IGHD cohort described here. Normal stature controls included University and Hospital staff, as well as medical students not tested for GH secretion levels. A written informed consent was obtained from the patient's parents, because they were all aged less than 18, and from the normal stature controls.

Detailed description of patients carrying the *GH1* variations

The variants identified in the three patients are reported in Table 1.

Case 1

This patient is a boy born at term with no perinatal complications. He came to our attention when he was 7.3 years old with a height of −2.7 SDS. He has normal stature parents: the father is −1.7 SDS and the mother −0.6 SDS. He was prepubertal. GHD was diagnosed based on low response to insulin tolerance test and clonidine provocative tests (8.9 and 8.6 ng/mL, respectively). The IGF-I level was 130 ng/mL. The bone age was delayed by 3.8 years. GH therapy was initiated with a good clinical response reaching a height of −1.7 SDS after 1.3 after 6 years.

Case 2

This case is a male born from nonconsanguineous parents by vaginal delivery after 41 weeks of gestation. At birth, he was adequate for gestational age (birth weight, 3330 g; length, 52 cm). Both parents presented normal height: father, 174 cm (−0.2 SDS) and mother, 155 cm (−1.3 SDS). He came to our attention at age 13 presenting short stature with a height of −3.4 SDS. His growth rate in the year preceding the diagnosis was 5.0 cm/y (−1.5 SDS). At the diagnosis, the pubertal stage was pubic hair development 2, genital development 2, and bilateral testicular volume 5. The bone age was delayed (11.6, TW2 method). GH secretion peaks after stimulus with arginine and clonidine were 3.9 and 8.3 ng/mL, respectively, with an IGF-I level of 65 ng/mL. He promptly started the recombinant human GH replacement therapy (0.22 mg/kg·wk) and reached the height of −1.8 SDS after 4 years.

Case 3

This boy came for the first time to our attention for growth failure at the age of 16.5 years. At diagnosis, his stature was −2.7 SDS. He was born to nonconsanguineous parents after a normal pregnancy at 40 weeks of gestation, adequate for gestational age (birth weight, 2950 g). His father and mother presented normal (−0.5 and −0.3 SDS, respectively). The pubertal stage was: pubic hair development 3, genital development 3, and bilateral testicular volume 8 mL. The growth velocity was 4.5 cm/y (1.6 SDS, 1.6 SDS, but low when corrected for pubertal stage), and x-ray revealed a bone age (TW2) of 14.2 years. Laboratory analysis revealed a normal IGF-I level (180 ng/mL) and low level of circulating GH after 2 provocative tests (7.7 ng/mL with arginine and 3.7 with clonididne). The recombinant human GH replacement therapy (0.24 mg/kg·wk) was begun with a good response, because he reached the height of −0.7 SDS after 3 years.

Table 1. Exon 3 Variations Detected in the IGHD Patients

Patient	Variation Nucleotide	Amino acid	Within a predicted ESE
1	c.261C>T	(p.Pro87Pro)	No
	c.272A>T	(p.Glu91Val)	Yes
2	c.255G>A	(p.Pro85Pro)	Yes
	c.261C>T	(p.Pro87Pro)	No
3	c.246G>C	(p.Glu82Asp)	No

Predicted ESE by the software ESE finder 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi).

Screening of *GH1*

Genomic DNA was amplified by PCR using previously described primers (18) and a proofreading Taq polymerase (Finnzymes). The resulting 2.7-kb product, including the whole *GH1*, was used as template for a series of nested PCRs using internal primers for the proximal promoter, 5 exons, 4 introns, and the untranslated regions of the *GH1*. These primers are designed specifically for the *GH1* and do not amplify other genes in the cluster. PCR conditions and primer sequences are available upon request. The PCR products were visualized on a 2% agarose gel and purified using ExoSAP-IT enzymatic PCR clean up system (Affymetrix). The purified products were then sequenced with the Big Dye Terminator kit (Applied Biosystems) and analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). The PCR products containing the variations in the heterozygous state were then cloned into the plasmid vector pTZ57R/T using the InsTAclone PCR cloning kit (Fermentas), and the 2 alleles were separately sequenced.

Site-directed mutagenesis

The 2.7-kb fragment containing *GH1* was inserted into the pcDNA 3.1(+) expression vector (Invitrogen). The constructs bearing the single variants, namely 246C, 255A, 261T, and 272T, and the variants combined as in the patients, namely 261T/272T (patient 1) and 255A/261T (patient 2), were generated from the GH(wt)-pcDNA 3.1 plasmid by the QuikChange Site-Directed Mutagenesis kit from Stratagene using mismatch complementary primers containing the desired mutations. PCR conditions and primer sequences are available upon request. DH5a competent cells were transformed with the different constructs and grown on Luria Broth/ampicillin media. After selecting the correct clones by colony PCR, the plasmid DNA was isolated using Maxiprep kit (QIAGEN). The desired mutation was confirmed by sequencing.

Cell culture, transfection, and isolation of RNA

The GH4C1 rat pituitary cell line was used for the transfection experiments. The stock culture was grown in Ham's F10 medium (Gibco-Life Technologies) supplemented with 15% horse serum, 2.5% fetal bovine serum, 100-U/mL penicillin, and 100-μg/mL streptomycin in a 5% CO_2. A day before transfection, 4×10^5 cells were seeded into each well of a 6-well tissue culture plate in 2.5-mL medium. The wells were previously treated with 1:10 diluted poly-L-lysine solution (Sigma-Aldrich) to allow the cells to completely attach to the plate. At 50%–70% confluency, cells were transfected with 2.5-μg DNA of the

wt-GH or the mutated constructs using the Trans IT-LT1 transfection reagent (Mirus Bio LLC). A green fluorescent protein control was used to test transfection efficiency. Forty-eight hours after transfection, total RNA was isolated and purified from the cells using the QIAGEN RNA mini kit (QIAGEN).

cDNA synthesis

cDNA was synthesized from 1.5 μg of RNA by the High Capacity cDNA Reverse Transcription kit (Applied Biosystems), according to the manufacturer's instructions. The different transcripts produced by alternative splicing were analyzed using primers specific for *GH1* cDNA (Figure 2A). The RT-PCR was performed with primers GH2 (5′-CGTCTGCACCAGCTGGCCTTT-3′) and GH7 (5′-AAGCCACAGCTGCCCTCCACAGA-3′), which amplify part of exon 2, exon 3, exon 4, and part of exon 5, allowing detection of both exon 3- and exon 4-skipped products.

Western immunoblot analysis

CHO cells were transiently transfected with wt-GH and mutated constructs, as described above. After 48 hours, whole-cell lysates (WCL) were collected using the standard radio-immuno precipitation assay (RIPA) lysis buffer containing 0.1% sodium dodecyl sulfate (SDS). A total of 20 μg of WCL were separated on 15% SDS-PAGE gel and blotted on Immun-Blot PVDF membrane (Bio-Rad). Membranes were probed with a polyclonal rabbit antihuman GH antibody (Abnova) and detected with a secondary horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (Millipore). A polyclonal antiactin antibody (Sigma-Aldrich) was used to normalize the protein loading. Protein bands were visualized using enhanced chemiluminescence reagent (Thermo Scientific) with image capture performed using a charge-coupled device camera linked to ChemiDoc apparatus (Bio-Rad).

CHO cells were used instead of GH4C1 cells for protein analysis, because the GH4C1 showed many unspecific bands after Western blotting, likely due to cross-reactions with endogenous proteins.

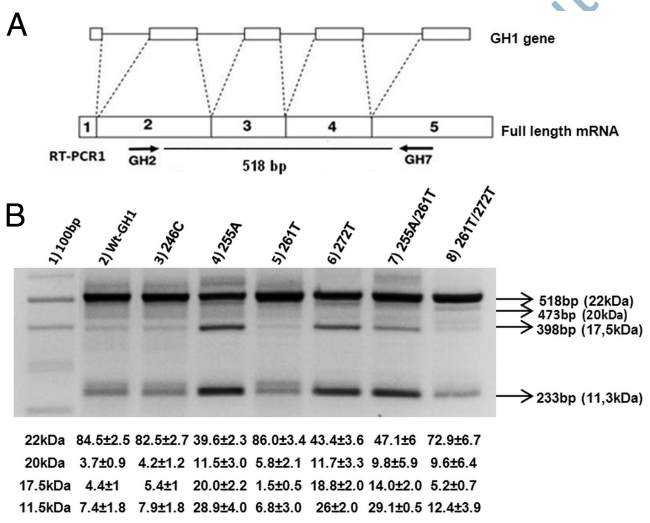

Figure 2. A, Scheme of GH1 pre-mRNA splicing showing the full-length mRNA and the corresponding RT-PCR product. B, RT-PCR performed on mRNA extracted from GH4C1 pituitary cells transfected with the wild-type *GH1* (lane 2) or constructs carrying different variations (lanes 3–8). The size of the bands are indicated on the right with the corresponding protein molecular weight of the different isoforms. Untransfected rat cDNA did not show any bands (data not shown). Each PCR product was eluted from the gel and characterized by sequencing. The relative band intensity (expressed as the percentage on the total yield) calculated as the mean ± SD over 4 different independent transfection experiments is reported for each construct below the corresponding lane.

ESE finder analysis

Analysis of the splicing regulatory motifs within exon 3 was performed using the software ESE finder 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi). The default thresholds were considered to identify sites responsible for the 4 serine/arginine-rich (SR) proteins alternative splicing factor 1/pre-mRNA splicing factor SF2 (ASF/SF2), serine/arginine-rich splicing factor 3 (SC35), serine/arginine-rich splicing factor 5 (SRp40), and serine/arginine-rich splicing factor 6 (SRp55).

Gel image analysis

The RT-PCR and Western blotting gel images were analyzed using the freeware ImageJ1.46r (http://rsb.info.nih.gov/ij/; National Institutes of Health), and the bands were quantified by measuring pixel intensity and normalized to the corresponding β-actin band intensity.

Results

Genetic analysis of the IGHD patients

A total of 103 sporadic IGHD patients with height ranging from −1.8 to −4.5 SDS were investigated for the presence of functionally relevant mutations in *GH1*, including coding regions and introns. Two patients harbored a combination of 2 variations within exon 3 (Table 1) that were not previously reported in public databases, including dbSNP (http://www.ncbi.nlm.nih.gov/SNP/), the Exome Variant Server (http://evs.gs.washington.edu/EVS/), and 1000 genomes (http://www.1000genomes.org/). Patient 1 carried c.261C>T (p.Pro87Pro) and c.272A>T (p.Glu91Val), and patient 2 carried c.255G>A (p.Pro85Pro) and c.261C>T (p.Pro87Pro).

A third patient (patient 3) (Table 1) carried the non-synonymous c.246G>C determining the substitution p.Glu82Asp in exon 3. This variation was reported in the dbSNP (rs61762497) only in 1 individual and in the Exome Variant Server in 2 out of 13 000 individuals.

All the above exon 3 variations were absent in a panel of 410 chromosomes sequenced from 205 normal stature Italian control individuals.

The analysis of the parents of patients 2 and 3 demonstrated that their variants were inherited from the unaffected fathers. Unfortunately, the parents of patient 1 did not give their consent to DNA analysis. By subcloning the patient's PCR products in a TA cloning vector system, we confirmed that in patient 1, as well as in patient 2, both

variants were on the same allele. The alignment of *GH1* with the paralogous genes of the GH cluster suggested that all the 4 variants in exon 3 were generated by nonallelic gene conversion from the *GH2* gene, because the other 3 genes (*CS-5*, *CS-2*, and *CS-1*) have the same *GH1* sequence at these sites.

The exon 3 variations fall within predicted ESEs

Because several exon 3 mutations lead to missplicing of mRNA and production of increased amounts of the 17.5-kDa protein, we evaluated the possible involvement in the splicing regulation of the variants identified in our patients within this exon. None of them was included in the previously described ESEs (ESE1 and ESE2). An in silico analysis using ESE finder 3.0 (Figure 1, A and B) revealed the presence of 2 high scoring ESE motifs from c.254 to c.260 (CGACACC) and c.269 to c.275 (GGGAGGA). These 2 sequences include the variations c.255G>A and c.272 A>T, respectively (Figure 1A). The position c.255 is located within an ESE motif recognized by the SR protein SRp40 and c.272 within a sequence recognized by SRp40 and SF2/ASF. Both putative ESE sequences showed an increased score compared with the threshold value (Figure 1B). By substituting the wild-type nucleotides with those found in the patients, namely c.255A and c.272T, the software predicted the complete loss of these 2 putative ESEs (Figure 1B).

Variants c.255A and c.272T affect *GH1* splicing in vitro

To evaluate whether exon 3 variants actually had some effect on mRNA splicing in vitro, we transfected GH4C1 rat pituitary cells with an expression vector containing 1) the wild-type allele (wt-*GH1*), 2) the alleles carrying the single mutations, and 3) the alleles containing the mutations combined as they were in patients 1 and 2. The mRNA from transfected cells was reverse transcribed, and the resulting cDNA was amplified with primers (Figure 2A) specific for the human *GH1* that did not amplify the rat mRNA. The RT-PCR on the wt-*GH1* mRNA (Figure 2B, lane 2) yielded an intense band corresponding to the *GH1* full-length transcript (518 bp) producing the 22-kDa protein and faint bands corresponding to the 20-kDa isoform (473 bp), the 17.5-kDa isoform (398 bp), and the 11.3-kDa isoform (233 bp). The average yield of the mRNA corresponding to the 17.5-kDa isoform over 4 independent experiments was 4.4 ± 1% of the total mRNA yield with a ratio 17.5 kDa/22 kDa of about 1:20. The 246C construct (Figure 2B, lane 3) was also tested, although the c.246C variation was not predicted to influence splicing. This construct generated a band pattern similar to that observed in the wild-type accordingly to the in silico prediction. In contrast, 255A and 272T constructs (Figure 2B, lanes 4 and 6, respectively) produced a higher level of the exon 3-skipped mRNA with a ratio 17.5 kDa/22 kDa of about 1:2 and 1:2.4, respectively. In these 2 constructs, there was also an evident increased production of the exon 3–4-skipped transcript (233-bp band). Interestingly, the exon 3-skipped mRNA was not evident in the transcripts from the 261T construct, suggesting that this variant might strengthen the correct splicing. When the variants were combined on the same construct to reproduce the status of patients 1 and 2, the exon 3-skipped mRNA produced by the construct 255A/261T (patient 2) (Figure 2B, lane 7) was less abundant than that observed for 255A but still evident, with a 17.5 kDa/22 kDa ratio of 1:3.3. Conversely, the 261T/272T construct (patient 1) (Figure 2B, lane 8) showed a splicing pattern very similar to the wild type.

We then investigated the consequences of the variants on protein synthesis (Figure 3). Western blot analysis confirmed that the transfected wt-*GH1*-produced mostly the 22-kDa protein with only traces of the 20 and 17.5 kDa. A higher level of the 17.5-kDa product was produced by 255A, 272T, and by the 255A/261T constructs. Although the 261T/272T plasmid carrying the 2 variations detected in patient 1 showed a mRNA pattern similar to the wild type (Figure 2B, lane 8), it exhibited a band in correspondence of the 17.5-kDa protein more intense than the wild type (about 10% of the total GH proteins) (Figure 3, lane 7). This band in the 261T construct was weak (as in the wt-*GH1*), but the 20-kDa isoform was more intense than in all the other constructs. The corresponding 11.3-kDa protein isoform was not detectable by the antibody used for Western blot analysis.

Discussion

In the present study, we sequenced *GH1* in 103 IGHD patients with clinically variable phenotypes and no family

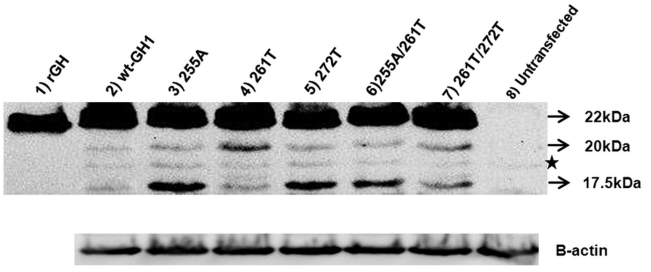

Figure 3. Western immunoblot analysis of the different GH isoforms encoded by the different splicing products. The CHO cells were transiently transfected with either wt-*GH1* (lane 2) or constructs carrying different variations (lanes 3–7). Untransfected CHO cells were used as a negative control (lane 8); *, nonspecific bands also present in the untransfected CHO cell lysate.

history to search for mutations that might be 1) inherited from healthy parents (incomplete penetrance), 2) arisen de novo, or 3) biallelic (recessive inheritance). We identified 3 patients that carried variations in exon 3, and 2 of them carried 2 variants on the same allele. All the variants were likely generated by gene conversion, an event that is frequently associated to *GH1* sequence variability (19–21). An in silico analysis was performed using ESE finder 3.0, which has been designed to identify SR protein binding sites by generating a motif score that reflects the binding site strength. The sequences from c.254 to c.260 and from c.269 to c.275, including the positions c.255 and c.272 mutated in patient 2 and in patient 1, respectively, were predicted to represent ESE motifs recognized by the SRp40 and SF2/ASF proteins with a score above the threshold (Figure 1B). The same factors were predicted to recognize ESE2 with the same strength (9). It has been demonstrated that SF2/ASF activates exon 3 inclusion and that disruption of this motif causes increased exon 3 skipping (22). The substitution of the wild-type nucleotides c.255G and c.272A with the mutants c.255A and c.272T, respectively, was predicted to abolish the binding of these proteins in both putative ESEs (Figure 1). The RT-PCR analysis (Figure 2) confirmed the influence of c.255A and c.272T on splicing by increasing the exon 3-skipped isoform to 18%–20% of the total GH mRNA (Figure 2B, lanes 4 and 6, respectively).

Mutations within ESE1 in *GH1* have been previously reported to cause either complete or partial exon 3 skipping and generation of increased amounts of the 20- and the 17.5-kDa isoforms at various concentrations (20%–37% and 35%–68%, respectively) (8, 9). The clinical variability observed in patients carrying these mutations has been explained by variable amounts of the 17.5-kDa isoform consequent to a weakened exon 3 recognition. Hamid et al (23) reported a large pedigree with the c.172G>A splicing mutation, and they suggested that the ratio of 17.5 kDa/22 kDa transcripts in the lymphocytes correlated with the height SDS before GH replacement therapy. In this family, there were individuals with a height SDS more than −2 that inherited the mutation (incomplete penetrance). Another heterozygous missense mutation, c.200A>G, within ESE2 induces exon skipping in about 20% of the transcripts, giving rise to different phenotypes ranging from short stature to normal stature in the same large pedigrees (19).

It is thus conceivable that mutations associated with variable expressivity and incomplete penetrance might be responsible for at least some milder forms of IGHD.

The effect determined by the here detected variants, c.255G>A and c.272 A>T, is comparable with the effect caused by most of the previously described mutations falling within ESE1 and ESE2 (8, 19, 23). However, when these 2 variants were combined with c.261T, as in patients 1 and 2, the effect on splicing was maintained, although weaker, only for the 255A/261T construct (patient 2) (Figure 2B, lane 7). In contrast, the 261T/272T construct (patient 1) (Figure 2B, lane 8) was very similar to the wild-type, although the Western blotting for the same construct showed a slightly increased amount of the 17.5-kDa band (Figure 3, lane 7). This discrepancy might be attributable to the different sensitivity of the two methods.

Thus, the c.261 variant, which is not included in the putative ESE sites, partially hides the negative effect of the c.272. It can be speculated that c.261 influences splicing by strengthening the affinity for other proteins involved in the correct splicing regulation.

It is worth considering that the 233-bp band corresponding to the bioinactive 11.3 kDa (exon 3–4-skipped isoform) is strongly increased in all mutant constructs, but 261T and 261T/272T (Figure 2B). The relative total amount of the 2 alternatively spliced mRNA (exon 3 and exon 3–4 skipped) in the 255A, 272T, and 255A/261T constructs can be roughly estimated to represent nearly 50% of the total *GH1* mRNA in contrast to the wild-type, where these transcripts represent about 12% of the *GH1* transcripts.

The low serum GH level detected in vivo in the patients might be in part determined by the effect on splicing of these variants and in part by other yet unidentified genetic factors. Notably, patient 1 carried on the other allele the *GH1* promoter haplotype 1 (data not shown) that has been associated to IGHD and to a reduced luciferase activity in vitro (24).

The functional significance of the Glu82Asp variant carried by patient 3 is uncertain. From our experiments, it does not seem to influence splicing (Figure 2B, lane 3), and it is not predicted to have an effect on SR protein binding (data not shown). This variant is very rare, because it was reported in the dbSNP database (rs61762497) only in 1 individual of African ancestry, it is present in 2 individuals out of 13 000 of the Exome Variant Server, and it was absent in our panel of 205 normal stature individuals. However, it is not predicted to exert a damaging effect on the protein function by the software Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/). It might, thus, represent either a rare benign polymorphism, or alternatively it might contribute to a multigenic form of IGHD in this patient.

In conclusion, the analysis of *GH1* in individuals with sporadic IGHD led to the identification of 2 novel *GH1* exon 3 variations, c.255 G>A and c.272 A>T, included within 2 novel putative splicing regulatory elements that increase the aberrant splicing in vitro. When combined in

cis, with c.261 C>T, as in the patients, their effect was reduced but still evident on the protein synthesis. It can be hypothesized that also a minimally increased amount of the 17.5-kDa protein might exert a dominant negative effect on the GH synthesis in vivo. The phenotype of patients 1 and 2 might be associated to these *GH1* splicing variations that by themselves only partially influence the amount of GH secretion but that might act in concert with other genetic variants.

Acknowledgments

We thank the patients and parents that participated to this study.

Address all correspondence and requests for reprints to: Professor Mara Giordano, Department of Health Sciences, University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy. E-mail: mara.giordano@med.unipmn.it.

This work was supported by the Cariplo Foundation Grant Project 2009–2609.

Disclosure Summary: The authors have nothing to disclose.

References

1. **Chen EY, Liao YC, Smith DH, Barrera-Saldaña HA, Gelinas RE, Seeburg PH.** The human growth hormone locus: nucleotide sequence, biology, and evolution. *Genomics*. 1989;4:479–497.
2. **Mullis PE.** Genetics of isolated growth hormone deficiency. *J Clin Res Pediatr Endocrinol*. 2010;2:52–62.
3. **Procter AM, Phillips JA 3rd, Cooper DN.** The molecular genetics of growth hormone deficiency. *Hum Genet*. 1998;103:255–272.
4. **DeNoto FM, Moore DD, Goodman HM.** Human growth hormone DNA sequence and mRNA structure: possible alternative splicing. *Nucleic Acids Res*. 1981;9:3719–3730.
5. **De Vos AM, Ultsch M, Kossiakoff AA.** Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*. 1992;255:306–312.
6. **Palmetshofer A, Zechner D, Luger TA, Barta A.** Splicing variants of the human growth hormone mRNA: detection in pituitary, mononuclear cells and dermal fibroblasts. *Mol Cell Biol Endocrinol*. 1995;113:225–234.
7. **Moseley CT, Mullis PE, Prince MA, Phillips JA.** An exon splice enhancer mutation causes autosomal dominant GH deficiency. *J Clin Endocrinol Metab*. 2002;87:847–852.
8. **Petkovic V, Lochmatter D, Turton J, et al.** Exon splice enhancer mutation (GH-E32A) causes autosomal dominant growth hormone deficiency. *J Clin Endocrinol Metab*. 2007;92:4427–4435.
9. **Ryther RC, Flynt AS, Harris BD, Phillips JA 3rd, Patton JG.** GH1 splicing is regulated by multiple enhancers whose mutation produces a dominant-negative GH isoform that can be degraded by allele-specific small interfering RNA (siRNA). *Endocrinology*. 2004;145:2988–2996.
10. **Ryther RC, McGuinness LM, Phillips JA, et al.** Disruption of exon definition produces a dominant-negative growth hormone isoform that causes somatotroph death and IGHD II. *Hum Genet*. 2003;113:140–148.
11. **Takahashi I, Takahashi T, Komatsu M, Sato T, Takada G.** An exonic mutation of the GH-1 gene causing familial isolated growth hormone deficiency type II. *Clin Genet*. 2002;61:222–225.
12. **Lee MS, Wajnrajch MP, Kim SS, et al.** Autosomal dominant growth hormone (GH) deficiency type II: the Del32–71-GH deletion mutant suppresses secretion of wild-type GH. *Endocrinology*. 2000;141:883–890.
13. **McGuinness L, Magoulas C, Sesay AK, et al.** Autosomal dominant growth hormone deficiency disrupts secretory vesicles in vitro and in vivo in transgenic mice. *Endocrinology*. 2003;144:720–731.
14. **Tanner JM, Whitehouse RH.** Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Arch Dis Child*. 1976;51:170–179.
15. **Tanner JM, Whitehouse RH, Cameron N, Marshall WA, Healy WA, Goldstein H.** *Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method)*. San Diego, CA: Academic Press; 1988.
16. **Ghigo E, Bellone J, Aimaretti G, et al.** Reliability of provocative tests to assess growth hormone secretory status. Study in 472 normally growing children. *J Clin Endocrinol Metab*. 1996;81:3323–3327.
17. **Growth Hormone Research Society.** Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH Research Society. GH Research Society. *J Clin Endocrinol Metab*. 2000;85:3990–3993.
18. **Vivenza D, Guazzarotti L, Godi M, et al.** A novel deletion in the GH1 gene including the IVS3 branch site responsible for autosomal dominant isolated growth hormone deficiency. *J Clin Endocrinol Metab*. 2006;91:980–986.
19. **Millar DS, Lewis MD, Horan M, et al.** Novel mutations of the growth hormone 1 (GH1) gene disclosed by modulation of the clinical selection criteria for individuals with short stature. *Hum Mutat*. 2003;21:424–440.
20. **Giordano M, Marchetti C, Chiorboli E, Bona G, Momigliano Richiardi P.** Evidence for gene conversion in the generation of extensive polymorphism in the promoter of the growth hormone gene. *Hum Genet*. 1997;100:249–255.
21. **Wagner JK, Eblé A, Cogan JD, Prince MA, Phillips JA 3rd, Mullis PE.** Allelic variations in the human growth hormone-1 gene promoter of growth hormone-deficient patients and normal controls. *Eur J Endocrinol*. 1997;137:474–481.
22. **Solis AS, Peng R, Crawford JB, Phillips JA 3rd, Patton JG.** Growth hormone deficiency and splicing fidelity: two serine/arginine-rich proteins, ASF/SF2 and SC35, act antagonistically. *J Biol Chem*. 2008;283:23619–23626.
23. **Hamid R, Phillips JA 3rd, Holladay C, et al.** A molecular basis for variation in clinical severity of isolated growth hormone deficiency type II. *J Clin Endocrinol Metab*. 2009;94:4728–4473.
24. **Giordano M, Godi M, Mellone S, et al.** A functional common polymorphism in the vitamin D-responsive element of the GH1 promoter contributes to isolated growth hormone deficiency. *J Clin Endocrinol Metab*. 2008;93:1005–1012.

CHRONIC RENAL FAILURE OF UNKNOWN ORIGIN IS CAUSED BY *HNF1B* MUTATIONS IN 9% OF ADULT PATIENTS: A SINGLE CENTER COHORT ANALYSIS

Claudio Musetti[1], Marco Quaglia[1], Simona Mellone[2], Alessia Pagani[3], Ileana Fusco[2], Alice Monzani[3], Mara Giordano[2] and Piero Stratta[1]

1. Departments of Translational Medicine, Nephrology and Transplantation, Hospital "Maggiore della Carità", Novara (NO), Italy
2. Laboratory of Genetics, Department of Health Sciences and IRCAD University of Eastern Piedmont "Amedeo Avogadro", Italy
3. Division of Pediatrics, Department of Health Sciences, University of "Piemonte Orientale Amedeo Avogadro", Italy

Running Head: *HNF1B* mutations in adult CKD

Correspondence to

Claudio Musetti

Nephrology and Transplantation

Department of Translational Medicine

University "Amedeo Avogadro", Novara

Ospedale Maggiore della Carità,

Corso Mazzini 18, 28100 Novara

Tel 0039 0321 3733798

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/nep.12199

Fax 0039 0321 3733916

e-mail: claudio.musetti@med.unipmn.it

Abstract

Background. *HNF1B* gene mutations might be an underdiagnosed cause of nephropathy in adult patients mainly because of their pleomorphic clinical presentations. As most studies are based on paediatric populations, it is difficult to assess the likelihood of finding *HNF1B* mutations in adult patients and consequently define clinical settings in which genetic analysis is indicated. The aim of this study was the search for mutations in the *HNF1B* gene in a cohort of unrelated adult patients with nephropathy of unknown aetiology.

Methods Patients were tested for the *HNF1B* gene if they had chronic kidney disease of unknown origin and renal structure abnormalities (RSA) or a positive family history of nephropathy. The *HNF1B* coding sequence and intron-exon boundaries were analyzed by direct sequencing. The search for gene deletions was performed by Multiple Ligation Probe Analysis (MLPA).

Results Heterozygous mutations were identified in 6 out of 67 screened patients (9.0%) and included two whole gene deletions, one nonsense (p.Gln136Stop), two missense (p.Gly76Cys and p.Ala314Thr) mutations and a frameshift microdeletion (c.384_390 delCATGCAG), the latter two (c.384_390 del and p.Ala314Thr) not ever being reported to date. Mean age of the mutated patients at screening was 48.5 years with a M/F ratio of 2/4. The clinical manifestations of affected patients were extremely pleomorphic, including several urological and extra-renal manifestations.

Conclusions Mutations of *HNF1B* could explain chronic kidney disease in up to 9% of adult patients with a nephropathy of unknown aetiology and RSA: therefore an *HNF1B* mutation analysis should be considered in this group of patients.

Key Words: Chronic kidney disease End stage renal disease (ESRD) Hepatocyte nuclear factor-1ß gene (*HNF1B)* Hereditary cystic diseases Renal malformations

Introduction

Hepatocyte nuclear factor-1 beta (HNF1B) is a transcription factor involved in the tissue-specific regulation of gene expression and embryonic development of various organs, including liver, kidney, intestine, pancreas, and the genitourinary system[1]. It regulates the expression of proteins of the primary cilia of epithelial cells of all segments of the nephron and collecting duct and exerts a direct control on the expression of several key genes involved in cystic kidney diseases, including *Pkhd1*, *Pkd2* and *Umod*[2].

The HNF1B protein consists of a N-terminal dimerisation domain, a Pit-1/Oct-1/Unc-86 (POU) domain, an homeodomain that mediates DNA binding and a C-terminal transcriptional activation domain. HNF1B molecules bind to promoters of target genes as a homodimers or heterodimers and either activate or (more rarely) repress transcription[2].

Heterozygous point mutations and deletions of the *HNF1B* gene were initially identified in the autosomal dominantly inherited maturity onset diabetes of the young (MODY-5) associated with progressive renal dysfunction,[3,4] leading to the definition of "renal cysts and diabetes syndrome" (OMIM #137920). *HNF1B* mutations were subsequently recognized as a cause of renal disease even without MODY and were associated with a wide range of extra-renal manifestations such as liver and exocrine pancreas alterations and reproductive system malformations[5,6,7]. This pleomorphic clinical presentation is consistent with a broad and complex role of this transcription factor throughout development phases[1].

The clinical spectrum of phenotypes associated with *HNF1B* mutations is very broad, without any clear genotype-phenotype correlation. "*De novo*" mutations are found in

50% of patients and more than 40% of them are large genomic deletions. Interestingly, the *HNF1B* gene is located in a region flanked by segmental duplications[8] which are susceptible to unequal crossing over between them. Most of the point mutations are private, of which missense mutations account for about 50% and frameshift, nonsense and splicing variants for the remaining 10%[9].

An impressively high frequency of *HNF1B* mutations (up to 30%) was recognized in pediatric cohorts with renal structure abnormalities (hyperechogenic, multicystic dysplastic, hypo-dysplastic kidney, oligomeganephronia and horseshoe kidneys)[10], while a systematic review calculated a global detection rate of *HNF1B* anomalies of 9.7% including both paediatric and adult studies. The frequency of mutations varied considerably among different reports, ranging from 1.4% in patients with MODY to 21.4% in patients with renal structure anomalies (RSA) and to 41.2% if RSA and MODY were both present[11,12,13]. However, most studies were performed in paediatric cohorts, making it difficult to assess the likelihood of finding *HNF1B*mutations in adult patients with a nephropathy of unknown aetiology[11].

Aim of this study was to look for mutations in the *HNF1B* gene in a cohort of unrelated adult patients with a nephropathy of unknown aetiology and either renal structure abnormalities or a positive family history for nephropathy, in order to evaluate the relevance of *HNF1B* mutations in patients without a defined cause of chronic kidney disease (CKD).

Subjects and Methods

Patients

Adult patients eligible for enrollment included those who attended the Nephrology Outpatient Clinic of the tertiary care hospital "Maggiore della Carità" of Novara (Italy), those who received a kidney transplant at the same institution and those who were evaluated for admission in the kidney transplant waiting list (Figure 1).

Inclusion criteria for this study were the presence of a chronic kidney disease (CKD) of unknown etiology, defined by an estimated glomerular filtration rate by CKD-EPI formula (eGFR), less than 60 ml/min/1.73m^2, and either a positive family history for renal disease or any renal structure abnormality (RSA), defined as renal cysts (RC), renal dysplasia, diffuse hyperecogenic kidneys (HK), uni- or bi-lateral multicystic kidney disease (MCD), cystic dysplasia or gross renal developmental abnormalities (single kidney, horseshoe kidney, renal hypoplasia).

Exclusion criteria included the diagnosis (either clinical or genetic) of another genetic disease that causes RSA, including autosomal recessive polycystic kidney disease, autosomal dominant polycystic kidney disease, nephronophthisis, UMOD-associated kidney disease, papillorenal syndrome and branchio-oto-renal syndrome. Moreover patients were excluded if they were diagnosed with a primary or secondary glomerulonephritis or showed signs of glomerular disease, including nephritic syndrome, nephrotic syndrome or a proteinuria greater than 1 gr/day confirmed at two separate measurements.

A written informed consent was obtained from all patients and local Ethical Committee approved the study.

Genetic analysis of HNF1B gene

Genomic DNA was extracted from lymphocytes using a QIAamp DNA Kit (Qiagen, Hilden, Germany). The entire coding region of the HNF1B gene and intron-exon boundaries were amplified by PCR. PCR conditions and primer sequences are available upon request. The PCR products were visualized on a 2% agarose gel and purified using affymetrix ExoSAP-IT enzymatic PCR clean up system. The purified products were then sequenced with Big Dye Terminator kit (Applied Biosystems, Foster City, CA) and automatic sequencer ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA).

Search for deletions of the single exons and of the entire gene was performed by multiplex ligation-dependent probe amplification (MLPA assay) using a MLPA commercial kit (SALSA MLPA KIT P018-E1 SHOX, MRC-HOLLAND MCR-Holland) following the manufacturer's instructions.

Results

Patient characteristics

A total of 67 adult patients referring to our centre over a 4-year period (2009-2012) fulfilled the inclusion criteria for the study of the *HNF1B*gene. Median age of the patients at time of genetic analysis was 51 years (49.6 ± 17.3 years); M/F ratio was 31:36 (Male 46.3%).

All patients had CKD of unknown etiology (on dialysis or KTR: 27/67 = 40.3%) with or without laboratory signs of tubular dysfunction (22/67 = 32.8%), defined biochemically as renal tubular acidosis, hypercalciuric hypocalcemia (without hyperparathyroidism) or renal salt or magnesium wasting. Moreover specific indications for the genetic analysis were RSA (58/67 = 86.6%), mainly represented by cysts (43 of 58 patients with RSA, 74.1%) and/or a positive family history of renal disease (33/67 = 49.3%). Eighteen patients (26.9%) had also other relevant extra-renal manifestations that have been described in association with *HNF1B* mutations, mainly represented by gout (10/18 = 55.6%), and genital abnormalities (6/18 = 33.3%). Seven patients (7/67=10.4%) had type 2 diabetes or a new onset diabetes after transplant (NODAT), diagnosed at an age of 45±2.8 years in mutated patients and 50.2±4.8 years in non-mutated patients. However none of the patients –either mutated or non- mutated- was diagnosed as having a MODY.

Among patients that turned out to carry an *HNF1B* mutation, CKD was often associated with low-grade proteinuria (<0,5 g/day), biochemical signs of tubular disease (renal magnesium wasting, hypercalciuric hypocalcemia or hypokaliemia), and RSA, including renal cortical cysts (CC) and/or medullary cysts (MC) in 3/6 patients. As for extrarenal phenotype, type 2 diabetes was present in 2/6 patients, biochemical

alterations of liver or exocrine pancreas exams in 4/6 patients, genital tract abnormalities in 2/6 patients, a thyroid disorder in 2/6 patients (Table 1).

HNF1B molecular analysis

Heterozigous mutations were identified in six unrelated individuals leading to a mutation detection rate of 8.96%. The mutations included two deletions of the entire gene, a frameshift micro-deletion (c.384_390 delCATGCAG) leading to a premature stop codon, two missense (p.Gly76Cys and p.Ala314Thr) and one nonsense (p.Gln136Stop) mutation. Two of these alterations had never been reported to date, namely c.384_390 del CATGCAG and p.Ala314Thr, whereas the others have already been identified in patients with renal disease[9].

Among the 6 mutated patients, four (66,6%) had at least one first-degree relative with nephropathy and/or extrarenal manifestation compatible with an *HNF1B* related disorder (Fig. 2). The segregation analysis performed on the relatives of cases #1 and #3 demonstrated that the mutations arose *de novo* in the probands. In case #1 the *"de novo"* mutation in the proband (II-1) was transmitted to both her siblings (Fig. 2).

Case #1 (GM, female)

The proband was a 17-year-old girl with CKD with small-sized hyperhecogenic kidneys, little medullary cysts (up to 10 mm), mild CKD (serum creatinine 1.4 mg/dL, eGFR = 52 ml/min/1.73m^2), low grade proteinuria (0.4 g/day) and hypomagnesaemia. Among extrarenal manifestations which have been rarely associated with *HNF1B* mutations[9], the proband had a polyclonal hypergammaglobulinemia and an Hashimoto thyroiditis.

The proband had two daughters (Fig. 2), the first one was born in 2007 with impaired renal function and multiple congenital abnormalities (ectopic anal opening, pilonidal dimple, occult spinal disraphism, strabismus) including bilateral hyperechogenic kidneys (HK) with multiple cortical cysts (CC). She later developed juvenile rheumatoid arthritis. The second daughter, born in 2011, had diffuse HK with some small CC and CKD (sCr = 0.73 mg/dl). She developed a complex neuropsychiatric disorder with motility impairment and slight mental retardation and was diagnosed with celiac disease and Hashimoto thyroiditis.

The analysis of *HNF1B* gene showed an heterozygous 7 bp deletion within exon 2 from nucleotide c.384 to c.390 (c.384_390delCATGCAG, considering as c.1 the first nucleotide of the ATG initiation codon) in the proband and in her two daughters This mutation causes a frameshift from codon 129 within the POU specific domain and it is predicted to result in a premature stop 14 aminoacid downstream. The putative protein, if translated, lacks most of the DNA binding domain extending from aminoacid 88 to 313 and the entire transactivation domain (aminoacid 314-557). Absence of the mutation in either the proband's parents revealed that the it arose de novo in the proband.

Case # 2 (GL, female)

The proband had frequent episodes of renal colic ever since her youth. At the age of 40, she was diagnosed with Rokitansky syndrome (primary amenorrhea, underdevelopment of vagina and uterus, normal secondary sexual characteristics), type 2 diabetes, obesity and CKD (serum creatinine 1.3 mg/dL, eGFR = 40 ml/min/1.73m^2), with a tendency to hypokaliemia and magnesium renal loss (magnesiuria 250-300 mg/day). She was diagnosed with a hypoplasic right kidney, bilateral HK and multiple bilateral CC (1-3

cm) at the age of 72. Abdominal MRI showed dilatation of the intrahepatic biliary tree and coledocus, without any stone.

The analysis of *HNF1B* genes showed the presence of the entire gene deletion at the heterozygous state. Both the brother and the deceased mother suffered from chronic nephropathy of unknown etiology (Fig. 2), suggesting a dominant transmission of the mutation. However the proband's brother didn't give his consent to genetic analysis.

Case #3 (FE, female)

The proband was diagnosed with bilateral kidney hypoplasia and CKD at the age of 4, but she did not undergo any genetic screening for CKD or RSA. She underwent thyroidectomy due to thyroid oncocytoma when she was 23. At the age of 31, she reached ESRD and during the pre-transplant work-up a gastric microcarcinoid and a breast cancer were diagnosed.

Analysis of *HNF1B* gene revealed a mutation at the heterozygous state, c.406C>T, leading to the creation of a premature stop codon at position 136 (p.Gln136Stop) and to a putative protein lacking most of the DNA binding domain. The mutation arose de novo in the proband as it was absent in both her parents. The same mutation had been already reported in two unrelated patients[14,15] and might thus represent a mutational hot-spot.

Case #4 (AE, female)

The proband had recurrent urinary tract infections leading to a diagnosis of chronic pyelonephritis secondary to bilateral pyelo-uretheral junction stenosis, which was never corrected. CKD was diagnosed at the age of 47 (serum creatinine 1.7 mg/dL, eGFR = 33 ml/min/1.73m^2) along with type 2 diabetes and hypouricuric hyperuricemia. Renal imaging showed bilateral hypercogenicity, some scars and persistent pielectasy.

Molecular analysis of *HNF1B* gene showed an heterozygous deletion of the whole gene. As the proband's deceased mother had a solitary kidney and developed ERSD it is likely that she transmitted the mutation although her DNA was not available.

Case #5 (BM, male)

The proband was diagnosed with CKD when he was 30-years old, together with urethral stenosis, left hydrocele and a III degree bilateral vesico-uretheral reflux. He started hemodialysis at 48 and received a kidney transplant at the age of 50. During the pre-transplant work up a chronic increase in pancreatic enzymes (PE) was investigated with CT and MR without any pathologic finding.

The analysis of *HNF1B* gene showed the heterozygous mutation c.G226T within exon 1, leading to the aminoacidic substitution p.Gly76Cys, a highly conserved residue located between the dimerization and the DNA binding domains. This mutation has been already reported in other three unrelated patients[14,15]. In the Exome Variant Server (http://evs.gs.washington.edu/EVS/), a genetic database collecting the information on the variants detected by sequencing the exomes of about 13.000 individuals of European and African ancestry, this mutation is reported in 5 out of 12.999 individuals (0.04%). This database contains controls but also the extremes of specific traits (LDL and blood pressure), and diseases (such as cardiovascular diseases and lung diseases) and it is not possible to obtain information about the phenotypes of individuals with any specific variants. However considering that Gly76Cy has been previously reported in 3 out of 1406 unrelated patients[9] and presently in 1 out of 67 patients (for a total of 4/1473), the frequency is significantly higher in patients with renal phenotypes (p= 0.008) with respect to a population of mixed phenotypes. It is thus conceivable that this represents a causative mutation that might be characterized by incomplete penetrance.

Case #6 (SG, male)

The proband had positive family history for nephropathy (the mother and the maternal uncle were diagnosed with CKD of unknown etiology) and was affected by CKD since he was 10-years old. He developed low-grade tubular proteinuria (< 1 g/day) and a tendency to chronic hypokalemia and hypomagnesemia. An IgG-kappa monoclonal gammopathy was diagnosed when he was 38 and then developed Bence-Jones proteinuria (kappa light chains). Renal ultrasound revealed a hypoplasic and ectopic left kidney with small parapyelic cysts and a single CC (4 cm). He started peritoneal dialysis at 47 and received renal transplant at 50. In the post-transplant period persistent increase of amylase and lipase was noted, but a CT study showed no evidence of anatomical alterations in pancreas or liver.

The analysis of the *HNF1B* gene revealed the presence of the missense mutation c.940G>A leading to the aminoacid substitution Ala314Thr. This represents a novel mutation, which has been never detected in the 13000 genomes of the Exome Variant Server (http://evs.gs.washington.edu/EVS/). It falls within the transactivation domain at a position which is highly conserved among species and it is predicted to exert a damaging effect on the protein function by the software Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/). For these reasons Ala314Thr is likely to represent a pathogenic mutation although the co-segregation with renal disease within the family could not be performed as the proband's relatives didn't give their consent to genetic testing. Genetic analysis was only performed in the 19 year- old healthy son who resulted negative for the father's mutation.

Discussion

In this study we demonstrated a relatively high prevalence (9%) of *HNF1B* mutations in adult patients with CKD of unknown etiology and either RSA and/or a positive family history of renal disease. Overall *HNF1B* mutations are a rare cause of CKD: out of 877 patients who were referred to our Nephrology Clinic/Transplant Centre during the study period, only six (0.68%) carried an *HNF1B* mutation. However the detection rate of *HNF1B* mutations could be increased adopting the selection criteria used in this study.

The mutation detection rate is comparable to that found in other cohorts of patients. In a recent systematic review of the literature taking into account 15 different studies, Chen et al.[9] reported a global detection rate of *HNF1B* anomalies of 9.7% with the highest frequency in patients with both MODY and renal phenotypes (41.2%).

With respect to previous adult studies, which mainly focused on renal structure abnormalities[9,11,13] or diabetes[9,16,17], we considered as a mandatory inclusion criteria the presence of a moderate or severe CKD of unknown aetiology associated with either RSA or a positive family history of renal disease. Even if CKD is described in up to 55.5% of patients with a renal phenotype[9,13], the relatively high detection rate found in the present study underlines the clinical relevance of this disease in the setting of CKD, predialysis and ESRD care of adult patients.

Considering the probands, a positive family history was reported in 4 out of 6 mutated patients (Cases #1,2,4,5) of which one "*de novo*" mutation which was transmitted to the offspring (Case #1) and a "*de novo"* mutation has been demonstrated to occur in one case (#3).

In general *"de novo"* mutations are reported in a considerable proportion (50%) of patients and more than 40% are whole gene deletions, suggesting an intrinsic instability

of *HNF1B* locus[18]. In our patients mutations which arose "*de novo"* were the 7 bp deletion (c.384_390 del CATGCAG) and a missense mutation leading to a premature stop codon (p.Gln136Stop). The former has never been described to date and causes a premature stop codon (at position 143 of the putative protein) within the DNA-binding domain leading to a protein lacking most of the DNA binding and the whole transactivation domains. As previously demonstrated for other truncating mutations within this domain, even if the ability to form hetero/homodimer is retained, the mutant protein might be unable to bind DNA[19]. Another possible consequence is nonsense-mediated mRNA decay (NMD), a surveillance mechanism that removes transcripts with termination signals[20]. The latter *"de novo"* mutation is a stop codon within the same domain (the DNA binding domain) that it is likely to exert effect on the protein functioning similar to that of c.384_390 del CATGCAG. The p.Gln136Stop mutation has been previously described in two unrelated families[14,15] and it is thus likely to represent a mutation hot spot.

It is widely reported that there is no a clear phenotype-genotype correlation in patients harbouring *HNF1B* mutations; on the other hand this approach would not be feasible in such a small sample of patients. Some of the here reported patients might exemplify the extremely variable phenotypes associated to *HNF1B* mutations. In particular the relatives of case #1 exhibited a high degree of variable expression within the same pedigree. The proband (GM) had a mild CKD with little medullary cysts; her first daughter had a complex picture of congenital abnormalities and CKD with bilateral renal dysplasia and multiple CC; the second daughter had only CKD, but no extra-renal morphological abnormalities. Interestingly, the mother and the two daughters had autoimmune disorders: while thyroiditis (mother and second daughter) has been rarely

reported in the patients with *HNF1B* mutations[9], juvenile rheumatoid arthritis (first daughter) and celiac disease (second daughter) have never been related to these genetic alterations. The possible mechanisms linking autoimmunity and *HNF1B* mutations are not known and might deserve further *in vitro* studies.

This study is entirely based on adult patients, providing a peculiar prospective on the evolution of this nephropathy: mean age at screening was 48 years and three out of six positive patients had ESRD. Typical but subtle signs of this condition had been present for many years in all of these cases and delayed diagnosis could have been avoided if a higher index of suspicion had been adopted[16,21]. The misleading concept that relegates *HNF1B* related disorders to the paediatric setting is probably contributing to a significant under-diagnosis of this condition. It is likely that not all patients with RSA diagnosed during infancy have been tested for genes involved in RSA and some patients may not draw medical attention until they develop CKD in their adult life.

The evolution to ESRD has been described in a variable percentage of cases carrying *HNF1B* mutations, ranging from 3%[4] to 15%[11]; CKD was already present in more than 90% of patients in their third decade of age in the latter series. In our study CKD was present in all patients at time of diagnosis: it was diagnosed at a mean age of 24.6 years (4-47 years) and progressed to ESRD by the age of 42 years (31-48) in 3 out of 6 patients. Therefore, although usually characterised by a slow decline of renal function, *HNF1B* related nephropathy can cause ESRD at a young age.

Notably, CKD and RSA were both present in all the patients carrying *HNF1B* anomalies, highlighting the relevance of performing *HNF1B* molecular analysis in adult patients with nephropathy of unknown aetiology and RSA. Extrarenal manifestations already described as associated to *HNF1B* mutations were present in 4/6

of mutated patients (Cases# 1, 2, 3 6), while MODY should not be considered as an essential element for the molecular study of *HNF1B*, because, at least in our experience, it was absent in all positive patients.

Another equally challenging subset of pleomorphic manifestations is represented by the wide spectrum of urological and reproductive abnormalities[8,19,22]. While presence of RSA such as cysts and hypodisplasia are hallmarks of this disorder, most uretheral abnormalities can pass unnoticed: in our study one patient (Case #4) had a bilateral pyelo-uretheral junction stenosis while another one (Case #5) had a urethral stenosis, left hydrocele and bilateral vesico-uretheral reflux.

The main limit of this study lies in the fact that it was centre rather than population-based and patients were referred on the basis of a set of phenotypic features which always included CKD. As a consequence, it cannot be used to infer prevalence of this genetic disorder in the general population with undetected nephropathy or without renal manifestations. However, it probably provides a realistic picture of several clinical settings in which genetic testing should be considered in an adult nephropathic patient.

In conclusion, we found a prevalence of *HNF1B* gene mutations of 9% in an adult population of patients with CKD of unknown aetiology, who were selected for genetic analysis mainly based on the presence of either RSA or family history of renal disease: therefore *HNF1B* mutation analysis should be considered by the adult nephrologist when dealing with a CKD of unknown origin, especially if cysts or other key-features are present. However future studies on cost-effectiveness of screening programs for *HNF1B* mutation in adults with CKD of unknown aetiology are needed.

References

1) Igarashi P, Shao X, McNally BT, Hiesberger T. Roles of HNF-1beta in kidney development and congenital cystic diseases. Kidney Int. 2005;68(5):1944-7.

2) Hiesberger T, Shao X, Gourley E, Reimann A, Pontoglio M, Igarashi P. Role of the hepatocyte nuclear factor-1beta (HNF-1beta) C-terminal domain in Pkhd1 (ARPKD) gene transcription and renal cystogenesis. J Biol Chem. 2005;280(11):10578-86.

3) Fajans SS, Bell GI, Polonsky KS. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. N Engl J Med 2001; 345: 971-980.

4) Bellanné-Chantelot C, Clauin S, Chauveau D, Collin P, Daumont M, Douillard C et al. Large genomic rearrangements in the hepatocyte nuclear factor-1beta (TCF2) gene are the most frequent cause of maturity-onset diabetes of the young type 5. Diabetes. 2005 ; 54: 3126-32.

5) Bellanne-Chantelot C, Chauveau D, Gautier JF. Clinical spectrum associated with hepatocyte nuclear factor-1beta mutations. Ann Int Med 2004; 140: 510-517.

6) Bingham C, Hattersley AT. Renal cysts and diabetes syndrome resulting from mutations in hepatocyte nuclear factor-1beta. Nephrol Dial Transplant 2004; 19: 2703-2708.

7) Thomas CP, Erlandson JC, Edghill EL, Hattersley AT, Stolpen AH A genetic syndrome of chronic renal failure with multiple renal cysts and early onset diabetes. Kidney Int. 2008 Oct;74(8):1094-9.

8) Raile K, Klopocki E, Holder M, Wessel T, Galler A, Deiss D, Müller D, Riebel T, Horn D, Maringa M, Weber J, Ullmann R, Grüters A. Expanded clinical spectrum in

hepatocyte nuclear factor 1b-maturity-onset diabetes of the young. J Clin Endocrinol Metab. 2009 Jul;94(7):2658-64. 20)

9) Chen Y, Gao Q, Zhao X, Chen Y, Craig LB, Xiong X, Mei C, Shi Y, Chen Y. Systematic review of TCF2 anomalies in renal cysts and diabetes syndrome/maturity onset diabetes of the young type 5. Chin Med J 2010; 123(22): 3326-3333.

10) Zaffanello M, Brugnara M, Franchini M, Fanos V. TCF2 gene mutation leads to nephro-urological defects of unequal severity: an open question. Med Sci Monit. 2008 Jun;14(6):RA78-86.

11) Faguer S, Decramer S, Chassaing N, Bellanné-Chantelot C, Calvas P, Beaufils S, Bessenay L, Lengelé JP, Dahan K, Ronco P, Devuyst O, Chauveau D. Diagnosis, management and prognosis of HNF1B nephropathy in adulthood. Kidney Int. 2011 Oct;80(7):768-76.

12) Edghill EL, Oram RA, Owens M, Stals KL, Harries LW, Hattersley AT, Ellard S, Bingham C. Hepatocyte nuclear factor-1 beta gene deletions – a common cause of renal disease. Nephrol Dial Transplant 2008; 23: 627-635.

13) Heidet L, Decramer S, Pawtowski A, Morinière V, Bandin F, Knebelmann B, Lebre AS, Faguer S, Guigonis V, Antignac C, Salomon R. Spectrum of HNF1B mutations in a large cohort of patients who harbor renal diseases. Clin J Am Soc Nephrol. 2010;5(6):1079-90.

14) Ulinski T, Lescure S, Beaufils S, Guigonis V, Decramer S, Morin D, et al. Renal phenotypes related to hepatocyte nuclear factor-1beta (TCF2) mutations in a pediatric cohort. J Am Soc Nephrol 2006; 17: 497-503.

15) Bellanne-Chantelot C, Clauin S, Chauveau D, Collin P, Daumont M, Douillard C, et al. Large genomic rearrangements in the hepatocyte nuclear factor-1beta (TCF2) gene

are the most frequent cause of maturity-onset diabetes of the young type 5. Diabetes 2005; 54: 3126-3132.

16) Adalat S, et al. HNF1B mutations associate with hypomagnesemia and renal magnesium wasting. J Am Soc Nephrol. 2009 May;20(5):1123-31.

17) Edghill EL, Stals K, Oram RA, Shepherd MH, Hattersley AT, Ellard S. HNF1B deletions in patients with young-onset diabetes but no known renal disease. Diabet Med. 2013 Jan;30(1):114-7.

18) Ma Z, Gong Y, Patel V, Karner CM, Fischer E, Hiesberger T, Carroll TJ, Pontoglio M, Igarashi, P. Mutations of HNF-1-beta inhibit epithelial morphogenesis through dysregulation of SOCS-3. Proc Nat Acad Sci 2007;104: 20386-20391.

19) Nakayama M, Nozu K, Goto Y, Kamei K, Ito S, Sato H, Emi M, Nakanishi K, Tsuchiya S, Iijima K. HNF1B alterations associated with congenital anomalies of the kidney and urinary tract. Pediatr Nephrol. 2010 Jun;25(6):1073-79.

20) Shyu AB, Wilkinson MF, van Hoof A. Messenger RNA regulation: to translate or to degrade. EMBO Journal. 2008; 27: 471- 81.

21) Edghill EL, Bingham C, Ellard S, Hattersley AT. Mutation in hepatocyte nuclear factor-1b and their related phenotypes. J Med Genet 2006; 43: 84-90.

22) Thomas R, Sanna-Cherchi S, Warady BA, Furth SL, Kaskel FJ, Gharavi AG. HNF1B and PAX2 mutations are a common cause of renal hypodysplasia in the CKiD cohort. Pediatr Nephrol. 2011 Jun;26(6):897-903.

Table 1. Extra-renal phenotype at diagnosis in patients with *HNF1-β* mutation

Patient	Diabetes (age at diagnosis)	Pancreas abnormalities	Liver abnormalities	Genital Tract abnormalities	Other
Pt 1 GM†	Absent	Absent	Absent	Absent	Chronic tyroidits
Pt 2 GL	Present (43)	Absent	Chronic cholestasis	Rokitansky syndrome	Absent
Pt 3 FE	Absent	Absent	Absent	Absent	Thyroid oncocytoma; gastric carcinoid
Pt 4 AE	Present (47)	Absent	Chronic steatohepatitis‡	Absent	Absent
Pt 5 SG	Absent	Chronic PE increase	Absent	Absent	Absent
Pt 6 MB	Absent	Chronic PE increase	Absent	Urethral stenosis; right hydroceles;	Absent

Pt: patients; PE = pancreatic enzymes

† Several abnormalities were found in one of the two daughters, who inherited the same mutation, and led to genetic analysis: occult spinal dysraphism, ectopic anal opening, strabismus

‡ defined by ultrasound imaging, with normal liver function and enzymes

Legend to figures

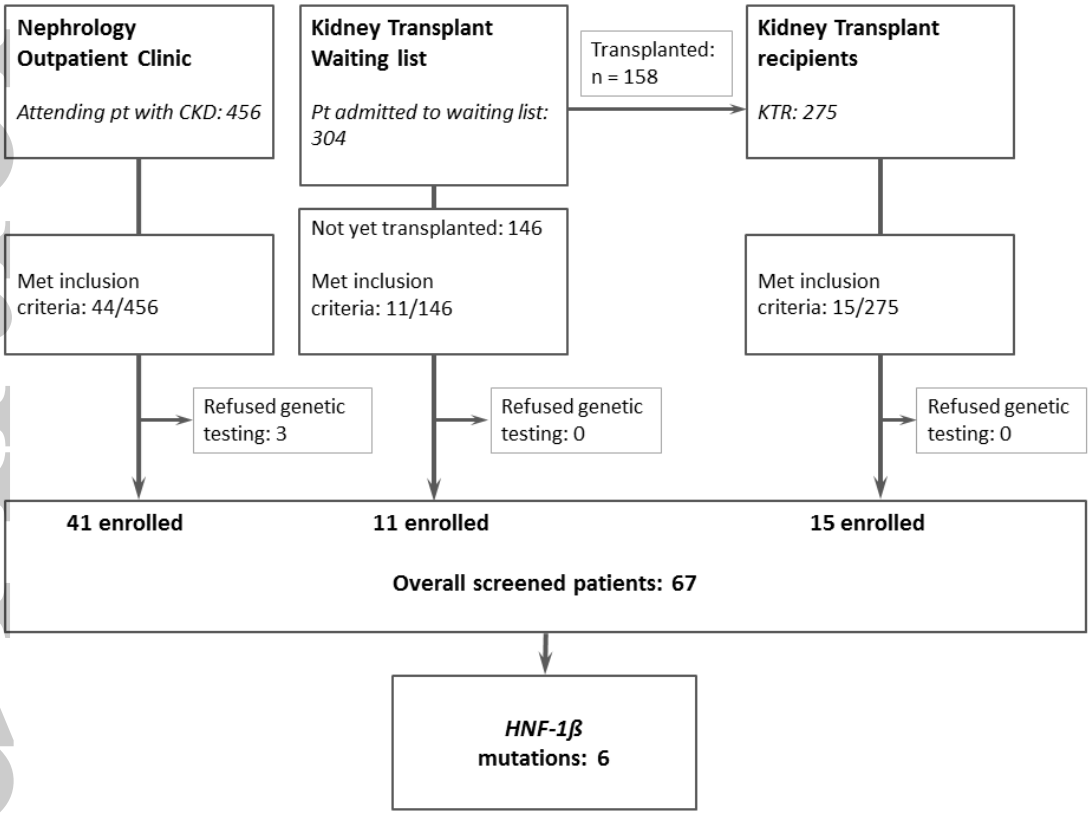

Figure 1. Patient enrollment flowchart. Patients attending our nephrology outpatient clinic, evaluated for admission in the kidney transplantation waiting list and who have been transplanted in the study period (January 2009 - December 2012) were clinically evaluated for the presence of chronic kidney disease, RSA and family history. Patients who were put in the transplant wait-list and then transplanted have been included only in the KTR group. Patients who met the inclusion criteria have been screened for *HNF1-beta* mutations (n= 67). Pt: patients. KTR: kidney transplant recipients. RSA: renal structural abnormalities

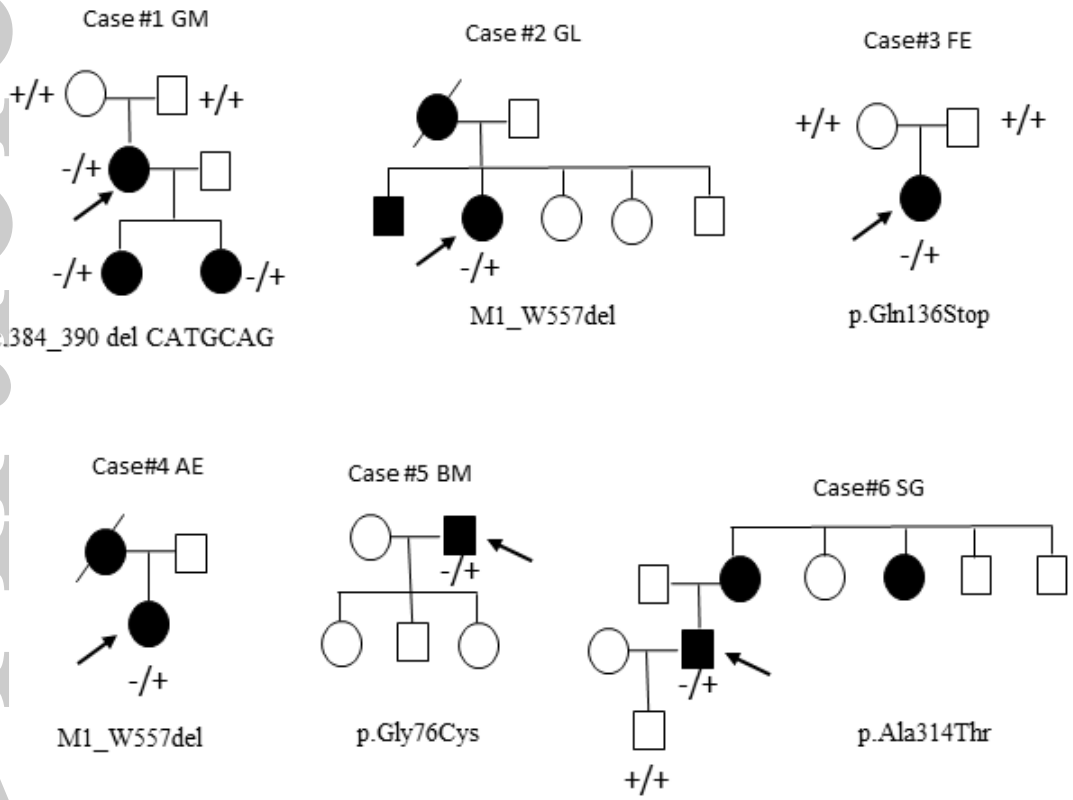

Figure 2. Overview of pedigrees of probands families. Affected subjects are indicated by filled symbols, healthy subjects by open symbols and probands by an arrow. HNF-1β genotype is reported for each individual (+ = wild type allele, - = mutated allele).