UNIVERSITÀ DEGLI STUDI DI GENOVA

Scuola di Scienze Mediche e farmaceutiche

Corso di Laurea in Medicina e Chirurgia

PhD Thesis

Autism Spectrum Disorder and other Neurodevelopmental Disorders: cytogenetic and genomic approaches

Ph.D. Course

NEUROSCIENCE

Curriculum

CLINICAL and EXPERIMENTAL NEUROSCIENCES

Cycle

XXXIII (2017-2021)

Candidate:

Dr. Livia Pisciotta, MD

Tutors:

Prof. Lino Nobili

Prof. Elisa De Grandis

Co-Tutor:

Prof. Aldamaria Puliti

Coordinators: Prof. Angelo Schenone, Prof. Tullio Florio

Summary

0. ABSTRACT

Introduction: Neurodevelopmental disorders (NDDs) are a heterogeneous class of conditions involving the brain, including intellectual disability (ID) and autism spectrum disorder (ASD), that affect about 1%-3% of children (Miller et al., 2010). The genetics of NDDs is complex and include copy number variations (CNVs), pathogenetic mutations in single genes. To date, more than 1000 genes have been implicated in the etiopathogenesis of NNDs. Preliminary investigations have suggested that the majority of Developmental Disorders, in particular ASD, are actually polygenic; in addition, the genetic and environmental interplay in defining the phenotype clearly classifies NDDs such as ID and ASD as complex disorders. In this dissertation, I sought to explore the contribution of rare de novo and inherited coding variation in neurodevelopmental disorders and use these genetic variations to identify neurodevelopmental disorder associated genes and new/unknown oligogenic mechanisms.

Methods: In a retrospective review of data, we re-evaluated all the results of diagnostic array-CGH tests on 700 cases with NDDs, focusing on variants previously interpreted as VOUS. Furthermore a series of 68 patients with autism spectrum disorder were recruited to perform whole exome sequencing and eventual whole genome sequencing. A deep analysis of VOUS, mainly consisting in a revision of gene expression/function annotation, and chromatine organization data, was performed. New candidate genes were analysed by GeneCodis4 to evidence enrichment for known NDD-associated GeneOntology terms and pathways. Whole exome sequencing was performed and potentially deleterious variants prioritized by custom filtering strategies including the use of ORVAL (Oligogenic Resource for Variant Analysis Platform) and enrichment analysis of candidate genes with GeneCodis4.

Results: In about 42% of cases pathogenic CNVs were found, while in 58% identified CNVs remained initially VOUS. New potential genes and mechanisms such as double-hit mechanisms were found in our patients. In our 34 analysed ASD patients 11 cases showed possible deleterious rare variants, in different and, in the majority of cases, in multiple genes. The role of X chromosome and neurotransmitter pathways appears important.

Conclusion: In our cohort of NDDs patients CNV-mediated double-hit mechanisms seem to play a relevant role in elucidate complex phenotypes. About 10% of patients from our ASD cohort also showed rare deleterious variants in multiple genes that seem to fully explain their complex phenotype.

1. NEURODEVELOPMENTAL DISORDERS

1.1. INTRODUCTION

Neurodevelopmental disorders (NDDs) are conditions characterized by a primary dysfunction of the neurological system and brain. This definition puts together a very wide range of neurological and psychiatric problems that are clinically and causally different.

"Developmental disorders" were nominated for the first time in DSM-III[1], for the category that included autistic disorder. "Neurodevelopmental disorders" (NDDs) were introduced as a new disorder category in DSM-5[2]. This replaced a previous chapter "Disorders usually first diagnosed in infancy, childhood, or adolescence." In ICD-11[3], the latest revision of the International Classification of Diseases published by the WHO, NDDs have become an integral part of: "Mental, behavioural or neurodevelopmental disorders" chapter.

In this manuscript, I adopt the approach of DSM-5. Then NDDs include intellectual disability (ID), communication disorder, autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), learning disabilities and motor/coordination disorder[2].

Children with these disabilities can experience difficulties with language, motor skills, behaviour, memory, executive functions, learning, or other neurological functions.

One of the key defining characteristics of these neurodevelopmental disorders is that they typically onset in childhood, before puberty.

The symptoms and behaviours of these conditions often change or evolve during evolutive age, but some impairments are permanent and constitutive. The level of overlap between these problems and their constituent symptoms is high, there has been a general trend to move away from the classification of single disorder as discrete entity, to placing it within a spectrum.

Some researchers have stated that the prevalence of certain neurodevelopmental disabilities, specifically ASD and ADHD, has been increasing over the last four decades[4567].

The percentage of intellectual disability diagnosed in children fluctuated between 0.6% and 0.9% from 1997 to 2010, and was between 1.3% and 1.4% from 2011 to 2013[7]. In 2010–2013, the number of boys with intellectual disability (1.6%) was higher than for girls (0.8%).

In 2013, 8.2% of children (5-17 years) had a diagnosis of a learning disability. There was only a little change in this percentage between 1997 and 2013[7]. For the years 2010–2013, the number of boys reported to have a learning disability (10.4%) was higher than for girls (6.6%).

From 1997 to 2013, the percentual of children (5-17 years) diagnosed with attention-deficit/hyperactivity disorder (ADHD) increased from 6.3% in 1993 to 10.7% in 2012 and 9.9% in 2013[7]. This rate was higher for the boys reported to have ADHD (13.7%) than the girls (6.0%).

The percentage of children (5 - 17 years) with diagnosis of autism rose from 0.1% in 1997 to 1.2% in 2013[7]. This increasing trend was statistically significant. The rate of this diagnosis was more than 4 times higher in boys than in girls (1.9% versus 0.4%).

Some experts suggested the DSM-5 criteria for Autism Spectrum Disorders require a higher threshold of symptoms[8]. Previous studies based on clinical or research samples in fact have reported that a proportion of individuals who meet DSM-IV-TR[9] criteria for ASD fail to meet the DSM-5 criteria. Despite this, the prevalence of the disorder from 2013 to today has markedly increased.

In 2020, the CDC reported that approximately 1 in 54 children in the U.S. is diagnosed with an autism spectrum disorder (ASD), according to 2016 data[10]: 1 in 34 boys and 1 in 144 girls identified with autism. Also in Europe, a new report on autism in U.K. schools indicates sharp increases in prevalence rates over a nine-year period through 2019[11].

Northern Ireland had the highest prevalence throughout, reaching 3.20% of all students by 2018/19, while the lowest rate in 2019 was 1.92% in Wales. England, with the largest population of school children (8,180,469 in 2019) had a rate of 2.25%.

Overall NNDs have an estimated prevalence of approximately 1–3% in the general population and according to the WHO about one in 270 people has an ASD.

The substantial increase in prevalence estimates is probably due to the change in diagnostic criteria and tools in recent years, accompanied by a greater availability of services and social sensitivity with respect to this condition.

The same genetic and early environmental risk factors that are associated with a diagnosis of ADHD or autism spectrum disorder can predict trait levels in the general population[12]. Sub-threshold diagnoses (in people with traits and so insufficient symptoms to have a diagnosis, but often with impairments) are common and are, in a sense, clinically important in terms of prognosis and early treatment. Maybe this tendency to diagnose these disorders more easily may partly justify their increase.

Most neurodevelopmental disorders have complex and multiple contributors rather than any one clear cause, in fact these disorders likely result from a combination of genetic, biological, psycho-social, and environmental risk factors. A broad range of environmental risk factors may affect neurodevelopment, including maternal abuse of alcohol, tobacco or drugs during pregnancy, preterm birth in particular with low birthweight, prenatal or childhood exposure to some environmental contaminants[13][14][15].

Genetics can play an important role in many neurodevelopmental disorders, and some cases of these are associated with specific genes.

Some neuropsychiatric disorders like intellectual disability and autism spectrum disorder are genetically heterogeneous and frequently caused by rare de novo mutations.

1.1.1. INTELLECTUAL DISABILITY

The most commonly used definitions of Intellectual Disability (also referred to as mental retardation) emphasize subaverage intellectual functioning before the age of 18, usually defined as an IQ less than 75-70 and impairments in adaptive skills such as communication, self-care, home living, and social or interpersonal relationships[2]. Different severity categories, ranging from mild to severe retardation, are defined based on IQ scores. According to the American Psychiatric Association, 1% of the population have intellectual disability. Around 85% of these people have mild cases. Researchers have identified some causes of ID, including genetic disorders, traumatic perinatal and/or prenatal events such as maternal infection or exposure to alcohol.

However, the causes of intellectual disability are unknown in 30–50% of all cases[16]. The causes are more frequently identified for cases of severe retardation (IQ less than 50), whereas the cause of milder retardation (IQ between 50 and 70) is unknown in more than 75% of cases[17].

1.1.2. SPEECH DISORDER

Diagnostic categories for communication disorders in DSM 5[2] include Language Disorder (language delay, expressive and receptive language disorder), Speech and Sound Disorders (phonetic and phonological language disorder), Childhood-Onset Fluency Disorder (Stuttering), Social (Pragmatic) Communication Disorder, and Unspecified Communication Disorder. These categories changed from the DSM-IV[9] categories of Expressive Language Disorder and Mixed Receptive-Expressive Language Disorder.

1.1.3. ATTENTION-DEFICIT AND HYPERACTIVITY DISORDER

Attention-deficit and hyperactivity disorder (ADHD) is characterized by symptoms of inattention and/or hyperactivity-impulsivity, occurring in several settings and more frequently and severely than the other children in the same stage of development[2]. ADHD often disrupts family and peer relationships, diminishes school performance. Children with ADHD frequently have other problems, with parents reporting that about half of children with ADHD have a learning disability and about one in four have a conduct disorder[18]. Other disorders, including anxiety, depression, and learning disabilities, can be expressed with signs and symptoms that resemble those of ADHD.

A great deal of research on ADHD has focused on aspects of brain functioning that are related to the behaviours associated with ADHD. These children generally have trouble with some skills involved in problem-solving and executive functions. These skills include working memory (keeping information in mind while briefly doing something else), planning (organizing a sequence of activities to complete a task), response inhibition (suppressing immediate responses when they are inappropriate), and cognitive flexibility

(changing an approach when a situation changes). Maternal smoking during pregnancy has been associated with increased risk of ADHD in the child in numerous studies[19].

Preterm birth and low birth weight have also been found to increase the likelihood that a child will have ADHD[20].

Psychosocial adversity (representing factors such as low socioeconomic status and in-home conflict) in childhood may also play a role in ADHD[21].

1.1.4. LEARNING DISORDER

Learning disability (or learning disorder) is a neurological disorder in which a brain has an impairment in the way in which receive, process, retain, and respond to information. A child with this disability may have trouble learning and using certain skills, including reading, writing, listening, speaking, reasoning, and doing math. These children usually have average or above-average intelligence, but there are differences in the way their brains process information. DSM-5[2] combines the diagnoses, which were separate in DSM-IV[9], of reading disorder, disorder of written expression, mathematics disorder, and learning disorder not otherwise specified.

Problems during pregnancy and birth, such as maternal drug or alcohol use, low birth weight, lack of oxygen, or premature or prolonged labour, may also lead to learning disabilities[22].

1.1.5. COORDINATION DISORDER

Developmental Coordination Disorder (DCD) is the condition described on DSM 5 in which there were severe difficulties in learning everyday motor skills, cannot be explained by physical, sensory, or cognitive impairment[2]. Motor milestones are often not in delay; the delay is in fact in the acquisition of motor skills. Organizational skills can be less well developed than other children. Motor skills require effort so kids with DCD are often slow to complete tasks at school and may appear inattentive. Children with DCD usually begin to withdraw from and avoid sports and motor activities. They often seem verbally advanced but impaired socially and might have emotional or behavioural problems.

Developmental Coordination Disorder may exist in isolation or co-occur with other disorders such as learning disabilities or ADHD. The Prevalence is 5-6% of the school-aged population, more frequent in the boys.

1.1.6. AUTISM SPECTRUM DISORDER

Autism, or autism spectrum disorder (ASD), is a neurodevelopmental disease characterized by impairment in communication, social skills, and behaviour with involving rigidity of interests and repetitive, stereotypical comportments. Associated symptoms may contain obsessive-compulsive, sleep, hyperactivity, attention, mood, gastrointestinal, self-injurious, ritualistic, and sensory integration disorders.

The term “spectrum disorders” refers to the fact that although people with ASDs share some common symptoms, ASDs affect different people in different ways, with some experiencing very mild symptoms and others experiencing severe symptoms[2]. ASDs encompass autistic disorder and the generally less severe forms, Asperger’s syndrome and pervasive developmental disorder-not otherwise specified (PDD-NOS)[9].

At the moment there is no specific treatment for autism. Behavioural therapy is used, but the outcomes are still unsatisfying[23]. In many severe cases with comorbidities such as aggressivity, self-harm, attention deficit, anxiety, mood disorders, pharmacologic therapies such as antidepressants and antipsychotics are recommended but they may cause adverse effects[24].

Autistic symptoms and traits typically manifest during early childhood (i.e. around the age of 2years), although most individuals with ASD remain diagnosed after the age of 5 years[25]. There is consensus that the onset of autistic symptoms and their progression over the human life span is accompanied by differences in brain anatomy and connectivity. Longitudinal studies based on total brain volume calculation demonstrated atypical neurodevelopmental trajectory of brain maturation in ASD compared with typically developing children. MRI studies in ASD have demonstrated increased brain overgrowth in the early developmental phases: toddlers with ASD (age 2–4 years) have, on average, a larger brain volume than typically developing children.

However, this increased brain volume seems to disappear around the age of 6–8 years, when growth curves intersect, and is followed by a period of accelerated age-related decreases in cortical thickness in various brain regions in ASD across the remaining life span. Shen et al. (2013) reported that infants with ASD had significantly greater extra-axial fluid (i.e. excessive CSF in the subarachnoid space) at 6–9months[26]. Other studies have demonstrated that early brain enlargement is accompanied by significantly increased white matter volume in ASD, that may be particularly prominent in compartments that are located in close proximity of the cortical sheet [27] [28] [29]. On the other hand, several studies have shown that early brain overgrowth is due to atypical development of the cortical grey matter in ASD, which is determined by two distinct neuroanatomical features, namely cortical thickness and cortical surface, or a combination of both[30]. Interestingly, recent revealed that regional rather than global neuroanatomical variation may be more suited to characterize the neurobiological underpinnings of ASD and to provide neuroanatomical marker for ASD across the human life span. Increased GM volume (GMV) and WM volume (WMV) have been found in the frontal and temporal lobes and are less pronounced in the parietal and occipital lobes. Trajectory of development of amygdala and hippocampi follows overall trajectory of total brain volumes, showing increase in volume accompanied by more severe impairments in the social and communication domains. Frazier et al. demonstrated decreased corpus callosum volume in ASD suggesting abnormal interhemispheric connectivity causing reduced integration of information and slower processing[31]. Cerebellar volume differences have also been detected in ASD, while enlargement of the caudate nucleus in ASD has been associated with the severity of repetitive and stereotyped behaviours in these patients.

Molecular studies have demonstrated dysmaturation of the WM characterized by microstructural changes or disorganization in the brains of autistic populations. It has been reported that synaptogenesis is altered in children with ASD, affecting myelination, and thus compromising WM integrity. Compared with the control group, both the autism and sibling groups had widespread, significantly reduced white matter FA values in the frontal parietal and temporal lobes and included, but were not restricted to, regions known to be important for social cognition. Hong et al. demonstrated significantly less WM density, higher ADC, lower FA values in the corpus callosum of ASD patients compared to typically developing children[32]. Moreover, it was found that perturbations in the development of white matter are already present at 6 months in infants with a confirmed diagnosis of ASD at the age of 2 years, relative to infants who do not have an ASD[33].

In view of the abnormal development of grey and white matter in autism, affected brain regions will inevitably be unable to generate connections that give rise to fully effective functional networks. fMRI studies have demonstrated reduced activation in regions comprising the social brain network during tasks related to emotional processing or social cognition, the perception of biological motion, as well as self-referential cognition and empathy. Moreover, there is convergent evidence that the ASD group exhibited lower levels of functional connectivity and less network integration between frontal, parietal, and occipital regions.

Failure to develop normal language comprehension is an early warning sign of autism, but the neural mechanisms underlying this signature deficit are unknown. At-risk toddlers later diagnosed as autistic display deficient left hemisphere response to speech sounds and have abnormally right-lateralized temporal cortex response to language; this defect worsens with age, becoming most severe in autistic 3- and 4-year-olds. Typically developing children show opposite developmental trends with a tendency towards greater temporal cortex response with increasing age and maintenance of left-lateralized activation with age. Sensory over-responsivity, defined as a negative response to or avoidance of sensory stimuli, is both highly prevalent and extremely impairing in youth with ASD. fMRI studies revealed that ASD patients have greater responsivity in primary sensory regions as well as areas related to emotion processing and regulation (i.e., amygdala, hippocampus, and orbital-frontal cortex). Finally, studies of the whole brain network using graph theory, regional homogeneity analysis[34], or other analytic techniques[35], mostly found enhanced local connectivity over multiple brain regions in ASD, especially in the frontal lobe. While studies on networks of interest indicated region-specific reduced long-range connectivity, mainly on default mode network (DMN), salience network (SN), and executive control network (ECN)[36].

In a minority of cases (less than 10%) the Autism Spectrum Disorder is associated with diseases caused by a single gene, including the Syndrome Fragile X and Tuberous Sclerosis (TSC).

Some studies showed that children with autism manifest increased oxidative stress[37], maybe after environmental factors exposure (pollution, radiation, for example). The excess reactive oxygen species (ROS) are known to be a genotoxic stress that might induce DNA damage and mutation following ineffective repair of DNA damage[38].

Many studies have found a link between increasing paternal and maternal age and increased risk of ASDs[39]. Advanced parental age can contribute significantly to the frequency of de novo mutations in a parent's germ cells[40].

2. UNRAVELLING THE GENETICS OF NEURODEVELOPMENTAL DISORDERS

The phenotypical heterogeneity in NDDs is reflected in marked genetic heterogeneity, with an etiological diagnosis not possible for most cases. These complex disorders, unlike Mendelian disorders, are defined by a phenotype that is not caused only by one or two pathogenic variants in a single gene, but it is necessary that many genetic events happen with significant contribution from environmental factors. In fact, disorders with a simple genetic cause are a small fraction of the spectrum of NDDs[41] and the challenge is to better understand the correlation between phenotype and genotype for NDDs.

2.1. EPIDEMIOLOGICAL STUDIES

Epidemiological studies on Neurodevelopmental Disorders, in particular Intellectual Disability, ADHD and Autism Spectrum Disorders, have provided the first compelling evidence of a strong genetic component in these disorders. The leading observation is that NDDs run in families and often different types of disorders may be present in the same family, for example ADHD and autism[42]. Traits or milder phenotypes present in relatives of individuals with autism may reflect inheritance of allelic subsets of those present in the affected individual. Furthermore, first-degree relatives of autistic probands have almost a recurrence risk of 11-19% for the disorder, compared to 1%–2% risk for ASD in the larger US population[4344].

The confirmation that familiarity is due to genetic factors has come also from twin and adoption studies. In multiple cohorts of ASD twins it has been observed a significant higher concordance in affection status for monozygotic brothers compared to dizygotic ones, respectively 88%–95% in monozygotic compared to 31% in dizygotic[45 46]. Both type of twin pairs is considered to have the same environment, including in utero and perinatal conditions, thus the higher concordance rate in genetically identical subjects clearly indicate the role of genes in the risk. Similarly, adopted away children whose biological parents or siblings have NDDs have a risk to be diagnosed with the same disorder higher than that of the general population.

To date, more than 1000 genes have been implicated in the etiopathogenesis of NNDs. Many of these are involved on common pathways and protein networks, a fact that mirrors the clinical and molecular variability of NDDs. Among them, genes important for neuronal migration, development, branching of the neurites, synaptic function, transcriptional regulation, and construction of neuronal network are strongly represented[47].

Human Gene Module lists 881 genes implicated in ASD, many of these are responsible for the synaptic connectivity, with mutations leading to microscopic neuronal dysconnectivity.
Since ASD is a neurodevelopmental disorder often diagnosed before the age of three years, and previous studies have demonstrated that ASD risk genes are expressed primarily during foetal development, genes with a prenatal to early postnatal expression trajectory are good candidates, particularly in those brain regions that are involved in the development of autistic features, such as repetitive restricted behaviours.

2.2. DETERMINE THE GENETIC COMPONENTS OF A COMPLEX DISEASE

Since the establishment of heritability, considerable efforts have been undertaken to identify the loci involved in Developmental Disorders and their relative contribution, thus the genetic architecture of the disorders. The hypothesis of a major locus with a Mendelian-like inheritance, has been almost immediately rejected, because inconsistent with the high prevalence of the disorders in the population. Preliminary investigations have suggested that the majority of Developmental Disorders, in particular ASD, are actually polygenic; in addition, the genetic and environmental interplay in defining the phenotype clearly classifies NDDs such as ID and ASD as complex disorders. Thus, the observed patters of inheritance can be described by the so-called gene-environment multifactorial threshold models. The set of multiple genetic and environmental elements constitute the individual predisposition, which represents one continuous variable with Gaussian distribution in the general population. Therefore, each person has a certain degree of predisposition, while only a minor percentage of individuals is placed at the curve extremes, respectively with a very low or a very high predisposition to develop the disease. Individuals whose predisposition exceeds a certain level, defined as "threshold effect", manifest the disorder.
To identify the genes responsible for the pathology susceptibility, either Mendelian or complex genetic components, different strategies are used, which can be used in combination or sequence. The approaches that have been used so far for the identification of susceptibility genes for oligo- or polygenic diseases include:

- Functional candidate gene hypothesis: the involvement of a gene important in the pathophysiological process either
- Studies on chromosomal aberrations and on structural variants: cytogenetic abnormalities may result fundamental in identifying the genes responsible for the condition. The presence in the human genome of numerous sub-microscopic structural variants, as chromosomal deletions/duplications ranging from 1Kb to 3Mb, which vary in terms of number of copies, appears important. Such structural variants called Copy Number Variants (CNVs), in some cases, can be associated with diseases.

- Linkage studies: allow to localize the chromosomal position of the disease locus compared to that of polymorphic markers, observing the co-segregation of loci in families with at least two affected children.
- Studies with massive DNA/RNA sequencing techniques (also called Next Generation Sequencing, NGS). It is a series of technologies, in part still used as research protocols, that make it possible to sequence many genes (gene panel), whole exomes (whole exome sequencing, WES) and/or whole genomes (whole genome sequencing, WGS) in a short time, of the order of weeks.

The loci of susceptibility for a disease can be identified based on the indication of the known function and / or the expression profile of the gene. When the role of a given protein in the pathophysiological process of a disease is known, it is possible to hypothesize its direct involvement. It is therefore possible to trace the gene and analyse possible mutational events affecting the gene sequence and formulate a hypothesis for the disease aetiology. Studies conducted on animal models are interesting in this field. For many pathologies, the mouse model is used because it mimics the clinical characteristics of the condition under examination. Knock-in mice are for example used to analyse the acquisition of a new function by mutated genes (gain of function), while knock-out mice are useful when the abolition of the gene function is at the basis of a disease (loss of function).

2.2.1. COPY NUMBER VARIANTS

Array-comparative genomic hybridization (array-CGH) is widely used to detect copy number variants (CNVs) associated with intellectual disability (ID) and autism spectrum disorders (ASDs).

CNVs are DNA deletions or duplications with respect to the reference genome, usually in contiguous positions. The size of these structural variants ranges from 1 kb to several Mb, so they are sub-microscopic and hardly identified with standard karyotype analysis. The majority of CNVs are common changes of the human genome, with a frequency of more than 1%, and defined as copy number polymorphisms (CNPs)[48]. Rare CNVs, instead, have long been known to be involved in neurodevelopmental disorders, causing intellectual disability and autism spectrum disorder. Some recurrent rearrangements were shown to be consistently associated with ASD in unrelated individuals and families. The examination of structural variants in NDDs and ASD have been trigged by some precedents from cytogenetic studies (15q11-q13 associated with Prader-Willi and Angelman syndromes, 22q11.2 deletion associated with Velo-Cardio-Facial syndrome, 22q13 deletion associated with Phelan-Mc Dermit syndrome, 16p11.2 deletions and duplications and 22q11.2 deletions).

Structural anomalies determine functional genetic changes that can occur in various ways:

• “dosage effect”: as a consequence of the change in the number of copies of a gene following deletions or duplications.

• “position effect”: because of a chromosomal rearrangement, one or multiple genes can be separated by regulatory sequences.

• breakpoints directly associated with a chromosomal rearrangement can destroy the gene.

• deletions in one of the two alleles can show a point mutation located in the corresponding region on the homologous non-deleted chromosome.

In the diagnostic routine the analysis of the standard and molecular karyotype (array-CGH) allows to identify the genetic origin in less than 10% of the ASD patients, while results in an average diagnostic yield between 15% and 20% in the overall NDDs patients[49]. Furthermore, some CNVs have no clear causative effects and remain of uncertain significance (VOUS).

2.2.2. NEXT GENERATION SEQUENCING (NGS) TECHNOLOGIES

In the last years, several innovative techniques for sequencing DNA have become available in various fields of medical research, as well as in clinical practice. These methods, called Next Generation Sequencing (NGS), are progressively replacing traditional Sanger sequencing, and have made possible the study of an unprecedent amount of DNA sequences at relatively low cost, contributing to an enormous expansion of our knowledge on different genetic diseases. It is possible to distinguish three different technique[5051]

1 Whole Genome Sequencing (WGS), which is the most complex approach, allowing the sequencing of the entire human genome.

2 Whole Exome Sequencing (WES), that allows the analysis of the exome, the coding part of the human genome (~2% of the entire genome).

3 Targeted sequencing/sequencing panels of multiple genes responsible for a specific group of diseases.

These technologies permit for the first time the investigation of variants impossible to detect before: rare sequence variants, that include single nucleotide variants (SNVs) and small insertions/deletion (indels). The analysis of such rare alleles in case-control cohorts has proved arduous. As previously described for CNVs, in fact, a main issue is the discrimination of low frequency variants from control ones. Additionally, unlike structural variants, sequence variants have shown an unexpected heterogeneity. Thus, huge samples sizes are necessary to achieve sufficient statistical power; for this reason, association studies with sequencing data have been mainly performed so far on NDDs cases. To overcome these problems, first studies focused on the identification of de novo coding mutations in family trios of affected probands with unaffected parents. Justified by the hypothesis of a balance of negative selection, the approach offers in fact an important technical advantage, with the possibility to focus on a small subset of variants. Several works have supported an increased frequency of potentially damaging de novo alleles in NDDs probands compared to controls[42 52].

Exome sequencing is particularly useful in finding diagnoses for patients with Mendelian disease phenotypes for whom the causal variant has not been identified through other means. WES of patients with previously undiagnosed neurodevelopmental diseases has resulted in a genetic diagnosis of up to~40% in some cohorts[53]. Generating and analysing sequence data is still computationally expensive and requires particular expertise. It may also be difficult to identify clinically relevant variants given the large number of variants in the exome, particularly when data are not available for parents.

The above-mentioned NGS methods are characterized by specific sample preparation protocols and data analysis. All NGS technologies share a basic workflow in which an initial high molecular weight DNA sample is shattered into a fragment library, and single strand molecules are amplified and sequenced in parallel. DNA fragments are generated by mechanical or enzymatic methods. In exome sequencing and resequencing of customized gene panels, an additional crucial step involves the enrichment of the target DNA fragments (capturing). Libraries are obtained and specific adaptors are attached to both ends of each fragment. This passage allows the fragments to be more easily PCR amplified using primers or to be hybridized to a surface using complementary adaptors. Then the sequences of the fragments are cyclically read using fluorescence or electrical signals. For DNA resequencing, bioinformatic analysis includes the alignment of the raw reads against the reference sequence and the comparison of aligned reads against the reference to obtain a list of genomic variations (variant calling). Subsequently, descriptive information is added to the identified variants (variant annotation). The large amount of data generated is then processed using bio-informatic tools. One of the main drawbacks of the WES approach in fact is the large number of genomic variants detected in each analysis, which is by no means manually manageable. Automated variant prioritization strategies are therefore required. The first filtering step consists in selecting only variants with a minor allele frequency (MAF) <1%, other tools include filters that recognize non-synonymous and splicing variants, the analysis of evolutionary conservation of the affected amino acid throughout species, *in silico* analyses, and interrogation of publicly available population and disease-specific database. After this process, variants surviving the filtering need a more precise characterization to define their pathogenicity. Segregation of the variant within the patient's family, replication of the same finding (i.e. same mutation or same mutated gene found in additional subjects with a similar phenotype) and functional studies in cellular or animal models are often needed to fully support the effects of a specific variant.

The still high costs and the large amount of data resulting from the sequencing of the whole genome, which requires more complex bioinformatics analysis, and is associated with the difficulty in attributing a functional meaning to many of the variants detected, mean that the WGS currently finds its greater applicability in the field of research. Although in the future the costs related to genomic sequencing are progressively destined to decrease, the transfer of this method into clinical practice still appears difficult.

A valid alternative to WGS is represented by the possibility of selecting and sequencing only the portion of the genome corresponding to the coding regions through exome sequencing (WES). In general, WES allows

us to reduce costs and sequencing times without the probability of success being significantly reduced. The main limitation of the study of the exome is the lack of determination of variants present in the non-coding part, such as structural variants and chromosomal rearrangements.

Many rare causative variants found within coding sequences are missense or nonsense substitutions or small insertions and deletions for which a serious or deleterious impact on a functional level can be predicted.

The identification of causative genes for rare genetic diseases for which mutations are not known in literature require filtering strategies that allow for the identification of presumably pathogenic variants by selecting them from a background including non-pathogenetic variations (polymorphisms) and sequencing errors.

Variant filtering criteria in the case of mono or oligogenic diseases are generally based on the following considerations:

- Mutations in one or a few genes are generally sufficient to cause disease,
- Such mutations should be rare and high penetrance,
- Due to their impact they are probably located in the gene-coding part (missense, nonsense, frameshift or splicing mutations).

Studies of *de novo* sequence variation using whole-exome sequencing (WES) have proven to be a powerful approach to systematic gene discovery in neurodevelopmental disorders, particularly autism spectrum disorders[47].

2.2.3. NEW STUDY APPROACHES

Studies of genetic diseases traditionally followed a pathway from phenotype to genotype to gene, a so-called "forward genetics" approach. In the last years, most disease genes have, however, been identified by a different approach called "reverse genetics": the identification of a gene via its chromosomal location, without prior knowledge of its protein product or the molecular pathway in which it functions.

The development of techniques to generate and study large mole of genetic data by NGS has allowed genetic research to concentrate on the use of genome-wide approaches to identify genes. This conduct to the "reverse phenotyping"[54]. In this approach, the objective is to define phenotypic groupings that are distinguished by higher rates of allele-sharing, linkage data, or more deviant allele frequencies, association data, than are described in the traditional diagnostic groups. This "genotype-first" approach was introduced by Stessman et al as a way of defining the different phenotypes of a complex disorder[55]. They called a "genetic subtype" a gene in which recurrent variants show an overbalance of burden in patients versus controls. A "molecular subtype" represents a group of genetic subtypes that are linked together in a same pathway[56].

While genetically informed targeted therapies are the ultimate goal of precision medicine, there are important clinical and psycho-social benefits to the genotype-first approach. For families this conducts into better diagnosis, counselling and support strategies.

3. SCOPE OF THE THESIS

In this dissertation, I sought to explore the contribution of rare de novo and inherited coding variation in neurodevelopmental disorders and use these genetic variations to identify neurodevelopmental disorder associated genes and new/unknown oligogenic mechanisms.

During the PhD course, my activity as a clinician has been mainly focused on Neurodevelopmental Disorders. As a Child Neuropsychiatrist, the activities I carried out allowed me to broaden my knowledge in the field of NDDs, that represents my main topic of interest.

In terms of clinical duties, I worked in the Child Neuropsychiatric ward and day-hospital of IRCCS Istituto Giannina Gaslini. Meanwhile, I collaborated with the Medical Genetic Unit of the same Institute and with the Department of Neuroscience and Brain Technologies of the Istituto Italiano di Tecnologia (IIT), where I contributed to the interpretation of the results of genetic analyses carried out by geneticists for diagnostic purposes, establishing an active link between clinicians and the lab.

This thesis collects doctoral studies about array CGH studies on neurodevelopmental disorders (NDDs) and whole exome studies (WES) on autism spectrum disorders (ASD).

The work has been organized in three parts.

In chapters 4 and 5, studies relating to the re-evaluation of the data obtained through array CGH, which led to suggest the involvement of new genes and new pathogenetic mechanisms in NDDs. In particular in chapter 4 there are two case descriptions, the first report already published and the second in course of publication.

Chapter 5 is the summary of our work to identify new genetic mechanisms.

Chapter 6 reports the preliminary data of an ongoing doctoral study on WES in children with ASD, that include the case report already published.

3.1. AIM OF THE RESEARCH

I list the following aims for my dissertation:

- Identify the most likely pathogenic microdeletion and microduplication CNVs associated with NDDs, among diagnostic variants previously classified as VOUS, by combining mining of literature and searching in databases reporting deleterious and benign variants, and ID/ASD-associated genes.
- Identify known genetic / genomic mutations associated with ASD in a group of selected patients, using Whole Genome Sequencing (WGS) and Whole Exome Sequencing (WES) analysis technologies.
- Discover / recognize potential new candidate genes and new pathogenetic mechanisms involved in NDDs and ASD.

3.2. RESEARCH DESIGN AND METHODS FOR ACHIEVING THE STATED OBJECTIVES

Written informed consent was obtained from the patient's parents or legal representative.

The studies were reviewed and approved by the Ethics Committee of the Italian Regione Liguria.

In a retrospective review of data, we re-evaluated all the results of diagnostic array-CGH tests on 700 cases with NDDs admitted at the Gaslini Institute, Genova, Italy, over a 5-year period (2014-2018) focusing on variants previously interpreted as VOUS.

Furthermore a series of 68 patients with autism spectrum disorder, associated with intellectual disability or enrolled to another IIT study based on movement, without clear syndromic pictures (Fragile X Syndrome, RETT syndrome) and pathogenetic deletion / duplication, were recruited by the Neuropsychiatry Unit and by the Medical Genetic Unit from 2018 to 2019 to perform whole exome sequencing and, eventually, whole genome sequencing.

3.2.1. PRELIMINARY ASSESSMENT

The clinical assessment of the patients comprised a thorough medical history also attentive to investigating systemic symptoms and sleep, a neurological-behaviour examination, a developmental/cognitive assessment, a basic metabolic screening.

Prior to the array-CGH analysis, all patients were first evaluated by using different genetic tests for the common genetic causes of NDDs, including karyotype analysis, and tests for Fragile X syndrome.

3.2.2. ARRAY CGH

Array-CGH analysis was performed on DNA samples, extracted from peripheral blood, using a whole-genome 180 K Agilent array with ~13 Kb overall median probe spacing (Human Genome CGH Microarray, Agilent Technologies, Santa Clara, CA, USA). Data were analysed using Agilent CytoGenomics and genomic positions reported according to the human genome assembly (GRCh37/hg19). All detected CNVs were tested for inheritance by hybridization of the parental DNA with the same array platform. De novo CNVs were confirmed either by FISH analysis or by quantitative real-time PCR.

3.2.3. WHOLE EXOME SEQUENCING

Genomic DNA extraction and subsequent WES and data analysis for affected individuals and their parents were carried out by Italian Institute of Technologies (IIT) in Genoa, Italy. Briefly, the exomes were captured using the xGen® Exome Research Panel v1.0 – IDT KIT. Sequences were enriched using the Illumina Nextera

Flex for Enrichment KIT. Sequencing was performed on an Illumina NovaSeq 6000 platform (Illumina Inc., CA, USA). The library preparation and its sequencing were performed simultaneously for the parents and the proband using barcode adapters. Alignment of raw paired end reads to the reference genome (version hg19) was performed with bwa (version 0.7.17)[57]. Duplicated reads were marked with Picard (version 2.18.20). Variant discovery was then performed with the GATK4 utility Haplotype Caller, using the appropriate file containing the coordinates of the sequences targeted by the exome sequencing. Finally, all variants were annotated using ANNOVAR (databases updated to 27/05/2019)[58]. The resulting file was used for a manual evaluation of the variants. The sequencing provided a 60x medium coverage.

3.2.4. CNV ANALYSIS

CNVs identified by diagnostic array-CGH in 518 NDDs cases were re-analysed. A deep analysis, mainly consisting in a revision of gene expression/function annotation, and chromatin organization data, was performed. New candidate genes were analysed by GeneCodis4 to evidence enrichment for known NDD-associated GeneOntology terms and pathways.

3.2.5. WES DATA ANALYSIS

After whole exome sequencing, potentially deleterious variants were prioritized by custom filtering strategies including the use of ORVAL (Oligogenic Resource for Variant Analysis Platform) and enrichment analysis of candidate genes with GeneCodis4.

3.2.6. REVISION OF DATA AND GENOTYPE-PHENOTYPE CORRELATION

The analysis of filtered data and the genotype-phenotype correlation was performed by integrating multiple bioinformatics resources and a synergistic effort of our team of neuropsychiatrists and geneticists.

4. EVALUATION OF CNVS VARIANTS IN A COHORT OF ISOLATED AND SYNDROMIC INTELLECTUAL DISABILITY/AUTISM SPECTRUM DISORDERS REVEALS NOVEL POSITION EFFECTS AND CANDIDATE DISEASE GENES

4.1. INTRODUCTION

Array-comparative genomic hybridization (array-CGH) is widely used to detect copy number variants (CNVs) associated with intellectual disability (ID) and autism spectrum disorders (ASDs). However, some CNVs have no clear causative effects and remain of uncertain significance (VOUS). Their low frequency or their size does not allow us to easily understand which are the causative genes among those included in the range. Comparison with data in the literature by consulting public databases (DGV, DECIPHER, ClinGen, OMIM and PubMed) is essential for the interpretation of the pathogenetic significance of the variants found.

4.2. MATERIALS AND METHODS

We retrospectively reviewed the results of diagnostic array-CGH test on 700 cases with isolated or syndromic forms of ID and ASDs. The analyses were performed with the Agilent 180K array-CGH method (Agilent Technologies, Palo Alto, CA, USA) with a resolution of 13 kb.

We assessed pathogenicity of identified CNVs by combining mining of literature and searching in databases reporting deleterious and benign variants, and ID/ASD-associated genes. VOUS variants encompassing no genes or genes without any apparent brain expression/function were further investigated for potential long range position effects through involvement of Topologically Associating Domains (TADs), and other chromatin activation signatures by using the web-based 3D Genome Browser and UCSC Genome Browser.

4.3. RESULTS

We identified non-benign CNVs in 314 patients, of whose 158 patients showed the presence of uncommon CNVs.

Variants classified as pathogenetic included those already reported in the literature as causative of known syndromes (table 1) and those reported as statistically relevant for the phenotype in the work of Coe and collaborators[59]. We also considered some variants that, although to date not reported as pathogenetic, could be classified as pathogenic on the basis of some parameters such as: size of the variant; inheritance; comparison with databases showing rare variants not associated with neurological problems (DGV); function and expression profile of genes involved in deletions and / or interrupted by duplications.

Among the pathogenetic and the probably pathogenetic variants (24%), we found CNVs involving genes recently associated with ID/ASDs, because of missense mutations identified by whole exome sequencing (WES), and variants involving novel candidate genes and/or regions of regulatory expression of flanking ASD-associated genes.

Table 1: CNVs identified in our patients and already reported in the literature.

N°	Chromosome	CNV	Diagnosis	Inheritance	Description
1	2q13	dup	ID	mat	2q13 duplication syndrome
1	2q13	del	DD and dysmorphisms	pat	2q13 deletion syndrome
2	2p16.3	del	DD	pat/pat	2p16.3 deletion syndrome, including *NRXN1*
1	3q28q29	dup	DD	dn	3q28q29 duplication, involving *FGF12*
1	3q29	del	DD and microcephaly	mat	3q29 duplication syndrome
1	7q11.23	dup	DD and dysmorphisms	NA	WBS duplication syndrome
1	7q11.23	del	DD	dn	William syndrome deletion
1	13q31.3	del	DD and growth deficit	dn	Feingold syndrome type 2
1	15q11.2	del	DD	mat	15q11.2 deletion syndrome
1	15q11.2q13.1	dup	DD	dn	PWS duplication syndrome
2	15q13.1q13.3	del	DD and dysmorphisms	dn/NA	15q13.3 deletion syndrome
2	16p11.2	del	epileptic seizures	dn/dn	16p11.2 deletion syndrome
1	16p11.13	del	DD	pat	16p11.13 deletion syndrome
1	16p11.13	dup	DD	pat	16p11.13 duplication syndrome
3	17p13.3	dup	DD	dn/pat/dn	17p13.3 duplication syndrome, including *YWHAE*
1	22q11.1q11.21	del	atypical behaviours	dn	22q11 deletion syndrome

1	22q11.22q11.23	dup	speech delay	pat	22q11.2 distal duplication syndrome
2	22q13.33	del	DD	dn/dn	Phelan McDermid 22q13 deletion syndrome

Legend: del= deletion, dup= duplication, ID= intellectual disability, DD= developmental delay, dn= de novo, NA= not available, pat= paternal, mat= maternal

Here I will report 2 cases, both affected by NDDs and carrying de novo variants not reported to date and probably pathogenetic.

4.3.1. FIRST CASE

In this first article we described an adult patient affected by mild intellectual disability with history of absence epilepsy in adolescence, in whom we identified an intragenic duplication in *KCNQ5*.

Intragenic duplication of KCNQ5 gene results in aberrant splicing leading to a premature termination codon in a patient with intellectual disability[60]

Rosti G, Tassano E, Bossi S, Divizia MT, Ronchetto P, Servetti M, Lerone M, **Pisciotta L**, Mancardi MM, Veneselli E, Puliti A. Eur J Med Genet. 2019 Sep;62(9):103555. doi: 10.1016/j.ejmg.2018.10.007. Epub 2018 Oct 22. PMID: 30359776.

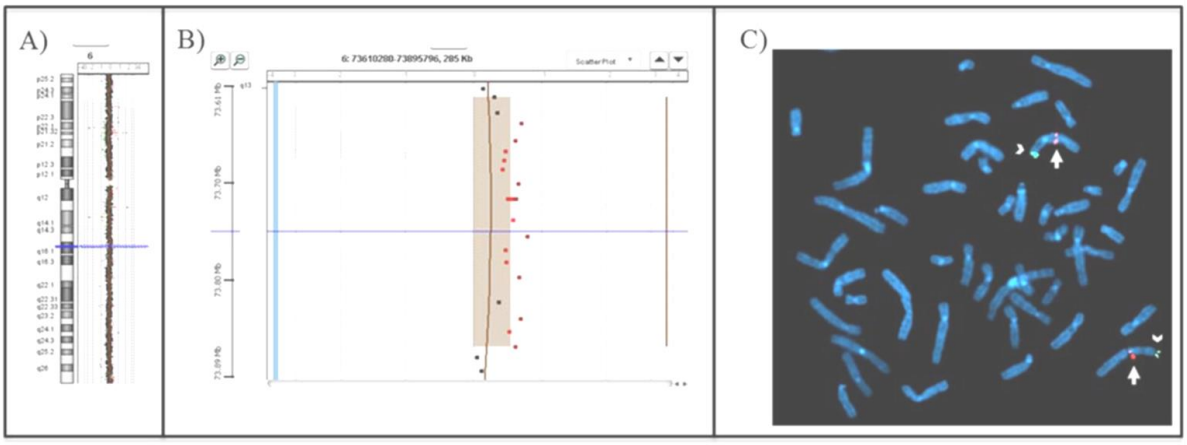
Here is a summary of the paper.

The *KCNQ5* gene lies in chromosome 6q13 and encodes a voltage gated potassium channel (Kv7.5) widely expressed in the brain. *KCNQ5* and other two voltage-gated potassium channels encoded, respectively, by *KCNQ2* and *KCNQ3* genes are known to play an important role in the regulation of M-type current and after hyperpolarization conductances which contribute to neuronal excitability. *KCNQ2* and *KCNQ3* can heterodimerize with one another and heterozygous mutations in either *KCNQ2* or *KCNQ3* result in benign familial neonatal seizures (MIM 121200 and MIM 121201) or epileptic encephalopathy (MIM613720) through either loss-of-function or gain-of-function mechanisms (Miceli et al., 2015; Orhan et al., 2014; Singh et al., 1998). Alteration and, in particular, reduction of M-type currents because of mutations in *KCNQ2* or *KCNQ3* can lead to enhanced neuronal excitability causing early-onset epileptic disorders (Weckhuysen et al., 2012). In mice, a dominant- negative *KCNQ5* mutation that makes homomeric and heteromeric non-functional channels was obtained (Tzingounis et al., 2010). Electrophysiological studies demonstrated that

homozygous mutated mice had significantly reduced medium and slow afterhyperpolarization currents in the hippocampus, although their brain morphology was normal and no seizures were observed (Fidzinski et al., 2015; Tzingounis et al., 2010). Similarly to *KCNQ2* and *KCNQ3* mutations, that proved to have deleterious effects on human phenotype, and on the basis of studies in *KCNQ5* mutated mice, *KCNQ5* mutations are expected to have a deleterious effect associated with a neurological phenotype. A recent paper by Lehman and collaborators (Lehman et al., 2017) reported de novo heterozygous missense mutations identified by exome sequencing in four probands with intellectual disability, abnormal neurological findings, and treatment-resistant epilepsy. These variants were not present in GnomAD database (Lek et al., 2016) and all of them were associated with a high score of pathogenicity prediction. Functional analyses of the four *KCNQ5* variants, expressed in frog oocytes, revealed shifts in the voltage dependence of activation, including altered activation and deactivation kinetics. Three variants seemed to have loss-of-function and one gain-of-function effects. Based on these findings, Lehman and collaborators suggested that both loss-of-function and gain-of-function *KCNQ5* mutations could have deleterious effects on Kv7.5 channel activity. Therefore, these heterozygous mutations in the *KCNQ5* gene were associated with an autosomal dominant mental retardation form that was classified as MRD46 in the OMIM database (MIM 617601). No other deleterious molecular mutations or copy number variations (CNVs) involving *KCNQ5* have been associated with intellectual disabilities so far (Deciphering Developmental Disorders, 2017; Iossifov et al., 2014). The findings by Lehman and collaborators represent the only reported evidence supporting *KCNQ5* involvement in neurodevelopmental disorders.

Herein we report the first *KCNQ5* intragenic duplication, identified by array-CGH, in a patient with intellectual disability and history of absence epilepsy in adolescence.

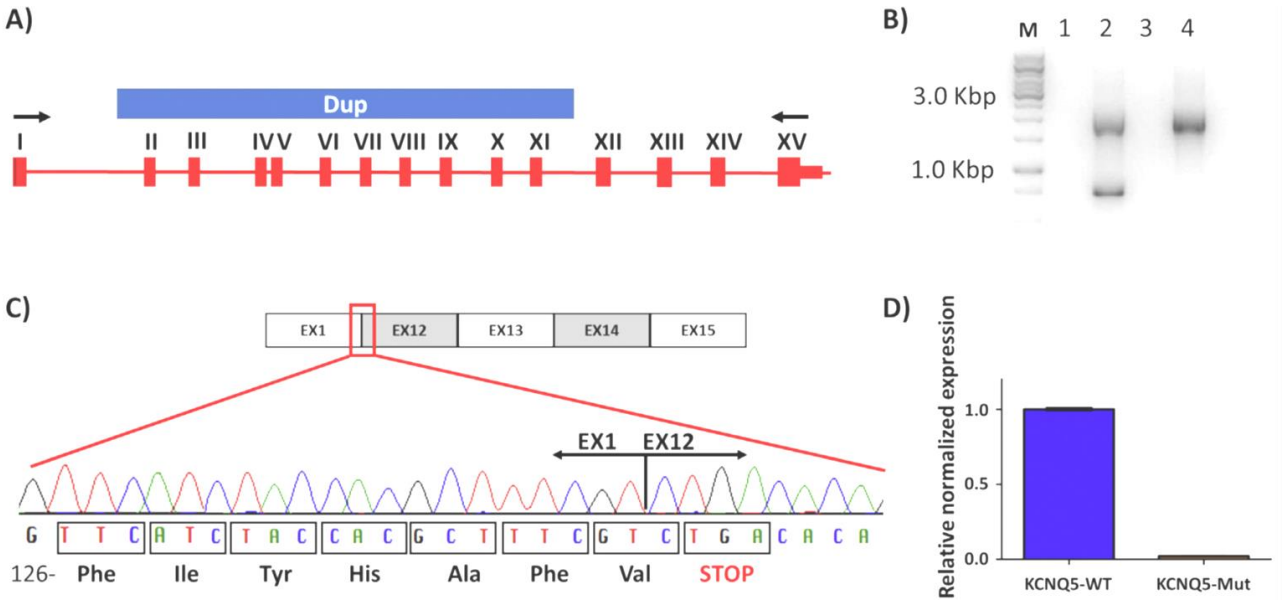
Figure 1 Results of array-CGH and FISH analyses.



(A) Array-CGH profile of chromosome 6 and (B) of a ~239 Kb duplicated region at 6q13 band. (C) FISH using BAC probe RP11-964B9 (chr6:73,626,399–73,791,626) (red), mapping to the 6q13 duplicated region and a subtelomere specific probe 6p (green) used as marker of chromosome 6.

The array-CGH results showed the presence of an intragenic duplication of the chromosome region encompassing exons 2–11 of KCNQ5 (Fig. 2a).

Figure 2 Analysis of KCNQ5 cDNA expression in the patient lymphoblasts.



(A) Schematic representation of KCNQ5 genomic structure (GenBank NM_001160133.1)
with exons represented by boxes and introns represented by lines. The position of intragenic duplication is indicated by the large box (top of the scheme). The position of primers used in RT-PCR is indicated by arrows.
(B) One per cent of agarose gel displaying KCNQ5 cDNA exon 1–15 PCR products. M; 1-Kb DNA ladder.
Lanes 1 and 3: empty wells; lane 2: cDNA derived from patient lymphoblasts; lane 4: cDNA derived from a control individual with normal phenotype.
(C) Schematic representation of aberrant structure of RNA obtained from the patient (upper). Electropherogram trace obtained from the Sanger sequencing of gel-purified aberrant cDNA showing the junction point between exon 1 and exon 12. After the splicing alteration, the hypothetical transduced protein introduces a valine residue in position 133, at the exon 1-exon 12 junction, and loses the frame including a new stop codon in the first codon of exon 12.
(D) Quantitative real-time RT-PCR analyses were performed to determine the proportion between the wild-type (WT) and mutant (Mut) mRNA levels of KCNQ5 in the patient lymphoblast cells. The relative expression levels of the two KCNQ5 mRNA molecules were calculated using the ΔΔCt method normalizing to B2M. A 12-fold difference in the expression was detectable between the WT and the mutated forms, the expression of the mutated KCNQ5 was at the limit of detection. Data represent the means±SD (fold change expression mutated KCNQ5 versus WT).

Therefore, we predicted that transcription of *KCNQ5* could be heavily affected by this duplication and result in aberrant mRNA. According to GTEx Consortium data, showing gene expression in 53 tissues and available

through the UCSC Genome browser, *KCNQ5* is expressed in EBV-transformed lymphocytes (GTEx Consortium, 2013). In order to investigate the effects of intragenic *KCNQ5* duplication on RNA transcription of the gene, we established a lymphoblastoid cell line from the patient's lymphocytes. In fact, cultured lymphoblastoid cells from our patient and a normal control showed the presence of *KCNQ5* expression and allowed studies of *KCNQ5* as RNA. After RT-PCR across exons 1 to 15, gel electrophoresis revealed two bands from the patient and one band from the control (see Fig. 2b). Both samples showed a band corresponding to a ~1800 bp amplicon, consistent with that expected from the wild-type (WT) transcript. Sanger sequencing of this band from both patient and control was consistent with the *KCNQ5* reference transcript NM_001160133.1, which includes all exons from the first to the 15th. An additional band corresponding to a ~900 bp amplicon was generated only from the patient's cDNA (Fig. 2b). Sanger sequencing of this band revealed an aberrant transcript exhibiting exon 2–11 (1127 bp) skipping and the direct junction of exon 1 to exon 12. This rearranged RNA, upon translation, could generate a protein terminating at amino acid residue 133 (Val), the last coding triplet lying at the junction between exon 1 and exon 12 (Fig. 2c). Thus, the hypothetical translated protein could be a truncated form consisting of the intracellular NH2-terminal region and only the first 8, out of 21 amino acids, which constitute the first transmembrane alpha-helical structure domain of the potassium voltage-gated channel. To better quantify the proportion between wild-type and mutant mRNA we performed qRT-PCR assays in cultured lymphoblastoid cells. The results indicated that the WT allele was predominantly expressed compared to the aberrant transcript, this latter being at the limit of detection (Fig. 2d).

CONCLUSIONS

We showed that the genomic duplication seems to generate a short transcript with exon 2–11 skipping and a premature stop codon causing, most likely, haploinsufficiency. Thus, our data further support the implication of *KCNQ5* in the pathogenesis of intellectual disability and absence epilepsy, and demonstrate that not only missense mutations but also CNVs affecting the gene can be involved in the disorder.

References for first case

Deciphering Developmental Disorders, S., 2017. Prevalence and architecture of de novo mutations in developmental disorders. Nature 542 (7642), 433–438. https://doi.org/10.1038/nature21062.

Fidzinski, P., Korotkova, T., Heidenreich, M., Maier, N., Schuetze, S., Kobler, O., et al., 2015. KCNQ5 K(+) channels control hippocampal synaptic inhibition and fast network oscillations. Nat. Commun. 6, 6254. https://doi.org/10.1038/ncomms7254.

GTEx Consortium, 2013. The genotype-tissue expression (GTEx) project. Nat. Genet. 45(6), 580–585. https://doi.org/10.1038/ng.2653.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., et al., 2014. The contribution of de novo coding mutations to autism spectrum disorder. Nature 515 (7526), 216–221. https://doi.org/10.1038/nature13908.

Lehman, A., Thouta, S., Mancini, G.M.S., Naidu, S., van Slegtenhorst, M., McWalter, K., et al., 2017. Loss-of-Function and gain-of-function mutations in KCNQ5 cause intellectual disability or epileptic encephalopathy. Am. J. Hum. Genet. 101 (1), 65–74. https://doi.org/10.1016/j.ajhg.2017.05.016.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., et al., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536 (7616), 285–291. https://doi.org/10.1038/nature19057.

Miceli, F., Striano, P., Soldovieri, M.V., Fontana, A., Nardello, R., Robbiano, A., et al., 2015. A novel KCNQ3 mutation in familial epilepsy with focal seizures and intellectual disability. Epilepsia 56 (2), e15–20. https://doi.org/10.1111/epi.12887.

Orhan, G., Bock, M., Schepers, D., Ilina, E.I., Reichel, S.N., Loffler, H., et al., 2014. Dominant-negative effects of KCNQ2 mutations are associated with epileptic encephalopathy. Ann. Neurol. 75 (3), 382–394. https://doi.org/10.1002/ana.24080.

Singh, N.A., Charlier, C., Stauffer, D., DuPont, B.R., Leach, R.J., Melis, R., et al., 1998. A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. Nat. Genet. 18 (1), 25–29. https://doi.org/10.1038/ng0198-25.

Tzingounis, A.V., Heidenreich, M., Kharkovets, T., Spitzmaul, G., Jensen, H.S., Nicoll, R.A., et al., 2010. The KCNQ5 potassium channel mediates a component of the after hyperpolarization current in mouse hippocampus. Proc. Natl. Acad. Sci. U. S. A. 107 (22), 10232–10237. https://doi.org/10.1073/pnas.1004644107.

Weckhuysen, S., Mandelstam, S., Suls, A., Audenaert, D., Deconinck, T., Claes, L.R., et al., 2012. KCNQ2 encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. Ann. Neurol. 71 (1), 15–25. https://doi.org/10.1002/ana.22644.

4.3.2. SECOND CASE

Unpublished data

In this paragraph I described an 11-year-old male patient with mild intellectual disability and behaviour immaturity in whom array-CGH analysis showed a heterozygous *de novo* chromosome 2q24.1 deletion. The deletion involves the *NR4A2* gene and the first untranslated exon of the *GPD2* gene.

Purpose: De novo deletions in 2q24.1 have been described in individuals with severe language impairment and reporting cognitive and behavioural symptoms. *NR4A2* is involved in several processes including development of the central nervous system, especially language-related brain regions.[l] *GPD2* encodes the

glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) which is located at the inner mitochondrial membrane and catalyses the conversion of glycerol-3-phosphate to dihydroxyacetone phosphate and, together with a cytosolic NAD-linked GPD, leads to the reoxidation of NADH formed during glycolysis. Mice with null mutation of both mitochondrial and cytosolic dehydrogenases exhibit hypoglycemic ketosis and glyceroluria and die within the first week of life.[II]

Recently, other patients with a similar phenotype have been described, eight harboring a 2q24.1 deletion encompassing *NR4A2* and *GPD2* genes, and three including only the *NR4A2* gene (see Table 2), thus suggesting that this gene might be responsible for the neurodevelopmental phenotype in these subjects.

Table 2. **Genotype-phenotype correlation of 2q24.1 deletion.**

We reported a review of literature and all information regarding patients with 2q24.1 deletion.

PATIENT	SEX	COGNITIVE DISABILITY	LANGUAGE IMPAIRMENT	BEHAVIOR PROBLEMS	ASD	MOTOR IMPAIRMENT	EPILEPSY	BRAIN IMAGING	DYSMORPHIC FEATURES	MUTATION	REFERENCE
1	F	ID	+	-	-	+	-	NA	+	- inh. interstitial 19p13 duplication - 3.9 Mb deletion on 2q23.3q24.1	Lybæk et al. 2009
2	M	ID	+	-	-	+	+	N	-	- inh. interstitial 19p13 duplication - 3.9 Mb deletion on 2q23.3q24.1	Lybæk et al. 2009
3	F	ID	+	-	-	+	-	NA	-	-inh. interstitial 19p13 duplication - 3.9 Mb deletion on 2q23.3q24.1	Lybæk et al. 2009
4	F	ID	+	-	+	-	-	NA	+	- de novo 298 kb deletion on 2q24.1	Barge-Schaapveld et al. 2012
5	F	DD	+	+	-	-	.	N	+	- de novo 5.4 Mb deletion on 2q23.3q24.1	Milani et al. 2015
6	F	ID	+	-	-	-	-	N	+	- de novo 4.5 Mb deletion on 2q23.3q24.1	Shimojima el al. 2016
7	NA	ID	+	-	+	-	-	NA	NA	- de novo1 Mb deletion on 2q24.1	Leppa et al. 2016/AGRE
8	NA	ID	+	-	+	-	+	NA	NA	- de novo3,3 Mb deletion on 2q24.1	Leppa et al. 2016/AGRE

9	NA	ID	+	-	+	-	-	NA	NA	- 4,6 Mb deletion on 2q24.1	Leppa et al. 2016/ ClinGen
10	NA	DD	+	+	-	-	-	macrocephaly	NA	- de novo 2.2Mb deletion on 2q24.1	Leppa et al. 2016/Decipher
11	NA	ID	+	+	-	-	-	NA	NA	- de novo170kb deletion on 2q24.1	Leppa et al. 2016/Decipher*
12	NA	DD	-	-	+	-	-	NA	NA	-1Mb deletion on 2q24.1	Leppa et al. 2016/Decipher
13	NA	NA	NA	NA	NA	NA	NA	NA	NA	- 4.7Mb deletion on 2q24.1	Leppa et al. 2016/Decipher
14	F	ID	+	-	-	-	-	NA	-	- de novo 89 kb deletion on 2q24.1	Reuter et al. 2017
15	M	ID	+	+	+	-	-	NA	-	- de novo 122 kb deletion on 2q24.1	Levy et al. 2018
16	F	ID	+	+	-	-	-	NA	-	- de novo 75 kb deletion on 2q24.1	Levy et al. 2018
17	M	ID	+	+	+	-	-	NA	-	- de novo 174 kb deletion on 2q24.1	Levy et al. 2018 *
18	F	NA	NA	NA	+	NA	+	NA	NA	- de novo heterozygous variant in NR4A2 (NM_006186): c.601_602insGTCC:p.P201Rfs*82	Guo et al. 2018
19	M	ID	+	-	-	-	+	N	-	- de novo heterozygous variant in NR4A2: c.326dupA: p.S110Vfs*2	Ramos et al. 2019
20	NA	NA	NA	NA	NA	NA	NA	NA	NA	- de novo heterozygous variant in NR4A2 (NM_006186): c.692delG:p.G231Afs*70	Feliciano et al. 2019

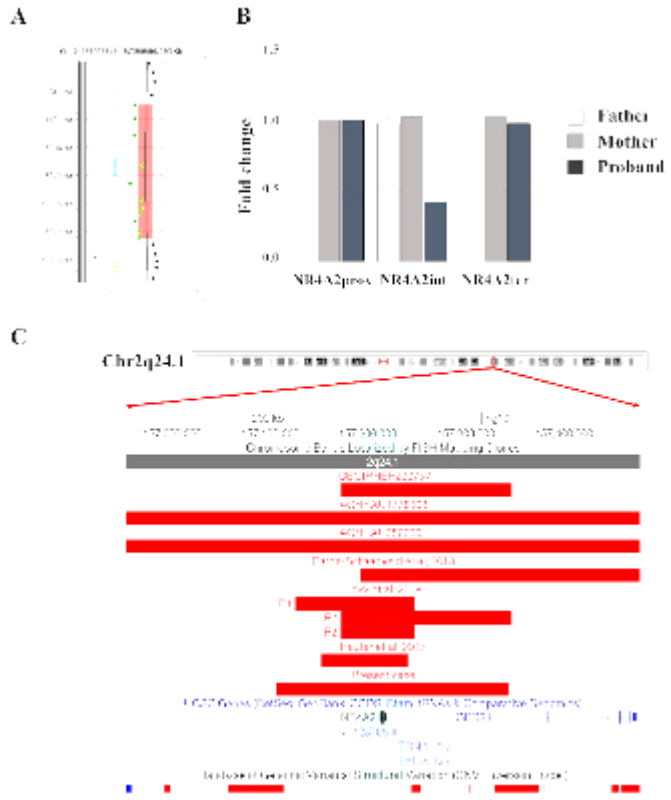
Legend: NA= not available, ID= intellectual disability, DD= developmental delay, ASD= autism spectrum disorder, inh= inherited

Indeed, *NR4A2* deleterious single nucleotide variants have been reported in patients with neurodevelopmental disorder.

Main findings: Our patient presented with mild intellectual disability, impairment of executive functions, hyperactivity and early-onset mild ataxia, expanding the clinical phenotype of this rare genetic variation. Array-CGH analysis of the subject and his parents showed the presence of a *de novo* interstitial microdeletion at 2q24.1, arr[GRCh37] 2q24.1(157,075,578_157,311,774)×1, (Figure 3A). The patients' CNV was validated by qPCR (Figure 3B). The deleted region contains 2 OMIM-reported genes: *NR4A2* (MIM 601828, Nuclear receptor subfamily 4, group A, member 2) and *GPD2* (MIM 138430, mitochondrial Glycerol-3-phosphate

dehydrogenase 2). Of note, terminal breakpoint localizes in the first intron of *GPD2* gene, that follows the first untranslated exon (Figure 3C).

Figure 3. Identification of a *de novo* chromosome 2q24.1 deletion in a patient with intellectual disabilities.



A) Array-CGH profile of chromosome 2 deleted segment. **B**) qPCR results for a region encompassed by the deletion (specific for the *NR4A2* gene: F: TTGAGGCGAGGACCCATACT; R: TACATGCTCTCTGACTGCCG), a 5' flanking region (specific for RefSeq Gene *LINC01876*: F: AGACCTGCGGAAAAGGTGAG*;* R: TCCACACCCTCAGATGCCTA), and a 3' flanking region (specific for intron 2 of the *GPD2* gene: F: GAGGGCTTTTAGCGTGGTCT; R: TCTGCTTGCTCCATTCCCAG) obtained in the proband and his parents. A region on chromosome 12 encompassing the *GAPDH* gene (NM_002046.3) was used as internal control to determine copy number and normalize primer efficiency. The copy number changes were calculated using the comparative DDCt method. Fold change of about 1 is expected for a diploid sample, and of about 0.5 for a haploid sample. **C**) Screenshot of chromosome 2 corresponding to the genomic position (chr2:156,593,532-158,119,396; NCBI build 37) and reporting the deletions carried by the present and already reported patients, copy number variations (based on DGV, filtered to show only deletions) (red bars); RefSeq genes and basic gene annotations from GENCODE present in the studied region.

In 2013, Barge-Schaapveld and collaborators first described a patient with a 298 kb deletion on chromosome 2q24.1 including the *NR4A2* and *GDP2* genes, presenting with autism and dysmorphic features.[III] Subsequently, 11 individuals with deletions at 2q24.1, containing 2–32 genes, respectively, have been reported, all showing neurodevelopmental disorders such as speech and language delay and ID.[IV, V, VI] In particular, a two years-old male with a *de novo* deletion of 2q24.1 (indicated as DECIPHER29057 in Figure

2C) presented a phenotype comparable to that of our proband, characterized by behaviour abnormalities without autism, dysmorphic features, and language impairment. Recently, 4 subjects with the same phenotype harboring 2q24.1 deletions encompassing only *NR4A2* (one including also the first exon of *GPD2*) have been reported [VII, VIII], thus narrowing the critical region to this single gene. Conversely, intragenic CNV deleting different exons of the *GPD2* were observed in 3 healthy individuals, as reported in DGV (Figure 2C), and *GPD2* loss-of-function variants were reported in about 100 non affected (control) individuals (GnomAD). Finally, a direct role of *GDP2* in brain development, intellectual disability and ataxia is not supported by *in vivo* functional studies.[I]

Interestingly, in the present subject we noticed hypotonia and early onset of mild ataxia. Although no developmental motor impairment was described for *NR4A2* haploinsufficiency, the only adult individual reported with a deletion of *NR4A2* had increasing gait unsteadiness, shaking on voluntary movements, and regression of motor skills and coordination, first noticed around 50 years of age. [V] As suggested by Reuter and collaborators, late-onset movement disorders can be potential features of *NR4A2* haploinsufficiency,[VIII] since variations in this gene were previously associated with susceptibility for Parkinson's disease, and studies in mice suggested Nr4a2-mediated protection against loss of dopaminergic neurons. [IX] Our case represents the first one with deletion of *NR4A2* showing early-onset of gait impairment, thus suggesting a potential role of this gene in developmental motor disorders. However, further studies on larger series of patients with *NR4A2* haploinsufficiency are required to shade light on the role of this gene in motor function development.

Principal conclusions: This case and the review of literature further support the role of *NR4A2* in neurodevelopmental disorders with high penetrance.

References for second case

I. Brown LJ, Koza RA, Marshall L, Kozak LP, MacDonald MJ. J Biol Chem. Lethal hypoglycemic ketosis and glyceroluria in mice lacking both the mitochondrial and the cytosolic glycerol phosphate dehydrogenases. 2002;277(36):32899-32904

II. Feliciano P, Zhou X, Astrovskaya I, et al. NPJ Genom Med. Exome sequencing of 457 autism families recruited online provides evidence for autism risk genes. 2019;4:19

III. Barge-Schaapveld DQ, Ofman R, Knegt AC, et al. Am J Med Genet A. Intellectual disability and hemizygous GPD2 mutation. 2013;161A(5):1044-1050

IV. Leppa VM, Kravitz SN, Martin CL, et al. Am J Hum Genet. Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. 2016;99(3):540-554

V. Lybaek H, Orstavik KH, Prescott T, et al. Eur J Hum Genet 2009. An 8.9 Mb 19p13 duplication

associated with precocious puberty and a sporadic 3.9 Mb 2q23.3q24.1 deletion containing NR4A2 in mentally retarded members of a family with an intrachromosomal 19p-into-19q between-arm insertion. 17(7):904-910

VI. Shimojima K, Okamoto N, Yamamoto T. Congenit Anom. Possible genes responsible for developmental delay observed in patients with rare 2q23q24 microdeletion syndrome: Literature review and description of an additional patient. (Kyoto) 2017;57(4):109-113

VII. Levy J, Grotto S, Mignot C, et al. Clin Genet. NR4A2 haploinsufficiency is associated with intellectual disability and autism spectrum disorder. 2018;94(2):264-268

VIII. Reuter MS, Krumbiegel M, Schluter G, Ekici AB, Reis A, Zweier C. Am J Med Genet A. Haploinsufficiency of NR4A2 is associated with a neurodevelopmental phenotype with prominent language impairment. 2017;173(8):2231-2234

IX. Liu H, Tao Q, Deng H, et al. Eur J Neurol. Genetic analysis of NR4A2 gene in a large population of Han Chinese patients with Parkinson's disease. 2013;20(3):584-587

5. A RE-EVALUATION OF ARRAY-CGH RESULTS IN 518 PATIENTS WITH NEURODEVELOPMENTAL DISORDERS SUPPORTS A TWO-HIT MODEL MEDIATED BY INHERITED OR *DE NOVO* COPY NUMBER VARIANTS

Published as:

Servetti M, **Pisciotta L**, Tassano E, Cerminara M, Nobili L, Boeri S, Rosti G, Lerone M, Divizia MT, Ronchetto P, Puliti A. Neurodevelopmental Disorders in Patients With Complex Phenotypes and Potential Complex Genetic Basis Involving Non-Coding Genes, and Double CNVs. Front Genet. 2021 Sep 21;12:732002. doi: 10.3389/fgene.2021.732002. PMID: 34621295; PMCID: PMC8490884.

Here is a summary of the work and the paper.

5.1. ABSTRACT

We aimed to unveil genes and pathogenic mechanisms explaining cases with neurodevelopmental disorders (NDDs) with inconclusive array-comparative genomic hybridization (array-CGH) results.

In a cohort of 518 NDD patients with non benign copy number variants (CNVs), we found 334 individuals with variants of uncertain significance (VOUS). We re-analyzed all results, with updated gene annotation and chromatin conformation data, and found likely pathogenic variants in 12 patients.

Notably, we observed patients with multiple variants, namely with one *de novo* variant and one variant inherited from unaffected parents, or with two variants transmitted, respectively, by each of their healthy parents. Inherited variants were classified as VOUS though overlapping known or candidate NDD genes. Conversely, these observations supported a double-hit pathogenic model, present in 3% of patients with VOUSs, that could fully explain their phenotype, and in 8% of patients with syndromic/recurrent CNVs, possibly explaining incomplete penetrance and variable expressivity associated with single variants. Gene enrichment showed that genes overlapped by the two CNVs, including new candidate genes, have a role in different biological processes with synergistic effect during neurodevelopment; in one patient, genes of two CNVs converged in the same pathway.

Overall, double-hit mechanisms seemed relevant in elucidating complex NDD phenotypes.

5.2. INTRODUCTION

Neurodevelopmental disorders (NDDs) are a heterogeneous class of brain disorders, including autism spectrum disorder (ASD) and intellectual disability, that affect about 1%-3% of children (Miller et al., 2010). The genetics of NDDs is complex and include copy number variations (CNVs) and deleterious variants in single genes (Coe et al., 2019). Despite the advances in the identification of risk genes for NDDs by whole exome and whole genome sequencing, the array comparative genomic hybridization (array-CGH) has long been used as the first tier diagnostic test for NDDs patients, resulting in an average diagnostic yield between 15% and 20% (Miller et al., 2010). Interpretation of CNVs is complex and, despite the use of standard classification guides (Riggs et al., 2020), most identified variants remain of uncertain significance (VOUS).

The goal of this study was to identify the most likely pathogenic microdeletion and microduplication associated with NDDs among diagnostic variants previously classified as VOUS. To this end, we studied 518 non benign CNVs using multiple bioinformatics resources and the synergistic effort of a team of neuropsychiatrists and geneticists.

We identified potential pathogenic variants in novel candidate genes for neurodevelopmental disorders. Furthermore, CNV-mediated double-hit mechanisms seem to play a relevant role in the neurodevelopmental disorders affecting our patients, especially accounting for complex phenotypes.

Indeed, these co-occurring hits involved already known NDD-associated genes or hypothetical novel NDD candidate genes enriched for pathways or biological processes known to be implicated in neurodevelopment.

5.3. PATIENTS AND METHODS

5.3.1. PATIENTS

In a retrospective review of data, we re-evaluated all the results of diagnostic array-CGH tests performed on 518 patients with NDDs admitted to the Gaslini Institute, Genova, Italy, over a 5-year period (2014-2018) focusing on variants previously interpreted as VOUS.

The clinical assessment of the patients comprised a thorough medical history also attentive to investigating systemic symptoms and sleep, neurological-behaviour examination, developmental/cognitive assessment, basic metabolic screening as detailed in Supplementary materials Table S1. Prior to array-CGH analysis, all patients were first evaluated by using different genetic tests for the common genetic causes of NDDs, including karyotype analysis, and tests for Fragile X syndrome. Array-CGH analysis was performed on DNA samples, extracted from peripheral blood, using a whole-genome 180 K Agilent array with ∼13 Kb overall median probe spacing (Human Genome CGH Microarray, Agilent Technologies, Santa Clara, CA, USA). Data were analyzed using Agilent CytoGenomics and genomic positions reported according to the human genome assembly (GRCh37/hg19). All detected CNVs were tested for inheritance by hybridization of the parental DNA with the same array platform. De novo CNVs were confirmed either by FISH analysis or by quantitative real-time PCR as already reported (Tassano et al., 2015; Vaccari et al., 2014). Written informed consent was obtained from the patient's parents or legal representative. This study was reviewed and approved by the Ethics Committee of the Italian Regione Liguria (R. P. 001/2019).

5.3.2. CNV DETECTION AND ANNOTATION

To assess the clinical significance of the detected CNVs, we used several tools. Aberration segments were reviewed using GRCh37 hg19 of UCSC Genome Browser (http://genome.ucsc.edu/index.html).

We annotated all detected copy number variations (CNVs) and CNV-encompassed genes across public databases.

Literature mining was also performed. We considered only variants present in <0.1% of control individuals. We re-annotated all CNV genes, entirely or partially encompassed by deletions, or entirely duplicated. We checked CNVs for the presence of dosage-sensitive genes (haploinsufficiency/triple sensitivity score) by using Clinical Genome Resource (ClinGen) consortium resource (https://dosage.clinicalgenome.org/). We also considered probability of intolerance to loss-of-function (LoF) variants (pLI). Genes with pLI scores of 0.9 or higher are extremely intolerant to heterozygous LoF variation, and thus haploinsufficient (Lek et al., 2016).

We then investigated deleted/duplicated regions for encompassing topologically associating domains (TADs), or their boundaries (TDBs). Indeed, TADs data from genome-wide higher order chromatin interaction data in human embryonic stem cells (h-ESC) and h-ESC derived neural progenitor cells (H1NPC) (Dixon et al., 2015), and brain cortex (Schmitt, Hu, & Ren, 2016), were downloaded (http://chromosome.sdsc.edu/mouse/download.html) and mapped to hg19 coordinates using the UCSC browser. We then investigated identified CNV-genes, not previously associated with genetic diseases, for relevance to neurodevelopmental disorders by considering different criteria: being expressed in the brain; reported mutations in the mouse associated with a neurodevelopmental phenotype; interaction with genes

known to be associated with a neurodevelopmental disorder. We searched for gene-specific mouse models presenting a neurodevelopmental phenotype using Mouse Genome Informatics (MGI) database.

5.3.3. GENE ENRICHMENT AND PROTEIN-PROTEIN INTERACTION ANALYSES

We used GeneCodis4 (Tabas-Madrid, Nogales-Cadenas, & Pascual-Montano, 2012) to unveil enrichment of annotations, as already reported (Cerminara et al., 2021). Briefly, all the genes obtained from either duplicated or deleted regions were used as input in GeneCodis4 together with known NDD-associated genes, as those related to ASD and reported in SFARI database and those reported in OMIM as associated with intellectual disability.

5.3.4. PATHOGENICITY EVALUATION

All CNVs were initially classified as pathogenic or VOUS according to standard classification (Riggs et al., 2020). Pathogenic CNVs were represented by *de novo* variants of the following categories: syndromic CNVs, encompassing a specific region already reported in DECIPHER CNV Syndromes and/or in recurrent CNV SFARI databases; CNVs larger than 3 Mb; CNVs smaller than 3Mb but encompassing a gene already associated with NDDs according to OMIM, SFARI databases, and/or reported in PubMed; all recurrent gene-specific CNVs, including intragenic deletions/duplications, based on OMIM, SFARI databases, and/or reported in PubMed. Inherited CNVs reported in SFARI database, or in the literature, with incomplete penetrance were also included in this class. Remaining patients with VOUS variants were further studied and grouped on the basis of the following criteria:

(i) patients with multiple CNVs, enclosing NDD genes, which could have additive effects on the patients' phenotype;

(ii) patients with CNVs encompassing regions that control expression of known NDD genes, including those with potential long-range effects;

(iii) patients whose CNVs encompass genes not yet reported as NDD-associated, but whose expression profile and function suggest a possible involvement in the disease, and at least another patient reported with a similar CNV and phenotype in publically available databases. When, after re-annotation, the clinical interpretation remained uncertain, CNVs were classified as VOUS. This class generally enclosed inherited and *de novo* CNVs that do not contain candidate genes according to adopted criteria described above (Figure S1). Results of this analysis were reported in figures using UCSC and specific Custom Tracks. In particular we used RefSeq Curated for gene representation and OMIM genes to underline genes already associated with disease. Decipher track was manually modified; we selected patients with CNVs similar to those of our patients, that is either deletions or duplications according to those found in our patients and overlapping approximately the same region with the same genes. We selected Decipher CNVs only if their related patients had a NDD phenotype (we kept out patients for whom we had no information about their phenotype). The CNVs thus

selected were further investigated for being present in the same patient with other secondary CNVs overlapping additional NDD genes, and in this case reported in figures (marked with letters). Developmental Delay case track includes manually selected CNVs similar in size and gene content to those of our cases; we proceeded analogously also for Developmental Delay Control CNVs. A track reporting SFARI genes was added to underline the presence of a known ASD associated gene in the genomic region shown in figures.
Brain-expressed gene track represents genes with a prevalent expression in the brain according to the BrainSpan data processed as detailed above, useful to highlight possible new NDD candidate genes.

5.4. RESULTS

We re-evaluated the results of diagnostic array-CGH tests on 518 patients (346 males and 172 females) showing at least one non-benign CNV, initially classified in a group of patients with a pathogenic variant (184) and a group with a VOUS variant (334). In order to ameliorate the genotype-phenotype correlation, a re-evaluation of clinical data was also performed considering the phenotypic features and genetic/instrumental tests, collecting detailed data and annotating them also with updated clinical information (see Table 1 and Table S1).

Table 1. Major clinical features

ID	Major clinical features
IGGAC01	Psychomotor Delay
IGGAC02	Psychomotor Delay with Microcephaly and Epilepsy
IGGAC03	Intellectual Disability
IGGAC04	Psychomotor Delay with hypotonia and Epilepsy
IGGAC05	Intellectual Disability and mild cerebral dysmorphism
IGGAC06	Autism Spectrum Disorder
IGGAC07	Psychomotor Delay with Behavioural disturbances
IGGAC08	Speech Disorder and Hyperactivity
IGGAC09	Intellectual Disability and Epilepsy
IGGAC10	Speech Disorder
IGGAC11	Autism Spectrum Disorder and Psychomotor Delay
IGGAC12	Autism Spectrum Disorder with language regression
IGGAC13	Psychomotor Delay with Behavioural disturbances
IGGAC14	Mild Intellectual Disability
IGGAC15*	Psychomotor Delay with Behavioural disturbances
IGGAC16*	Mild Intellectual Disability and Speech Disorder
IGGAC17	Psychomotor Delay and Speech Disorder
IGGAC18	Autism Spectrum Disorder
IGGAC19	Psychomotor Delay with mild Behavioural disturbances and Epilepsy
IGGAC20	Autism Spectrum Disorder and Psychomotor Delay

IGGAC21	Speech Disorder and mild Behavioural disturbances
IGGAC23	Autism Spectrum Disorder
IGGAC24	Speech disorder and Hyperactivity with brain malformations

After CNV re-annotation, and following standard criteria (Riggs et al., 2020) (see Methods and Figure S1), patients were re-classified into those with pathogenic CNVs (215 cases), and those with VOUS variants (303 cases). More in detail, patients with pathogenic variants included cases with a large genomic rearrangements (> 3 Mb)(5% of total patients); a known syndromic variant (11%); a recurrent CNV in a known NDD gene (14%); a CNV of less than 3 Mb and encompassing genes known to be associated with ASD and/or ID (5%). We also included in this group 6% (31) of cases carrying a CNV inherited from unaffected parents, overlapping genes recently associated with NDDs with reduced penetrance.

The remaining cases (303 individuals) presented CNVs initially classified as VOUS that were further and deeply analysed to search for new candidate NDD genes and/or new genetic mechanisms that could account for neurodevelopmental phenotypes (Figure S1). We revealed possible pathogenic CNVs in 12 patients which resulted in a 2% increase of patients diagnosed with a pathogenic/likely pathogenic variant. We observed patients with multiple variants supporting a double-hit pathogenic model in 3% (8) of patients with VOUSs and 8% (11) of patients with syndromic/recurrent variants, overall accounting for about 4% of total patients with either VOUS or syndromic/recurrent CNVs.

5.4.1. ADDITIVE EFFECTS OF CNVS INVOLVING KNOWN NDD GENES

Among patients with VOUS variants, four presented with multiple CNVs overlapping known NDD genes and inherited from unaffected parents. We thought that in these cases, CNVs could additively contribute to the patient's phenotype.

The first case**,** IGGAC13, is a male patient with a maternal deletion of chromosome 7 involving *CNTNAP2* and a paternal deletion of chromosome 11 involving *LRRC4C* (Figure 1 a,b). Both genes are present in the SFARI

database. Few CNVs are reported to overlap these two genes in control individuals, and in NDD patients according to Developmental Cases and Decipher databases. Of note, NDD patients reported in Decipher and represented in Figure 1 carried one additional CNV involving other NDD-associated genes. Indeed, a case has been previously reported with two deleterious variants, one of which disrupting *LRRC4C* and one affecting another SFARI-reported gene, *DPP6* (Maussion et al., 2017). The two variants, as in our case, were inherited one from the mother and one from the father, both unaffected.

We hypothesized that, while half dosage of each of these two molecules is compatible with a normal development, half dosage of both genes can impair synapse correct development and function. In fact, gene enrichment and String analysis showed that both genes are members of the KEGG cell adhesion molecule pathway and interact with other known NDD genes (Figure 1c, Table S4).

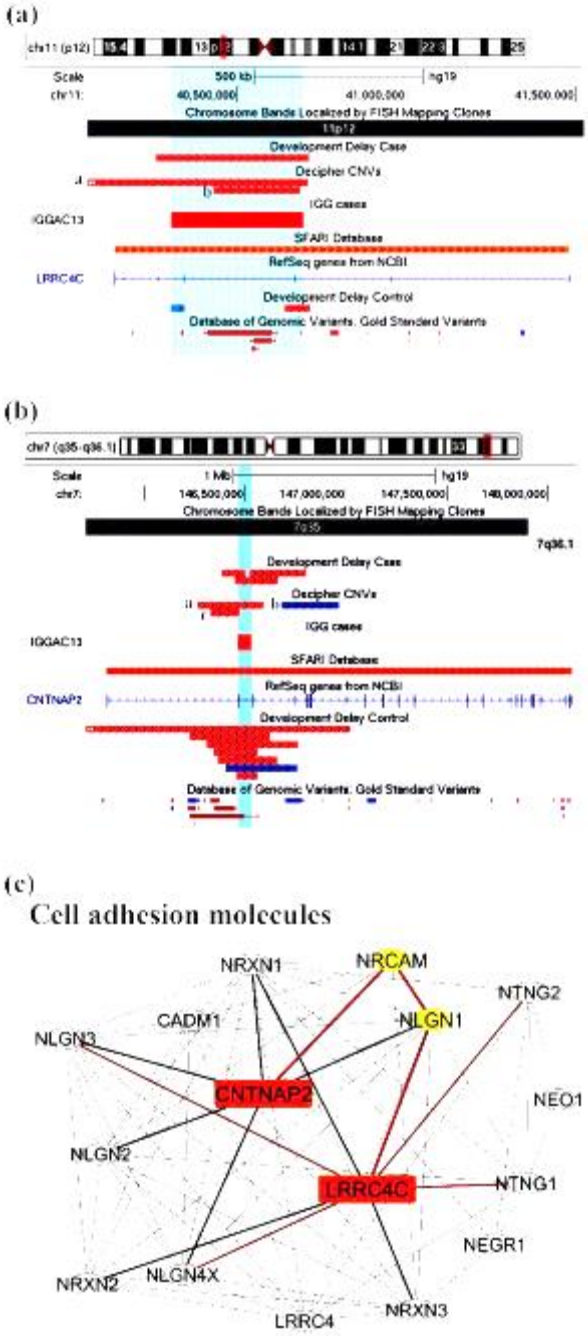


Figure 1. Additive effects between *CNTNAP2* and *LRRC4C* in patient IGGAC13

The second case is a male patient (IGGAC14) that inherited from his healthy father a deletion of chromosome 17 encompassing a SFARI gene, *DNAH17,* and its antisense gene *DNAH17-AS1*.

Gene enrichment evidenced a function of *SHOX* in the regulation of transcription by RNA polymerase II, and of *DNAH17* in microtubule-based movement, two important biological processes in neurodevelopment disorder (Figure 4, Table S4). As shown in Figure S2a, other patients with NDD phenotype are reported in Decipher database to have duplications of *SHOX* and a secondary CNV involving neurodevelopmental genes. Indeed, a deletion similar to that of our patient was reported in Decipher database in a patient with NDD phenotype having a secondary CNV also overlapping developmental genes (Figure S2b).

We observed a family with two affected siblings, a male (IGGAC15**)** and a female (IGGAC16**)**, who inherited from their unaffected mother a deletion encompassing a syndromic region at 15q11.2 (Figure S3). The sibs also shared a duplication involving the SFARI gene *SYNGAP1*. Of note, the *SYNGAP1* duplication was not observed in parents, at least in their blood-extracted DNA, thus leading to suppose mosaicism in either father or mother. Patient IGGAC16 also inherited a paternal, recurrent deletion involving the SFARI gene *RBFOX1* (Figure S3c). Interestingly, loss of function variants of *SYNGAP1* have been identified in patients with ASD and intellectual disability with or without epilepsy.

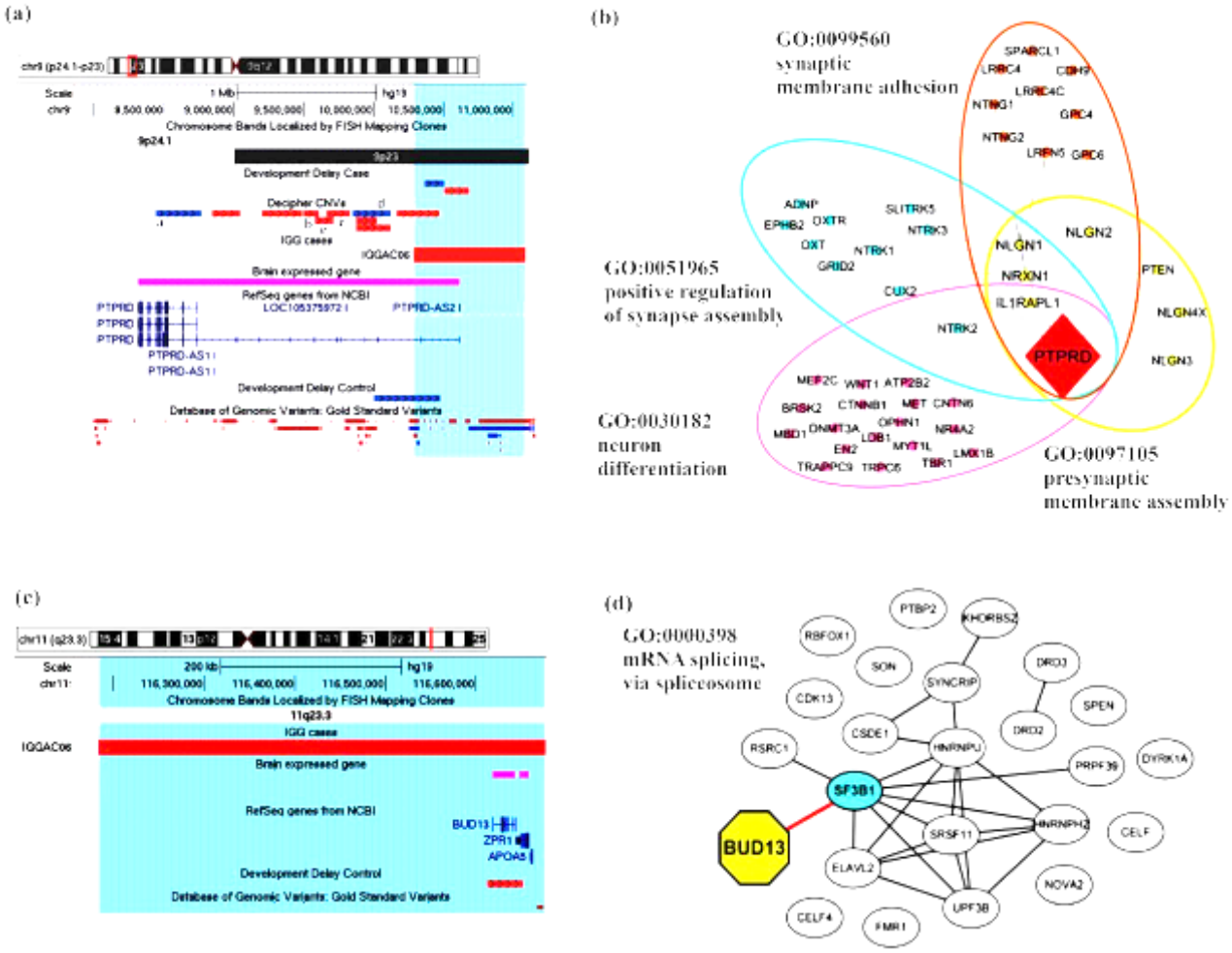
However, to the best of our knowledge, no cases with duplication or gain of function variants have been reported so far. Overexpression of SynGAP was reported to block neurite outgrowth by a mechanism that involves Ras-like GTPase cascade (Tomoda, Kim, Zhan, & Hatten, 2004), to produce a depression of AMPAR-mediated excitatory postsynaptic currents in neurons (Rumbaugh, Adams, Kim, & Huganir, 2006), and to regulate decreasing and increasing miniature excitatory synaptic currents in hippocampal neurons (McMahon et al., 2012). In fact, some phenotypic differences between the two sibs were observed. The male patient showed a mild phenotype characterized by a borderline cognitive functioning associated with atypical behaviours and psychomotor delay in early infancy, his sister had a more complex phenotype with mild cognitive impairment, irregular sleep rhythms, hyperactivity, coordination and language disorders. We could speculate that the two CNVs shared by the two siblings can explain the phenotypic features shared by the two siblings, while the third CNV could have worsened the phenotype of the female patient.

5.4.2. ADDICTIVE EFFECTS OF CNVS INVOLVING POTENTIAL NEW CANDIDATE GENES AND PATHOGENETIC MECHANISMS

We assessed whether some cases among the group with VOUS variants reported double CNVs involving genes or genomic regions not yet described as NDD-associated but that could act in concert, similarly to the four cases described above, with a potential impact on the patients' phenotype. We found four cases .

The first case (IGGAC06) carried a maternal deletion of chromosome 9 and a non-maternal deletion of chromosome 11; as genetic analysis of the patient's father was not available, we can not determine if the chromosome 11 deletion was *de novo* or paternally inherited (Figure 2).

Figure 2.



Additive effects between *PTPRD* and *BUD13* in patient IGGAC06

A total of 7 deletions encompassing exonic regions of *PTPRD* are reported in controls, while several CNVs and, in particular, 16 deletions and one intragenic duplication involving coding regions of the gene are present among cases in Developmental Delay and Decipher databases with NDD phenotypes. Five among Decipher cases have an additional CNV involving brain expressed genes that, analogously to those reported in patient IGGAC06, could have an additive effect in causing patient's phenotype. Interestingly, *PTPRD* is predicted to be intolerant to loss of function variants (pLI=1) and to haploinsufficiency mechanisms (score= 0.75; value above 0.5 are predicted to be haploinsufficient) (Shihab, Rogers, Campbell, & Gaunt, 2017). *PTPRD* has a role in synaptic adhesion and synapse organization and can bidirectionally induce pre- and postsynaptic differentiation of neurons by trans-synaptically binding to interleukin-1 receptor accessory protein (*IL1RAP*) (Yamagata et al., 2015; Yoshida et al., 2012). In the mouse, *PTPRD* was shown to regulate neurogenesis (Tomita et al., 2020) and mice lacking *PTPRD* showed impaired learning with enhanced hippocampal long-term potentiation (Uetani et al., 2000), thus suggesting a role for *PTPRD* in

neurodevelopmental disorders. In OMIM database, *PTPRD* is not associated with a disorder, and its implication in human disease is still to be completely clarified. Homozygous deletions of *PTPRD* were observed in a patient with intellectual disability and trigonocephaly (Choucair et al., 2015) and his family members carrying the same deletions were reported to be unaffected. Instead, a heterozygous *de novo* splicing variant of the gene was reported in a girl with moderate nonsyndromic developmental delay (Yan et al., 2019). Furthermore, genetic studies showed an oligogenic association of *PTPRD* variants with obsessive-compulsive disorder (Mattheisen et al., 2015) and with restless leg syndrome (Schormair et al., 2008). Overall, results from various genetic studies are consistent with both major/oligogenic and modest/polygenic contributions of common and rare *PTPRD* variations in neurological, behavioural, and neurodevelopmental disorders (Uhl & Martinez, 2019). Gene enrichment and String analysis showed *PTPRD* interacting with many NDD genes playing a role in synaptic membrane adhesion and neuronal development (Figure 2b, Table S4). The chromosome 11 deletion encompassed *BUD13,* which encodes a component of the retention and splicing (RES) complex of the spliceosomal complex. In zebrafish, the lack of *BUD13* caused defects in intron splicing specifically in genes with neurodevelopmental regulatory functions, thus resulting in a decrease of differentiated neurons and brain developmental defects (Fernandez et al., 2018).

GeneCodis analysis showed an enrichment of *BUD13* with various NDD genes in the GeneOntology category of mRNA splicing, via spliceosome, a biological process with an important role in neurodevelopment (Figure 2d, Table S4).

Patient IGGAC08 has a complex phenotype mainly characterized by language disorder, hyperactivity, learning difficulties, borderline cognitive level, motor skill impairment, and short stature (Table S1). His mother also had school difficulties and showed dyslexia in childhood. Mother and son share a deletion of chromosome 4q34.1 encompassing *GLRA3* which encodes the alpha-3 subunit of the neuronal glycine receptor, a ligand-gated ion channel. In addition, patient IGGAC08 had a *de novo* deletion involving *hsa-mir-4465* and a *de novo* duplication including *INPP5A* (Figure S4). *GLRA3* receptors are expressed in spinal cord, brainstem, hippocampus, amygdala, striatum, and cortex. The glycine receptor is a glycoprotein composed of 5 subunits, three α and two β subunits. The alpha subunits bind the glycine ligand while the beta subunits bind to gephyrin, a cytosolic protein required for a regulated synaptic aggregation and clustering of these receptors. Gephyrin has functional links with several synaptic proteins, mutations of which have been reported in various neurodevelopmental disorders (Choii & Ko, 2015; Kim et al., 2021). Half dosage of *GLRA3* may contribute to impair gephyrin-mediated aggregation and post-synaptic clustering.

Loss of *PTEN*, a SFARI gene, can cause postsynaptic and presynaptic changes in excitatory and inhibitory connectivity. In vitro analyses have recently demonstrated that a microRNA (microRNA- 301a) can downregulate *PTEN* after glycine receptor activation (Chen et al., 2016). Tao et al., 2019 showed that *miR-4465* significantly inhibited the expression of *PTEN*, upregulated phosphorylated AKT, and ultimately inhibited autophagy by activating mTOR in HEK293, HeLa, and SH-SY5Y cells (Tao et al., 2019). Half dosage of

hsa-mir-4465 in our patient is expected to cause upregulation of *PTEN,* contributing to postsynaptic impairment. MicroRNAs are important regulators of brain development and neuronal function, and have been associated with a variety of nervous system diseases, including ASD (Cheng et al., 2018; Wu, Li, & Zheng, 2020).

In mice, the deletion of *Inpp5a* causes perinatal lethality in 90 % of the homozygous mutants, and early onset ataxia and relatively small stature in surviving mutants. Heterozygotes do not exhibit obvious motor coordination impairment unless challenged in motor skill test (A. W. Yang, Sachs, & Nystuen, 2015). *Inpp5a* is a downstream effector of signalling from mGlu1 receptor, one of the two members of group I metabotropic glutamate receptors, that regulates synaptic plasticity, particularly in cerebellar Purkinje cells. The mGlu1 receptor activity needs to be fine-tuned and balanced for normal motor coordination. Alteration of expression, either enhancement or decrease, of group 1 mGlu metabotropic receptors in mice is known to alter their activity, which results in defects of motor coordination (Bossi et al., 2018; Rossi et al., 2013). The deletion of *INPP5A* in this patient may be in part responsible for his impairment in motor skills.

IGGAC10 patient showed a complex phenotype mainly characterized by intellectual disability and epilepsy (Table S1). He carried a maternal deletion encompassing *SYNCRIP* and a paternal duplication involving the upstream regulatory genomic region of *ADCY5*. Both CNVs overlap a TDB between two flanking TADs described in different tissues, including brain cortex, which could also contribute to dysregulating the expression of implicated genes (Figure S5).

Missense and truncating variants of *SYNCRIP* were found in ASD patients with more severe phenotypes (Guo et al., 2019) and in patients with severe non-syndromic sporadic intellectual disability (Rauch et al., 2012), respectively. *ADCY5* encodes a member of the membrane-bound adenylyl cyclase enzymes which mediate G-protein-coupled receptor signalling through the synthesis of the second messenger cAMP. According to SFARI database, missense and frameshift variants of *ADCY5* have been reported in ASD patients. Gain of function variants of *ADCY5* have been associated with a broad range of movement disorders, most notably chorea, dystonia, and myoclonus (Vijiaratnam, Bhatia, Lang, Raskind, & Espay, 2019).

Thus, the two CNVs could cause, by a mechanism implicating also a modification of chromatin conformation, a dysregulation of both *SYNCRIP* and *ADCY5* contributing to the complex phenotype observed in this patient.

IGGAC07 patient carried a *de novo* duplication of chromosome 14, and two maternally inherited duplications of chromosome 12 and chromosome 22, respectively (Table 2, Figure S6).

Interestingly, chromosome 14 duplication encompassed *ABHD4*, a gene recently shown to be involved in developmental anoikis, a mechanism of cell death in the prenatal brain preventing from survival of misplaced cells (Laszlo et al., 2020). Indeed, *ABHD4* is involved in N-acylethanolamine and related N-acyl-phospholipid metabolism including the production of the endocannabinoid molecule anandamide which has various physiological functions including the regulation of synaptic plasticity and apoptosis. *ABHD4* expression is tightly controlled spatially and temporally in the developing brain and a misregulation of its expression

because of the duplication could contribute to the patient's phenotype. Only a duplication encompassing *ABHD4* was annotated in one control individual, while various cases were reported both in Decipher and in Developmental Delay databases with duplications larger than that observed in our patient and a smaller one sharing with our patient *ABHD4* duplication.

The maternal duplication on chromosome 22 involved a region with triplo-sensitivity including *GNAZ,* the Guanine Nucleotide-binding protein (G protein) Alpha Z polypeptide encoding gene. In this region, many duplications are reported in both Decipher and Developmental Delay databases, most of them extending over the short region containing *GNAZ*. Only a total of 4 control individuals are reported in either Developmental Delay Control or in DGV databases with duplications overlapping *GNAZ*. *GNAZ* is a member of the Gαi subfamily of heterotrimeric G proteins, and couples to Gprotein- coupled receptors triggering important pathways in neuronal development (Hultman, Kumari, Michel, & Casey, 2014). The duplication of chromosome 12 entirely encompassed *WSCD2*, which encodes a WSC Domain-containing protein with a still unclear function. Interestingly, genome wide association analyses found significant results for variants in *WSCD2* in patients with psychiatric disorders (Lo et al., 2017). Finally, we could hypothesise that all the *de novo* and the inherited duplications could have jointly contributed to the NDD phenotype of patient IGGCA07.

5.4.3. ADDITIVE EFFECTS AMONG SYNDROMIC AND GENE-SPECIFIC RECURRENT CNVS

We evaluated in our cohort of patients, we could also find cases with secondary CNVs potentially modulating phenotypes otherwise due to syndromic or recurrent gene-specific CNVs. We found that 9% (5/57) and 8% (6/74) of patients respectively with a syndromic and with a recurrent gene-specific CNV carried a secondary CNV involving known NDD genes or brain-expressed genes potentially implicated in the disease (Table S2). In these cases, the two involved variants were inherited one from the mother and one from the father, both unaffected, or one of the two occurred *de novo*. Of note, one patient (IGGAC05) presented with two syndromic CNVs frequently observed in NDD patients, 15q13.3 deletion and 16p11.2 duplication, and both of them originated *de novo*.

In some cases, the secondary CNV was not a recurrent one and did not involve a known NDD gene.

Thus, to detect potentially candidate genes, we performed a gene enrichment analysis with CNV overlapped genes expressed in the brain, and known NDD-associated genes. We showed in Table S2 all cases presenting a secondary CNV involving a known NDD gene or a potential NDD candidate gene as determined on the basis of gene enrichment analysis.

5.4.4. CNV-MEDIATED COMPLEX PATHOGENIC MECHANISMS ASSOCIATED WITH COMPLEX NDD CASES

Among VOUS variants, we assessed whether some CNVs could affect expression of NDD genes throughout an indirect and/or long range effect, and explored, also through a deep clinical re evaluation of patients, if complex phenotypes observed in some patients could be explained by involvement of more than one gene.

We found one case (IGGAC04), presenting with very severe psychomotor delay, hypotonia, and epilepsy (Table 1 and Table S1), carrying a *de novo* heterozygous deletion of chromosome 5 (Figure 3a). This deletion encompasses 6 genes, two of them expressed in the nervous system: *MEF2C-AS1*, and *ADGRV1*.

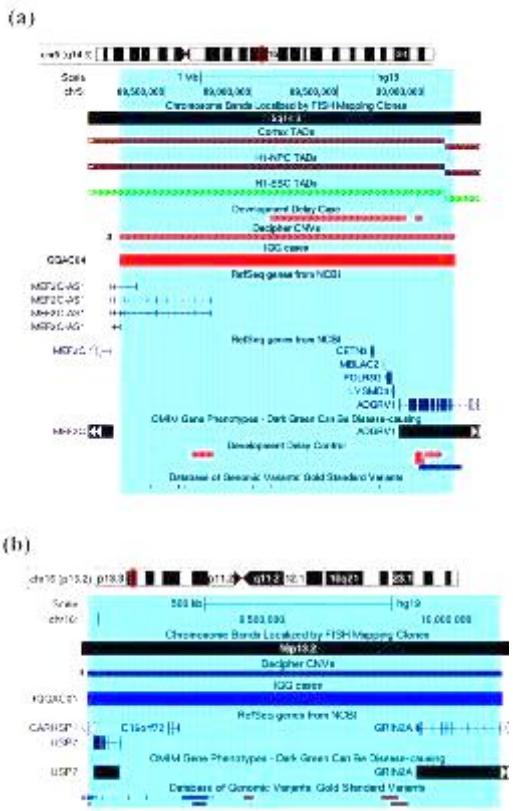


Figure 3. Potential candidate genes and pathogenetic mechanisms in patient IGGAC04 and IGGAC01

It has been recently demonstrated that, during myogenic differentiation a lncRNA, the *OIP5-AS1*,
through its complementarity with *MEF2C* 3'UTR, enhances *MEF2C* mRNA stability, thus promoting *MEF2C* expression and ultimately enhancing myogenesis (J. H. Yang et al., 2020). We hypothesized that during neurogenesis, analogously to *OIP5-AS1*, *MEF2C-AS1* could enhance *MEF2C* mRNA stability and promote *MEF2C* expression and neurogenesis. In this view, the heterozygous deletion observed in patient IGGAC04 could reduce the expression of *MEF2C-AS1* causing a decrease of *MEF2C*, thus interfering with neuronal differentiation and contributing to the patient's phenotype.

The deletion of *ADGRV1* could be responsible for the epilepsy phenotype. Of note, we found one patient in the Decipher database (Patient 251716) carrying a deletion overlapping that of our patient and having similar clinical features characterized by intellectual disability, muscular hypotonia, and seizures. No deletions overlapping the same region were present among controls (Developmental Delay and DGV Controls). The chromosome 5 region deleted in these patients overlaps a TDB described in different tissues, including those relevant for neurodevelopment (e.g. Cortex), according to their relative Hi-C maps (Figure 3A), which could further complicate the fine regulation of expression of the overlapped genes, including *MEF2C-AS*.

One case (IGGAC01) showed a complex phenotype mostly characterized by psychomotor delay, autistic features, hypotonia, epilepsy, gastrointestinal and sleep disturbances (Table 1 and Table S1), and carried a *de novo* heterozygous duplication of chromosome 16 (Figure 3b). The duplication involved, in particular, two

OMIM reported genes, *USP7,* encoding the Ubiquitin Specific Peptidase 7, and *GRIN2A,* encoding the NR2 subunit of the N-methyl-D-aspartate (NMDA) glutamate receptor.

Deletions, truncating and missense variants of *USP7* have been associated with the Hao-Fountain syndrome (OMIM 616863), characterized by speech delay, autistic spectrum disorder, attention-deficit hyperactivity disorder, sleep disturbances, and gastroesophageal reflux disease (Fountain et al., 2019). All these clinical features are present in our case who has the complete duplication of *USP7*. Notably, in syndromic forms of ASD, duplications in addition to deletions of *USP7* have been reported (Sanders et al., 2011), thus supporting the idea of dosage sensitivity, with both too little and too much *USP7* causing imbalances of neuronal homeostasis. The duplication present in patient IGGAC01 partially overlapped *GRIN2A*. Truncating as well as missense, activating variants of *GRIN2A* have been associated with an idiopathic form of focal epilepsy, EEG continuous spike-and-wave during sleep (CSWS), and speech disorder (OMIM 245570) (Lemke et al., 2013) whose clinical aspects are present in our patient. Different molecular alterations of NMDA receptor subunits seem to result in a deleterious dysregulation of NMDA receptor function (Balu & Coyle, 2011). Although the effect of partial duplications on gene function are not easily predictable, some clinical features of the IGGAC01 patient suggested the occurrence of a dysregulation of *GRIN2A* that could have contributed to the patient's phenotype. Of note, a patient with phenotype resembling that of patient IGGAC01 was reported in Decipher database with a similar duplication, including both *USP7* and *GRIN2A* (Figure 3b). We thus hypothesised the phenotype of these two patients could result from the impairment of both *USP7* and *GRIN2A.*

IGGAC24 had a *de novo* duplication of chromosome 19 that encompasses various genes, including 11 expressed in the brain, and a TDB described in the brain cortex and other different tissues, which could contribute to dysregulating gene expression of implicated genes (Figure S7a). In an attempt to prioritize enclosed genes and define possible candidate genes, we performed a gene enrichment analysis as described above (see Methods). Two genes emerged for their functions known to play a role in neurodevelopmental disorders, namely *SHISA7* and *U2AF2*.

Patient IGGAC23 showed a *de novo* duplication overlapping 7 brain-expressed genes and a TDB described in the brain cortex and other tissues (Figure S7b). After gene enrichment analysis of these 7 genes with known NDD genes, one emerged as the most interesting, *CTBP2*. The duplication present in patient IGGAC23 could have caused overexpression and/or dysregulated expression of overlapped genes, among which *CTBP2*, whose overexpression has already been associated with impaired neuronal defects.

5.5. DISCUSSION

The main goal of this study was to identify the most likely genes and genetic mechanisms underlying the NDD phenotype through the re-evaluation of microdeletions and microduplications previously classified as variants of uncertain significance in our cohort of patients. We re-analysed non-benign CNVs in 518 patients

in the light of newly acquired information on gene expression and function and on chromatin organization features of involved genomic regions. We also assessed the presence of potential deleterious CNVs acting in concert in the same patients, according to a two-hit model emerging from the recent

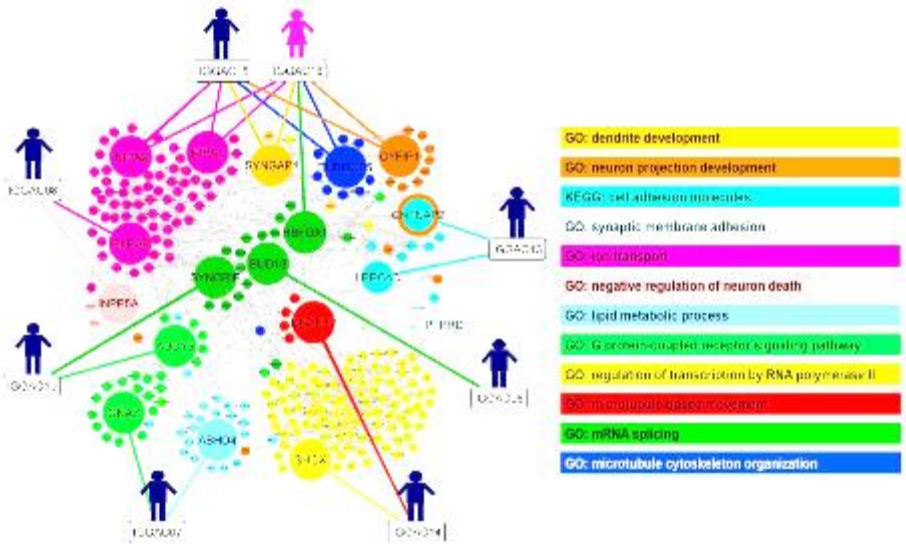


Figure 4. Overview of biological processes and pathways implicated in patients with CNV-genes acting according to double-hits model

literature as a possible pathogenetic mechanism with important roles in neurodevelopmental disorders. We also re-evaluated clinical data of patients to better interpret CNV causal effects in complex cases. We found potential deleterious variants in 2% (12) of patients previously classified with VOUS variants (Figure S1).

One important finding emerging from this re-evaluation of diagnostic array-CGH results is the presence of co-occurring CNVs that, according to the two-hit model, could have contributed to the patients' phenotype. In eight patients, we found two CNVs overlapping either known NDD genes and/or genes with potential roles in neurodevelopment, representing 3% (8/303) of patients with VOUS variants. In three patients, we found one *de novo* and one variant inherited from unaffected parents, the remaining five patients inherited one CNV from each of their unaffected parents. In three out of eight cases, namely IGGAC06, IGGAC07, and IGGAC08, involved genes have not been associated yet with disease and we presented data in favor of their implication in neurodevelopmental disorders. Indeed, we searched the Decipher database for cases with one CNV corresponding to that presented by our patients and a secondary CNV, involving genes with a role in neurodevelopment. We found some cases, reported in the figures (Figure 1, Figure 2, Figure S2, Figure S6), that further support the hypothesis that identified co-occurring CNVs in our patients act additively causing deleterious phenotypes according to a two-hit model.

In order to gain insight into possible interaction among genes implicated in the CNVs co-occurring in the same patient, and to evaluate whether these genes share molecular functions or biological processes, we performed gene enrichment and protein-protein interaction analyses (Figure 4, Table S4).

We also tested interaction with other known NDD associated genes, which helped to better define the role of new candidate genes in important neurodevelopmental pathways and biological processes.

We thus hypothesized that, in all these cases, the double-hit mechanism could involve two pathways/biological processes playing an important role in neurodevelopment. Although individually associated with NDD biological processes, genes affected by co-occurring CNVs require further functional analyses to understand their potential interactions and ultimately to clarify their contribution to patient phenotype.

Of note, co-occurring secondary CNVs overlapping NDD genes were also found in patients with syndromic or recurrent gene-specific CNVs, which could modulate penetrance and/or severity of the disease.

5.6. CONCLUSION

These findings support the usefulness of a continuous review of results of diagnostic genetic tests, including array-CGH, and suggest that genetic diagnosis of complex disorders requires complete evaluation of all the CNVs present in patients, and their parents, considering possible synergistic effects.

5.7. ADDITIONAL INFORMATIONS

DATA AVAILABILITY STATEMENT

Clinical and genetic variant details have been deposited in the Decipher database at https://www.deciphergenomics.org/. Accession numbers are reported in Table S5.

ACKNOWLEDGEMENTS

The authors wish to thank all families and referring nurses of Child Neuropsychiatric Unit for their help and collaboration. Funding for this research was provided by Italian Ministero della Salute (Project N. RF-2016-02361949 to A.P.; “Cinque per Mille” and Ricerca Corrente). This work was developed within the framework of the DINOGMI Department of Excellence (MIUR 2018-2022, legge 232 del 2016).

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

MS, LP and AP participated in the design of the study. PR and ET were responsible for array-CGH experiments. PR, MC and AP performed some array-CGH validation experiments. ML, MTD, LP, SB and LN collected all clinical phenotypes and MS managed the database of clinical and CNVs data.

MS, LP, MC and AP participated in data analysis and interpretation. MS, LP and AP wrote the manuscript. All authors critically revised the manuscript, and agreed to be accountable for the content of the work.

5.8. SUPPLEMENTARY MATERIALS

Table S1. Clinical Features

Patient ID	ASD	Intellectual disability (ID)	Other NDDs	Neurological signs/disorders	Others	Results of other laboratory tests
IGGAC01	NO	ID	Language disorder, Coordination disorder, Atypical behaviours, Attention disorder and Hyperactivity, Self-aggressivity	Hypotonia Epilepsy, Sleep disturbances	Gastrointestinal disorders,dysphagia and selective feeding	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC02	NO	ID	Language disorder	Microcephaly, Febrile convulsions, EEG abnormalities	Hyporexia, Breathing disorders	Karyotype, Fragile X analysis, *CFTR* gene analysis, globin genes analysis: negative
IGGAC03	NO	ID	Language disorder	EEG abnormalities	NO	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC04	NO	ID	NO	Hypotonia and Epilepsy	Prenatal gastrointestinal perforation	Normal blood amino acid, organic acids levels, acylcarnitine profile, guanidinoacetic acid. Karyotype, Fragile X analysis: negative
IGGAC05	NO	ID	Language disorder and Attention disorder	Mild EEG abnormalities	Brain midline defects	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative
IGGAC06	YES	ID	NO	Sleep disturbances	Congenital lymphedema	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative
IGGAC07	NO	ID	Atypical behaviours, Attention disorder and Hyperactivity	NO	NO	Findings of the amniocentesis
IGGAC08	NO	Borderline	Language disorder, Attention disorder and Hyperactivity	NO	Growth defect	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC09	NO	ID	NO	Febrile convulsions and Epilepsy	NO	/
IGGAC10	NO	ID	NO	Epilepsy	NO	Normal blood amino acid and organic acids levels. Fragile X analysis: negative

IGGAC11	YES	DD	Language disorder and Hyperactivity	NO	Allergies and Gastrointestinal disorders	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC12	YES	ID	Language disorder	NO	NO	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC13	YES	ID	NO	Epilepsy	NO	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative
IGGAC14	NO	ID	Learning disorder and Aggressivity	Chiari I malformation and Mild skull base abnormality	NO	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative
IGGAC15*	NO	Borderline	Psychomotor delay and mild Atypical behaviours	NO	NO	Fragile X analysis: negative
IGGAC16*	NO	ID	Hyperactivity	Sleep disturbances	NO	Fragile X analysis: negative
IGGAC17	NO	ID	Attention disorder and Hyperactivity	Tip-toe walking, Sleep disturbances, Enuresis, Febrile convulsions	NO	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC18	YES	Mild ID	Oppositive behaviour and Aggressivity	Epilepsy, Sleep disturbances and Enuresis	Hyporexia	Normal blood amino acid, organic acids levels, guanidinoacetic acid. Karyotype, Fragile X analysis, *SCN1A* gene analysis: negative
IGGAC19	NO	DD	Coordination disorder, Atypical behaviours and Executive Functioning disorder	Epilepsy	NO	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC20	YES	Mild ID	Language disorder and Hyperactivity	EEG abnormalities	NO	Normal blood amino acid and organic acids levels. Fragile X analysis: negative
IGGAC21	NO	Borderline	Verbal Dyspraxia and Social Pragmatic Communication disorder	NO	Periventricular leukomalacia	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative
IGGAC023	YES	ID	NO	Enuresis	NO	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative
IGGAC024	NO	ID	Language disorder and Hyperactivity	NO	Growth defect, Brain midline defects	Normal blood amino acid and organic acids levels. Karyotype, Fragile X analysis: negative

Table S2. Two-hit model in cases with known syndromic/recurrent CNVs

PATIENT ID	Gender	Known syndromic/ recurrent CNVs	Inh.	Additional CNV*	Inh.	Range / known and candidate genes in the additional CNV

IGGAC02	**M**	1q21.1 Deletion (Syndromic)	*de novo*	16p12.1 Deletion (Syndromic)	pat	456 Kb
IGGAC03	**F**	15q13.3 Deletion (Syndromic)	ND	Xq28 Deletion	ND	441 Kb / *RAB39B* (SFARI score 3)
IGGAC05	**F**	15q13.3 Deletion (Syndromic)	*de novo*	16p11.2 Duplication (Syndromic)	*de novo*	545 Kb
IGGAC09	**F**	15q13.3 Deletion (Syndromic)	pat	Xp11.4 Duplication	*de novo*	347 Kb / *BCOR* (Microphthalmia, syndromic 2 OMIM300166 XLD)
IGGAC11	**M**	22q11.2 Duplication (Recurrent)	mat	2p16.3 Deletion	pat	127 Kb / *NRXN1* (SFARI score 1) (Pitt-Hopkins-like syndrome 2 OMIM614325 AR)
IGGAC12	**F**	*CTNNA3* (Recurrent Deletion)	mat	1p36.33 Duplication	*de novo*	223 Kb / *GNB1* (Mental retardation, autosomal dominant 42, OMIM616973 AD);*NADK* candidate NDD genes
IGGAC20	**F**	*FHIT* (Recurrent Deletion)	pat	7q11.22 Deletion	mat	202 Kb / *KIAA0442* (SFARI score 1) (Mental retardation, autosomal dominant 26 OMIM615834 AD)
IGGAC19	**F**	*NRXN1* (Recurrent Deletion)	pat	8q24.21 Duplication	*de novo*	286 Kb/ *FAM49B* (candidate NDD gene)
IGGAC21	**M**	*NRXN1* (Recurrent Deletion)	mat	3p12.1 Deletion	pat	19 Kb / *CADM2* (SFARI score 3)
IGGAC17	**M**	*RBFOX1* (Recurrent Deletion)	pat	16q22.1-q22.2 Deletion	*de novo*	1153 Kp / *HYDIN* (SFARI score 3)
IGGAC18	**M**	*RBFOX1* (Recurrent Deletion)	mat	14q23.3 Deletion	pat	160 kp / *GPHN* (SFARI score 2) (Molybdenum cofactor deficiency C OMIM615501 AR)

*Additional CNV: additional CNVs were represented by known syndromic/recurrent CNVs or CNVs overlapping a known NDD gene;

#for known, large syndromic CNVs overlapped known and candidate NDD genes are not shown

Table S3. Complex NDD cases

PATIENT ID	Gender	CNV coordinates (hg19)	Minimal region (Kp)	del/dup	Inh	Total genes (RefSeqAll)	Candidate genes
IGGAC04	M	5:88232587-90181774	1,949	del	*de novo*	9	*MEF2C-AS1* and *ADGVR*
IGGAC01	F	16:8969984-10078802	1,109	dup	*de novo*	6	*USP7* and *GRIN2A*
IGGAC24	M	19:55948706-56471996	523	dup	*de novo*	24	*SHISA7*
IGGAC23	M	10:125121038-127119447	1,998	dup	*de novo*	13	*CTBP2*

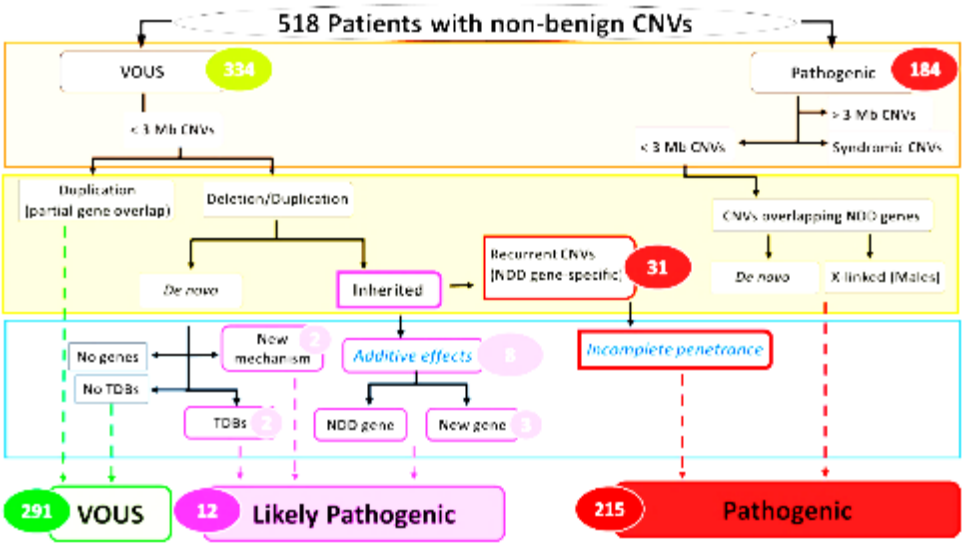
Table S4. Biological processes and KEGG pathways associated with genes implicated in the present cases and genes reported in the SFARI database and/or in OMIM as associated to neurodevelopmental disorders.

Annotation	Term	Genes found	Input size	Term genes	Genes universe	pval	pval_adj	Genes
GO:0016358	dendrite development	6	887	37	61686	0,00001381	0,00018191	GRIP1,**SYNGAP1**,MECP2,MAP1A,PPP1R9B,ACTL6B
GO:0007186	G protein-coupled receptor signaling pathway	27	887	682	61686	0,00000311	0,00005263	PLCB1,ARHGEF9,**ADCY5**,ADCY3,LADCI,NDPLHS,NEDMHM,TRIO,NS9,RORB,AVPR1A,PLXNB1,ACRDYS2,OXTR,ANXA1,SCA44,MRD42,**GNAZ**,NEDIM,

								GNAI1,OR52M1,**GPR123**,PREX1,COCPMR,MDCCAID,CMCS,CNR1
GO:0006811	ion transport	68	887	594	61686	0,00000000	0,00000000	MRXS32,MRXSRC,LKPAT,HMNDYT2,CHRNA7,SLC9A6,SCA42ND,CACNA1H,CACNB2,CACNA1E,CACNA1D,CACNA1C,CACNA1A,**GLRA3**,TRPC6,SLC1A2,SCN9A,SCN8A,FFEVF4,SCN2A,SCN1A,IECEE3,ATP2B2,AHC2,P2RX5,TRPM1,KCNS3,KCNQ3,KCNQ2,KCNJ10,KCNJ2,EPM7,KCNB1,MCIDDS,EIEE32,EIEE46,GRIN2B,GRIN2A,GRIN1,GRIK5,GRIK2,SCAR18,GRID1,NEDSGA,GRIA2,GRIA1,HCN1,GABRG3,FEB8,GABRB3,GABRB2,EIEE79,GABRA3,EIEE78,EIEE19,SLC38A10,**NIPA1**,COXPD18,MRX108,SSASKS,**NIPA2**,CDG2N,DA3,ENFL5,SLC12A5,MRD46,CACNA2D3,FHEIG
GO:0006629	lipid metabolic process	26	887	585	61686	0,00000057	0,00001162	DDHD2,CHKB,PLCB1,CMT2Z,SPG18,MEND,CBAS6,FANCS,MCT1D,PTPN11,PTEN,PON1,PCCA,OCRL,LRP1,HDLBP,HSD10MD,GBA,MRD55,DEDSM,MBOAT7,EIEE38,**ABHD4**,DHCR7,CYP27A1,PLDECO
GO:0000226	microtubule cytoskeleton organization	11	887	136	61686	0,00000488	0,00007692	CLASP1,CFEOM3A,CUL7,PEAMO,PEBAT,SON,MAP1A,PARD3B,**TUBGCP5**,TAOK1,DISC1
GO:0007018	microtubule-based movement	6	887	74	61686	0,00070194	0,00470892	SMALED2B,KIF14,**DNAH17**,MRD9,NEDMIBA,DYNC1H1
GO:0008380	mRNA splicing	18	887	264	61686	0,00000007	0,00000188	**SYNCRIP**,DDX48,SRSF11,CDK13,SON,PPP2CA,MRXS34,HNRNPU,AUKS,AFF2,FMR1,**BUD13**,PPP1R9B,MRX12,CELF4,**RBFOX1**,ALAZS,RSRC1
GO:1901215	negative regulation of neuron death	5	887	63	61686	0,00214674	0,01103270	PPP5C,PRKN,**INPP5A**,GBA,GABRB2
GO:0031175	neuron projection development	17	887	129	61686	0,00000000	0,00000000	**CNTNAP2**,**CYFIP1**,NLGN1,NRXN1,WASF1,MDFPMR,CHN3,PTEN,OPHN1,MECP2,LAMB1,NEDIM,GBA,CNTN4,PPP1R9B,FIME,TBC1D23
GO:0006357	regulation of transcription by RNA polymerase II	103	887	1111	61686	0,00000000	0,00000000	ZBTB20,NIPBL,ADNP,MED13L,EIEE67,CANPMR,ZC3H4,CIC,PHF8,SIHIWES,CHD3,CHD2,PILBOS,KDM4C,MYT1L,ZNF292,SRCAP,KDM5B,RAI1,TBR1,CTCF,DEAF1,MED13,NFE2L3,LDB1,CCNK,MRX106,CUL3,VEZF1,GDACCF,NEDISHM,CUGS,YY1,SWCOS,BMFS5,BBSOAS,TCF20,TCF7L2,TCF4,FANCS,TBX1,SOX5,CSS5,SMARCC2,SMARCA4,SMARCA2,**SHOX**,SATB1,RORB,RORA,RFX3,RAD21,PRKCB,POU3F3,PHF2,PAX6,PAX5,NR4A2,NFIX,MACID,IMDDHH,BRMUTD,EIEE72,NR3C2,MEIS2,MECP2,SMAD4,LMX1B,HOXA1,AUKS,HIVEP2,MKX,BRWD3,EBF3,ARX,FOXG1,NEDDFL,NACC1,FOXP2,EP300,HIES3,**NKX62**,TBL1XR1,IMD49,IRF2BPL,DLX3,FND2,HIVEP3,TSHZ3,ARID1B,ZMIZ1,SOX6,MRT58,PHIP,CUX1,CC2D1A,BCL11A,KDM3B,ACTL6B,BAZ2B,MRT69,FOXP1,ELP4
GO:0099560	synaptic membrane adhesion	6	887	23	61686	0,00000071	0,00001414	NLGN1,SPARCL1,**PTPRD**,GPC4,NTNG2,**LRRC4C**
hsa04514	KEGG: Cell adhesion molecules	12	887	63	61686	0,00000000	0,00000000	**CNTNAP2**,NEDCPMD,NLGN1,NRXN2,NRXN1,NRXN3,CHN3,NTNG2,**LRRC4C**,NLGN2,NLGN4X,NLGN3

Term genes found, number of annotated genes in the reference list; Input size, total number of genes in the input list, including the genes implicated in the present cases, SFARI reported genes and genes reported in OMIM as associated to NDD; Term genes, number of annotated genes in the reference list; Genes universe, total number of human reference genes; pval, hypergeometric test pValue; pval_adj, corrected hypergeometric pValue using FDR procedure.

Figure S1. Schematic representation of pipeline used to analyse CNVs in the present work



Here we represent the main steps of our analysis.

Figure S2. Additive effects between *SHOX* and *DNAH17* in patient IGGAC14

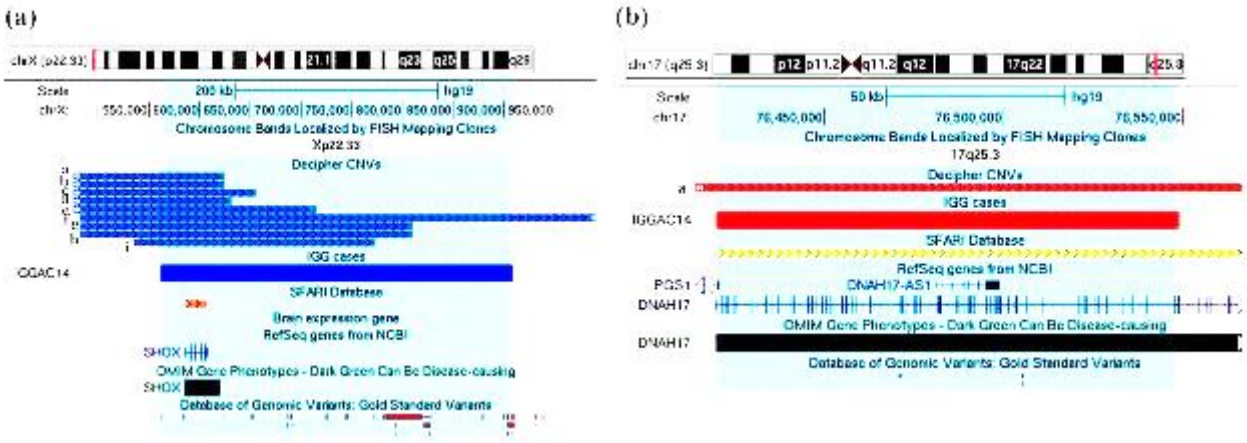


Figure S3. Additive effects in a familial case, siblings IGGAC15 and IGGAC16

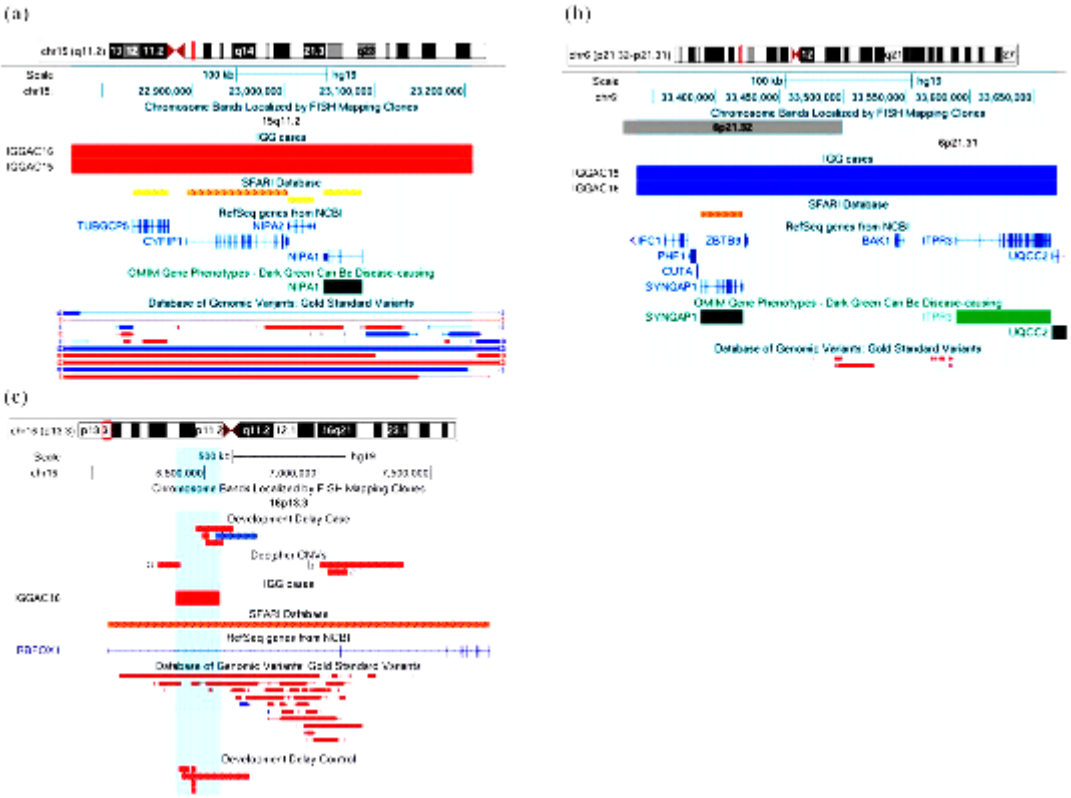


Figure S4. Additive effects among *INPP5A*, *MIR4465* and *GLRA3* in patient IGGAC08

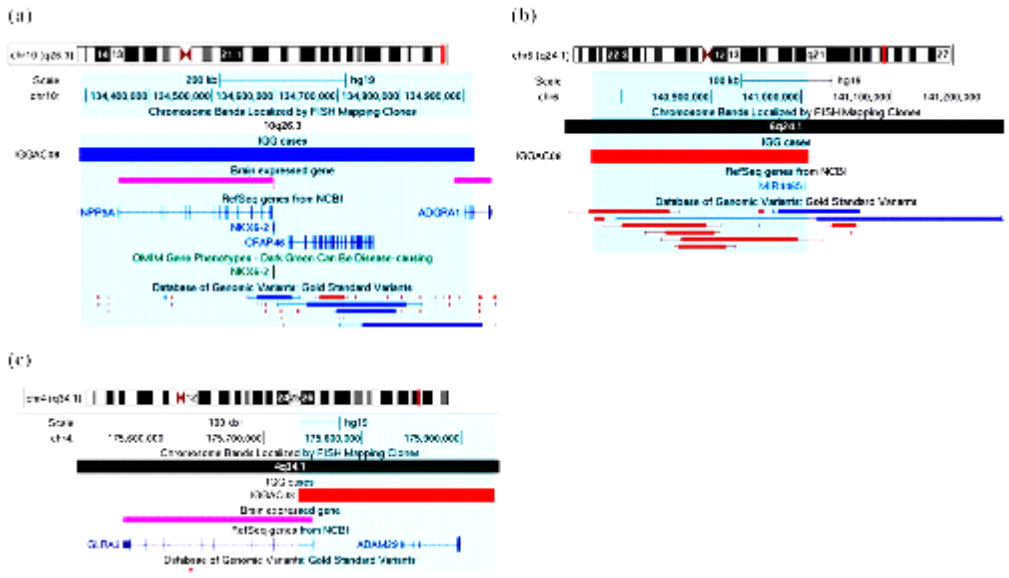


Figure S5. Additive effects between *SYNCRIP* and *ADCY5* in patient IGGAC10

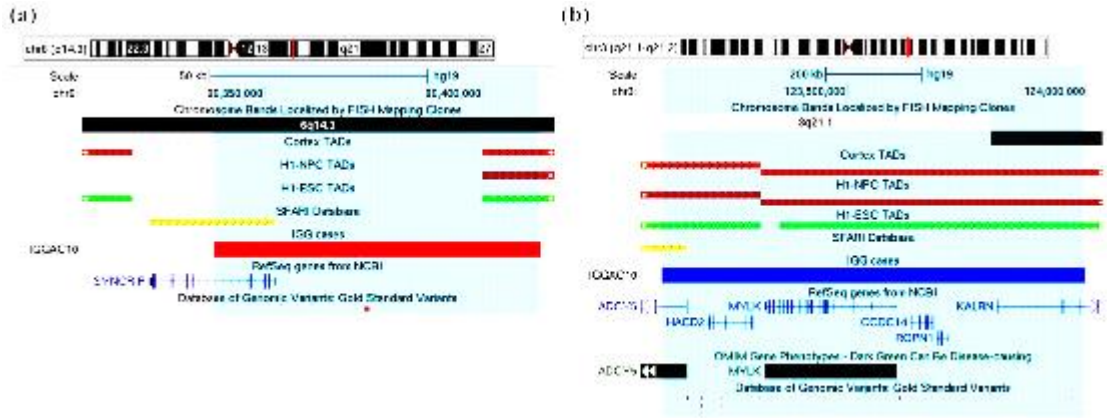


Figure S6. Additive effects among *ABHD4*, *GNAZ* and *WSCD2* in patient IGGAC07

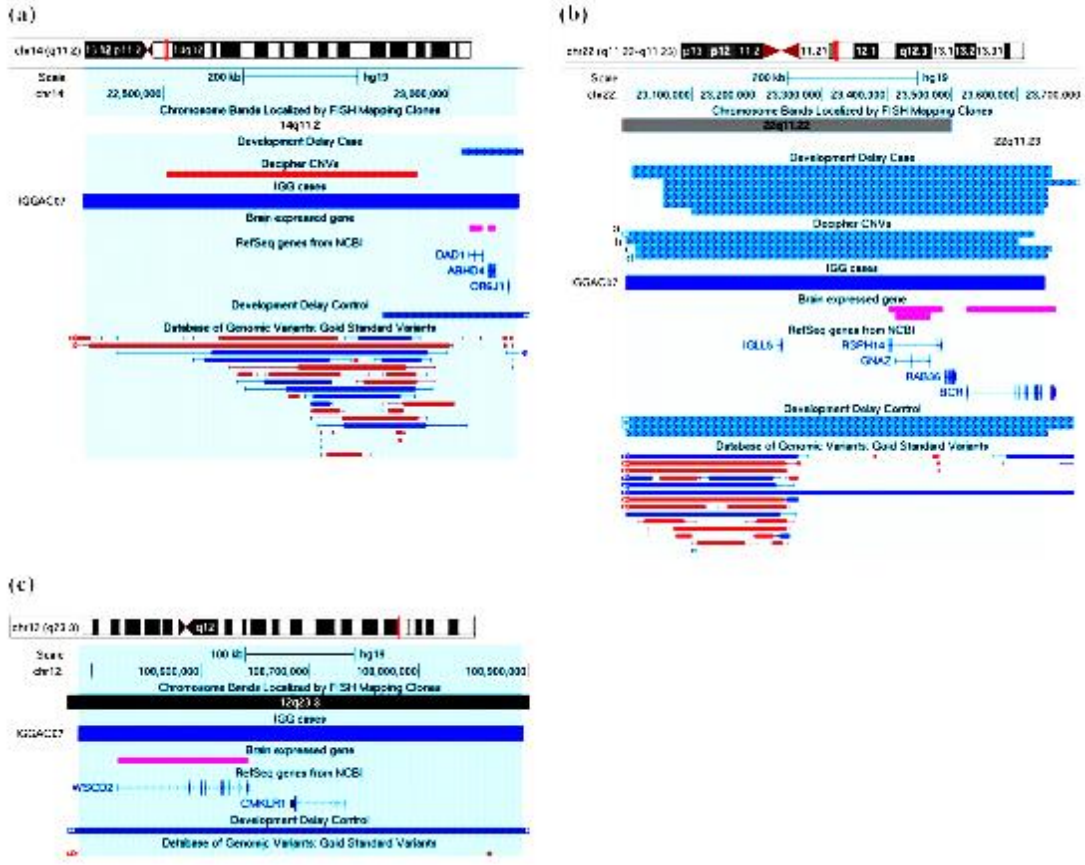
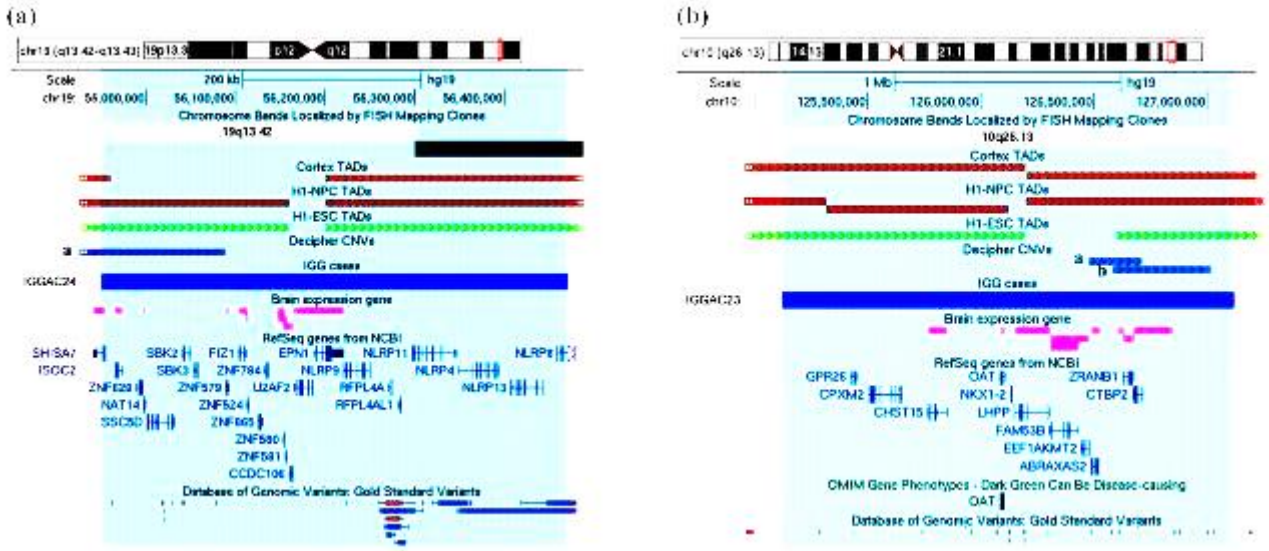


Figure S7. Duplications of topological domain boundaries (TDBs) potentially implicated in the phenotype of two patients (IGGAC24 and IGGAC23)



5.9. REFERENCES FOR THIS CHAPTER

Bossi, S., Musante, I., Bonfiglio, T., Bonifacino, T., Emionite, L., Cerminara, M., . . . Puliti, A. (2018). Genetic inactivation of mGlu5 receptor improves motor coordination in the Grm1(crv4) mouse model of SCAR13 ataxia. *Neurobiol Dis, 109*(Pt A), 44-53. doi:10.1016/j.nbd.2017.10.001

Cerminara, M., Spirito, G., Pisciotta, L., Squillario, M., Servetti, M., Divizia, M. T., . . . Puliti, A. (2021). Case Report: Whole Exome Sequencing Revealed Disease-Causing Variants in Two Genes in a Patient With Autism Spectrum Disorder, Intellectual Disability, Hyperactivity, Sleep and Gastrointestinal Disturbances. *Front Genet, 12*, 625564. doi:10.3389/fgene.2021.625564

Chen, J., Zhuang, Y., Zhang, Z. F., Wang, S., Jin, P., He, C., . . . Wan, Q. (2016). Glycine confers neuroprotection through microRNA-301a/PTEN signaling. *Mol Brain, 9*(1), 59. doi:10.1186/s13041-016-0241-3

Cheng, Y., Wang, Z. M., Tan, W., Wang, X., Li, Y., Bai, B., . . . Jin, P. (2018). Partial loss of psychiatric risk gene Mir137 in mice causes repetitive behavior and impairs sociability and learning via increased Pde10a. *Nat Neurosci, 21*(12), 1689-1703. doi:10.1038/s41593-018- 0261-7

Choii, G., & Ko, J. (2015). Gephyrin: a central GABAergic synapse organizer. *Exp Mol Med, 47*, e158. doi:10.1038/emm.2015.5

Choucair, N., Mignon-Ravix, C., Cacciagli, P., Abou Ghoch, J., Fawaz, A., Megarbane, A., . . . Chouery, E. (2015). Evidence that homozygous PTPRD gene microdeletion causes trigonocephaly, hearing loss, and intellectual disability. *Mol Cytogenet, 8*, 39. doi:10.1186/s13039-015-0149-0

Coe, B. P., Stessman, H. A. F., Sulovari, A., Geisheker, M. R., Bakken, T. E., Lake, A. M., . . . Eichler, E. E. (2019). Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nat Genet, 51*(1), 106-116. doi:10.1038/s41588-018-0288-4

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., . . . Ren, B. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature, 518*(7539), 331-336. doi:10.1038/nature14222

Fernandez, J. P., Moreno-Mateos, M. A., Gohr, A., Miao, L., Chan, S. H., Irimia, M., & Giraldez, A.J. (2018). RES complex is associated with intron definition and required for zebrafish early embryogenesis. *PLoS Genet, 14*(7), e1007473. doi:10.1371/journal.pgen.1007473

Fountain, M. D., Oleson, D. S., Rech, M. E., Segebrecht, L., Hunter, J. V., McCarthy, J. M., . . . Schaaf, C. P. (2019). Pathogenic variants in USP7 cause a neurodevelopmental disorder with speech delays, altered behavior, and neurologic anomalies. *Genet Med, 21*(8), 1797-1807. doi:10.1038/s41436-019-0433-1

Guo, H., Duyzend, M. H., Coe, B. P., Baker, C., Hoekzema, K., Gerdts, J., . . . Eichler, E. E. (2019). Genome sequencing identifies multiple deleterious variants in autism patients with more severe phenotypes. *Genet Med, 21*(7), 1611-1620. doi:10.1038/s41436-018-0380-2

Han, W., Li, J., Pelkey, K. A., Pandey, S., Chen, X., Wang, Y. X., . . . Lu, W. (2019). Shisa7 is a GABAA receptor auxiliary subunit controlling benzodiazepine actions. *Science, 366*(6462), 246-250. doi:10.1126/science.aax5719

Hultman, R., Kumari, U., Michel, N., & Casey, P. J. (2014). Galphaz regulates BDNF-induction of axon growth in cortical neurons. *Mol Cell Neurosci, 58*, 53-61. doi:10.1016/j.mcn.2013.12.004

Kim, S., Kang, M., Park, D., Lee, A. R., Betz, H., Ko, J., . . . Um, J. W. (2021). Impaired formation of high-order gephyrin oligomers underlies gephyrin dysfunction-associated pathologies. *iScience, 24*(2), 102037. doi:10.1016/j.isci.2021.102037

Laszlo, Z. I., Lele, Z., Zoldi, M., Miczan, V., Mogor, F., Simon, G. M., . . . Katona, I. (2020). ABHD4- dependent developmental anoikis safeguards the embryonic brain. *Nat Commun, 11*(1), 4363. doi:10.1038/s41467-020-18175-4

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., . . . Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-291. doi:10.1038/nature19057

Lemke, J. R., Lal, D., Reinthaler, E. M., Steiner, I., Nothnagel, M., Alber, M., . . . von Spiczak, S. (2013). Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet, 45*(9), 1067-1072. doi:10.1038/ng.2728

Mattheisen, M., Samuels, J. F., Wang, Y., Greenberg, B. D., Fyer, A. J., McCracken, J. T., . . . Nestadt, G. (2015). Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Mol Psychiatry, 20*(3), 337-344. doi:10.1038/mp.2014.43

Maussion, G., Cruceanu, C., Rosenfeld, J. A., Bell, S. C., Jollant, F., Szatkiewicz, J., . . . Ernst, C. (2017). Implication of LRRC4C and DPP6 in neurodevelopmental disorders. *Am J Med Genet A, 173*(2), 395-406. doi:10.1002/ajmg.a.38021

McMahon, A. C., Barnett, M. W., O'Leary, T. S., Stoney, P. N., Collins, M. O., Papadia, S., . . . Kind, P. C. (2012). SynGAP isoforms exert opposing effects on synaptic strength. *Nat Commun, 3*,
900. doi:10.1038/ncomms1900

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., . . . Strom, T. M. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet, 380*(9854), 1674-1682. doi:10.1016/S0140-6736(12)61480-9

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., . . . Martin, C. L. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med, 22*(2), 245- 257. doi:10.1038/s41436-019-0686-8

Ross, P. J., Zhang, W. B., Mok, R. S. F., Zaslavsky, K., Deneault, E., D'Abate, L., . . . Ellis, J. (2020). Synaptic Dysfunction in Human Neurons With Autism-Associated Deletions in PTCHD1-
AS. *Biol Psychiatry, 87*(2), 139-149. doi:10.1016/j.biopsych.2019.07.014

Rossi, P. I., Musante, I., Summa, M., Pittaluga, A., Emionite, L., Ikehata, M., . . . Puliti, A. (2013). Compensatory molecular and functional mechanisms in nervous system of the Grm1(crv4) mouse lacking the mGlu1

receptor: a model for motor coordination deficits. *Cereb Cortex, 23*(9), 2179-2189. doi:10.1093/cercor/bhs200

Rumbaugh, G., Adams, J. P., Kim, J. H., & Huganir, R. L. (2006). SynGAP regulates synaptic strength and mitogen-activated protein kinases in cultured neurons. *Proc Natl Acad Sci U S A, 103*(12), 4344-4351. doi:10.1073/pnas.0600084103

Sanders, S. J., Ercan-Sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., . . . State, M. W. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron, 70*(5), 863-885. doi:10.1016/j.neuron.2011.05.002

Schmitt, A. D., Hu, M., & Ren, B. (2016). Genome-wide mapping and analysis of chromosome architecture. *Nat Rev Mol Cell Biol, 17*(12), 743-755. doi:10.1038/nrm.2016.104

Schmitz, L. J. M., Klaassen, R. V., Ruiperez-Alonso, M., Zamri, A. E., Stroeder, J., Rao-Ruiz, P., . . . Spijker, S. (2017). The AMPA receptor-associated protein Shisa7 regulates hippocampal synaptic function and contextual memory. *Elife, 6*. doi:10.7554/eLife.24192

Schormair, B., Kemlink, D., Roeske, D., Eckstein, G., Xiong, L., Lichtner, P., . . . Winkelmann, J. (2008). PTPRD (protein tyrosine phosphatase receptor type delta) is associated with restless legs syndrome. *Nat Genet, 40*(8), 946-948. doi:10.1038/ng.190

Shihab, H. A., Rogers, M. F., Campbell, C., & Gaunt, T. R. (2017). HIPred: an integrative approach to predicting haploinsufficient genes. *Bioinformatics, 33*(12), 1751-1757. doi:10.1093/bioinformatics/btx028

Tabas-Madrid, D., Nogales-Cadenas, R., & Pascual-Montano, A. (2012). GeneCodis3: a nonredundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res, 40*(Web Server issue), W478-483. doi:10.1093/nar/gks402

Tao, Z., Feng, C., Mao, C., Ren, J., Tai, Y., Guo, H., . . . Wang, M. (2019). MiR-4465 directly targets PTEN to inhibit AKT/mTOR pathway-mediated autophagy. *Cell Stress Chaperones, 24*(1), 105-113. doi:10.1007/s12192-018-0946-6

Tassano, E., Gimelli, S., Divizia, M. T., Lerone, M., Vaccari, C., Puliti, A., & Gimelli, G. (2015). Thrombocytopenia-absent radius (TAR) syndrome due to compound inheritance for a 1q21.1 microdeletion

and a low-frequency noncoding RBM8A SNP: a new familial case. *Mol Cytogenet, 8*, 87. doi:10.1186/s13039-015-0188-6

Tomita, H., Cornejo, F., Aranda-Pino, B., Woodard, C. L., Rioseco, C. C., Neel, B. G., . . . Cancino, G. I. (2020). The Protein Tyrosine Phosphatase Receptor Delta Regulates Developmental Neurogenesis. *Cell Rep, 30*(1), 215-228 e215. doi:10.1016/j.celrep.2019.11.033

Tomoda, T., Kim, J. H., Zhan, C., & Hatten, M. E. (2004). Role of Unc51.1 and its binding partners in CNS axon outgrowth. *Genes Dev, 18*(5), 541-558. doi:10.1101/gad.1151204

Uetani, N., Kato, K., Ogura, H., Mizuno, K., Kawano, K., Mikoshiba, K., . . . Iwakura, Y. (2000). Impaired learning with enhanced hippocampal long-term potentiation in PTPdelta-deficient mice. *EMBO J, 19*(12), 2775-2785. doi:10.1093/emboj/19.12.2775

Uhl, G. R., & Martinez, M. J. (2019). PTPRD: neurobiology, genetics, and initial pharmacology of a pleiotropic contributor to brain phenotypes. *Ann N Y Acad Sci, 1451*(1), 112-129. doi:10.1111/nyas.14002

Vaccari, C. M., Romanini, M. V., Musante, I., Tassano, E., Gimelli, S., Divizia, M. T., . . . Puliti, A. (2014). De novo deletion of chromosome 11q12.3 in monozygotic twins affected by Poland Syndrome. *BMC Med Genet, 15*, 63. doi:10.1186/1471-2350-15-63

Vijiaratnam, N., Bhatia, K. P., Lang, A. E., Raskind, W. H., & Espay, A. J. (2019). ADCY5-Related Dyskinesia: Improving Clinical Detection of an Evolving Disorder. *Mov Disord Clin Pract, 6*(7), 512-520. doi:10.1002/mdc3.12816

Wu, X., Li, W., & Zheng, Y. (2020). Recent Progress on Relevant microRNAs in Autism Spectrum Disorders. *Int J Mol Sci, 21*(16). doi:10.3390/ijms21165904

Yamagata, A., Yoshida, T., Sato, Y., Goto-Ito, S., Uemura, T., Maeda, A., . . . Fukai, S. (2015). Mechanisms of splicing-dependent trans-synaptic adhesion by PTPdelta-IL1RAPL1/IL- 1RAcP for synaptic differentiation. *Nat Commun, 6*, 6926. doi:10.1038/ncomms7926

Yan, H., Shi, Z., Wu, Y., Xiao, J., Gu, Q., Yang, Y., . . . Wang, J. (2019). Targeted next generation sequencing in 112 Chinese patients with intellectual disability/developmental delay: novel mutations and candidate gene. *BMC Med Genet, 20*(1), 80. doi:10.1186/s12881-019-0794-y

Yang, A. W., Sachs, A. J., & Nystuen, A. M. (2015). Deletion of Inpp5a causes ataxia and cerebellar degeneration in mice. *Neurogenetics, 16*(4), 277-285. doi:10.1007/s10048-015-0450-4

Yang, J. H., Chang, M. W., Pandey, P. R., Tsitsipatis, D., Yang, X., Martindale, J. L., . . . Gorospe, M. (2020). Interaction of OIP5-AS1 with MEF2C mRNA promotes myogenic gene expression. *Nucleic Acids Res, 48*(22), 12943-12956. doi:10.1093/nar/gkaa1151

Yoshida, T., Shiroshima, T., Lee, S. J., Yasumura, M., Uemura, T., Chen, X., . . . Mishina, M. (2012). Interleukin-1 receptor accessory protein organizes neuronal synaptogenesis as a cell adhesion molecule. *J Neurosci, 32*(8), 2588-2600. doi:10.1523/JNEUROSCI.4637-11.2012

6. WES STUDY ON AUTISM SPECTRUM DISORDER

6.1. INTRODUCTION

Next generation sequencing (NGS) technologies enabled the extensive study of the genomics underlying human diseases. Namely whole exome sequencing (WES) represents a cost-efficient method which can lead to the detection of genomic variants and the discovery of novel disease-associated genes. This is particularly important in the case of common complex diseases such as autism spectrum disorders (ASD), whose genetic etiology is still poorly understood. Multiple studies support an oligogenic model of inheritance with a combination of inherited and / or de novo variants. The recent literature shows that the study of the entire exome (together with the analysis of the CNVs) can lead to a diagnostic evaluation in approximately 30% of the patients[61].

6.2. MATERIALS AND METHODS

A series of 68 children (60 trios) with autism spectrum disorder will be recruited at the Neuropsychiatry Unit and at the Medical Genetic Unit. The project consists of three phases:

1. Enrollment: this phase will be active until the last six months. This will allow us to involve the maximum number of cases, however having the security to complete the analysis in time for the closure of the work.
2. Analysis of WES and WGS: Sending of samples to the IIT Center for sequencing will be performed every month.
3. Data validation: it is the phase that requires more time, therefore it will be activated from the moment in which the first results will be obtained until the closure of the project.

The aim of the project is the identification of genetic / genomic mutations associated with ASD in this group of selected patients, using Whole Exome Sequencing (WES) and Whole Genome Sequencing (WGS) analysis technologies.

Written informed consent was obtained from the parents of the enrolled children. Genetic counselling and tests were performed considering the patient's medical history and phenotype.

We reviewed available data concerning clinical and neuropsychiatric status, presence of comorbidity, pharmacological treatments, associated systemic disorders, the severity of the developmental delay or intellectual disability considering physiological milestones and cognitive test.

6.3. RESULTS

The study is still ongoing, so these are preliminary data.

The samples of 34/60 trios have already been examined, of these, the results of the first 28 are being interpreted based on the clinical data I provided. I will discuss the results obtained at the moment on the 34 patients evaluated.

Clinical and neuropsychiatric characteristics

We evaluated the genetic results of 34 ASD patients, 5 females and 29 males, age ranged between 4 and 43 years (mean age 11,84 years).

We found in 8/34 a family of type “multiplex”, which have at least two individuals receiving validated ASDs diagnoses who are first to third degree relatives (for third degree, only consider cousins), the remaining part is “simplex” (they have only one known individual with ASD in first to third degree relatives) or unknown (Table 1). Twelve patients (35%) presented a psychomotor regression in their developmental history. Only two subjects presented a normal speech development and only 6/34 (18%) had a fluent language, while 20/34 (59%) had an unstructured or absent language. Despite 11/34 (32%) patients presented a delayed motor development, only 8/34 (23%) of our cohort presented a coordination impairment or abnormal gait (3 patients).

Table 1 Developmental History

ID	Age	SEX	Family type	Development regression/ arrest	Speech Development	Language	Motor development	Motor ability	Sphinter control
eSA201	8	F	1	/	delay	SW	delay	normal	yes
eSA202	5	M	1	x	delay	NV	normal	normal	no

eSA213	7	M	1	x	delay	SW	delay	normal	yes
eSA204	8	F	1	/	delay	SP	normal	normal	yes
eSA205	9	M	1	/	delay	SW	normal	normal	no
eSA206	7	M	1	x	delay	SP	normal	coordination impairment	yes
eSA207	7	M	2	x	delay	NV	normal	normal	no
eSA208A	8	F	2	x	delay	SW	normal	normal	yes
eSA208B	18	M	2	/	delay	SP	normal	normal	yes
eSA219	5	M	1	/	delay	NV	delay	coordination impairment	no
eSA2011	10	M	2	x	delay	SP	normal	normal	yes
eSA2012	7	M	1	x	delay	SW	normal	normal	NA
eSA2013	12	M	1	/	delay	F	normal	normal	yes
eSA2015	13	M	3	/	delay	SW	normal	coordination impairment	no
eSA2016	12	M	1	/	normal	SP	delay	normal	yes
eSA2017	6	M	1	/	delay	F	normal	normal	yes
eSA2018	12	M	3	x	delay	SW	normal	normal	yes
eSA2019	19	M	3	x	delay	NV	delay	abnormal gait	no
eSA2020	4	F	1	/	delay	SW	normal	normal	yes
eSA2021	11	M	3	/	delay	SP	normal	normal	yes
eSA2022	13	M	2	/	normal	F	normal	coordination impairment	yes
eSA2023	6	M	1	/	delay	SP	normal	normal	no
eSA2024	11	M	1	/	delay	F	normal	normal	yes
eSA2025	10	M	1	/	delay	F	normal	normal	yes

eSA2026	7	M	1	/	delay	F	delay	coordination impairment	yes
eSA2027	32	M	1	/	delay	SW	delay	normal	yes
eSA2028	6	M	2	/	delay	SW	delay	normal	yes
eSA2029	12	M	2	/	delay	SW	normal	normal	yes
eSA2032	7	M	1	/	delay	SW	delay	normal	no
eSA2033	9	M	1	/	delay	SP	normal	normal	yes
eSA2034	6	M	2	/	delay	SW	normal	normal	no
eSA2035	7	M	1	x	delay	SW	normal	normal	no
eSA2036	16	M	1	x	delay	SW	delay	abnormal gait	no
eSA2070	43	F	1	x	delay	SW	delay	abnormal gait	no

Legend: M= male, F= female; 1= simplex (Simplex families have only one known individual with ASD in first to third (cousin) degree relatives), 2= multiplex (Multiplex families have at least two individuals receiving validated ASDs diagnoses who are first to third degree relatives including families with affected dizygotic twins), 3= unknown (only proband affected but only up to 2nd degree); x= presence, /= absence; NV = not verbal; F= fluent; SW= single words; SP= simple phrases.

All patients had developmental delay or intellectual disability, except for 5 subjects. The patients showed different grades of cognitive impairment, either mild (10/34), moderate (10/34), severe (3/34) or profound (6/34) (Table 2). All subjects presented an inadequate adaptive level,
in the majority much lower than that of peers. Four patients presented also epilepsy.
Feeding difficulties such as selectivity were reported in 20/34 cases (59%) without apparent association with the developmental delay degree.
We observed also an high percentage of sleep disorders (20/34), mainly co-sleeping and sleep onset delay.
Gastrointestinal disturbances with constipation was observed in 15/34 patients (44%).
Behavioral problems were frequent, especially hyperactivity (50%) and attention deficit (88%).

Table 2 Clinical features

ID	Total Dev. Q (GS)	Total IQ (WISC IV)	Not verbal IQ (Leiter)	Adaptive level (Vineland)	Feeding	Sleep	Alvo	Other medical conditions	Auto-aggressive behaviours	Hetero-aggressive behaviours	Hyper-activity	Attention disorder
eSA201	54	NA	NA	L	N	N	C	/	/	/	x	x
eSA202	54	NA	NA	L	N	CS, W	N	/	/	/	/	x
eSA213	NA	NA	67	L	N	W	N	/	/	/	x	x
eSA204	82	NA	NA	L	S	CS	C	/	x	x	x	x
eSA205	66	NA	NA	L	N	CS	D	/	/	/	/	/
eSA206	50	NA	NA	L	S	CS	N	/	/	/	/	/
eSA207	65	NA	Na	L	S	N	N	/	x	x	x	x
eSA208 A	50	NA	NA	L	S	CS, R, O, W	C, P	allergies	x	x	x	x
eSA208B	NA	80	NA	ML	S	N	N	/	/	/	/	x
eSA219	33	NA	NA	L	S	CS, R, O, W	C, M, P	broncho-spasm	/	/	x	x
eSA2011	66	NA	NA	L	S	N	C	/	/	/	/	x
eSA2012	60	NA	NA	L	S	R,O	C	/	/	x	x	x
eSA2013	74	NA	NA	ML	S	N	N	/	/	/	x	x
eSA2015	85	66	NA	ML	S	O	N	/	/	/	x	x
eSA2016	82	62	NA	L	N	O	C	/	/	/	/	x
eSA2017	55	NA	NA	L	N	N	N	/	/	/	x	x
eSA2018	36	NA	NA	L	S	CS	N	/	/	x	/	x
eSA2019	40	40	40	L	S	CS, O, W	C, M, P	Epilepsy, GER	/	/	x	x
eSA2020	66	NA	NA	L	S	CS, R, O	C	/	/	/	x	x
eSA2021	NA	45	NA	L	S	R,O	C	/	x	x	x	x
eSA2022	NA	91	NA	ML	S	N	N	/	x	/	x	x
eSA2023	41	NA	NA	L	S	N	N	FC	/	/	/	x
eSA2024	NA	78	NA	ML	N	N	N	/	/	/	/	/
eSA2025	NA	51	NA	L	N	R,O	N	epilepsy	x	x	x	x
eSA2026	76	NA	NA	L	S	O	C	/	x	x	/	x
eSA2027	50	NA	NA	L	N	N	C	/	/	/	/	/
eSA2028	56	NA	NA	L	S	N	C	/	x	x	/	x
eSA2029	50	NA	NA	L	S	CS	N	/	/	x	x	x
eSA2032	47	NA	NA	L	N	N	N	/	x	/	x	x
eSA2033	71	NA	NA	ML	S	CS, W	N	/	/	/	x	x

eSA2034	50	NA	NA	L	N	N	N	/	/	/	/	x
eSA2035	22	NA	NA	L	N	O	N	/	/	/	/	x
eSA2036	NA	40	NA	L	N	N	C	epilepsy	/	/	/	x
eSA2070	NA	40	NA	L	N	W	C	epilepsy	/	/	/	x

Legend: Dev.Q= developmental quotient, GS= Griffiths scales, IQ= intellectual quotient; NA= not available, N= normal, L= low, ML= moderately low; S= selectivity, C= constipation, D= diarrhea, M= meteorism, P= pain, CS= cosleeping, R= bedtime resistance, O= sleep onset delay, W= night wakings; x= presence, /= absence

Genetic results

Eleven cases showed possible deleterious rare variants, in different and, in the majority of cases, in multiple genes (Table 3).

In two patients, who are old cases without diagnosis, we found a *de novo* heterozygous pathogenetic variants.

In nine patients (8 M, 1 F) we can hypothesize a role of some identified variants with the presence of modifier genes or we found till now a possible oligogenic mechanism.

Table 3 Summary

Patients	**N°**
Possible deleterious rare variants	12 patients (10 M, 2 F)
In Progress	16 patients (24 M)
Without significative results	6 patients (3 M, 3 F)
Tot	34 patients (29 M, 5 F)

Syndromic variants

In two patients we found pathogenetic variants in *SHANK3* and *SYNGAP1* respectively.

The first one is an 18-year-old boy with a clinical picture of psychomotor regression type Disintegrative Disorder (eSA2019 in Table 1 and 2), while the second is a 44-year-old woman with a Rett-like phenotype (eSA2070 in Table 1 and 2).

SH3 or the gene for multiple repeated ankyrin domains 3 (*SHANK3*), is located in the distal part of the long arm of chromosome 22 and its product acts as a bridge protein for the interaction of several synaptic

molecules including the N-methyl- D-aspartate receptors (NMDA-R), the glutamate metabotropic receptors (mGluRs) I class and the a-amino-3-hydroxy-5methyl-4-isoxazole (AMPA-R) GluR1 receptors. Loss of a functional pair (haploinsufficiency) of *SHANK3* results in a syndrome called Phelan-McDermid Syndrome (PMS - OMIM ID 606232)[62]. Phelan-McDermid syndrome is caused primarily by chromosomal microdeletion of 22q13 and manifests as neonatal hypotonia, severe speech delay or absence, minor facial dysmorphism, and moderate to severe intellectual disability, sometimes associated with autism spectrum disorder.

The synaptic Ras-GTPase-activating protein 1 gene (*SYNGAP1*) encodes for a brain-specific protein that is largely localized to dendritic spines in neocortical pyramidal neurons. This protein plays a critical role in the development of intellectual abilities, acting as a repressor of neural excitability[63].

The protein-truncating de novo variants in *SYNGAP1* caused non-syndromic mental retardation, named *SYNGAP1*-Related Intellectual Disability (SYNGAP1-ID). These mutations resulted in the production of proteins lacking important domains for the synaptic plasticity and spine morphogenesis. Furthermore, the resulting premature stop codons could destabilize the SYNGAP1 messenger RNA (mRNA) transcript through the non-sense-mediated mRNA decay mechanism.

X Chromosome variants

In six males we found possibly deleterious variants located in X chromosome (Table 4).

ID	Gene	Chr	Position	Ref	Alt	Func.refGene	ExonicFunc.refGene	OMIM	SFARI
eSA213	***BCOR***	chrX	39933592	G	A	exonic	nonsynonymous p S 336 L	Microphthalmia, syndromic 2 300166 XLD	score 0
eSA219	***HUWE1***	chrX	53563401	G	A	exonic	nonsynonymous p.R4122H	Mental retardation, Turner type 309590	score S
eSA2015	***ZC4H2***	chrX	64140024	C	A	exonic	nonsynonymous p R89L	Wieacker Wolff syndrome 314580 XLR	score 0
eSA2018	***BCORL1***	chrX	129159340	G	C	exonic	nonsynonymous p R 1355 P	Shukla Vernon syndrome 301029 XLR	score 1
eSA2021	***EBP***	chrX	48382186	C	G	exonic	nonsynonymous p H 9 Q	MEND syndrome 300960 XLR	score S
eSA2027	***KDM5C***	chrX	53225941	G	A	exonic	Stopgain p Gln 903 X	Mental retardation, X linked, syndromic, Claes Jensen	score 2

								type 300534 XLR	

As we can see in the Table 4, these genes are associated with specific OMIM disorders, not always involving autism. The SFARI score is, in fact, variable.

Neurotransmitter pathway genes

In some patients we found variants in genes involving neurotransmitter pathways. In particular genes of receptors or genes that regulate neurotransmitter release or synaptic activity (Table 5).
The most involved neurotransmitters were currently found to be serotonin, glutamate and gamma-aminobutyric acid.

Table 5

ID	Gene	Chr	Position	ExonicFunc.refGene	AAChange.refGene	Heredity	Gene_descr	Implication in neurotransmitter pathways
eSA213	*HTR1E*	chr6	87725456	nonsynonymous SNV	HTR1E:NM_000865:exon2:c.C404T:p.T135M	pt	5-hydroxytryptamine receptor 1E	serotonin receptor
eSA205	*SHANK1*	chr19	51165541	nonsynonymous SNV	SHANK1:NM_016148:exon23:c.C6167T:p.P2056L	mt	SH3 And Multiple Ankyrin Repeat Domains 1	regulation of excitatory glutamatergic synapses
eSA205	*SHANK2*	chr11	70319533	nonsynonymous SNV	SHANK2:NM_133266:exon11:c.T3227C:p.I1076T	pt	SH3 And Multiple Ankyrin Repeat Domains 2	regulation of excitatory glutamatergic synapses
eSA205	*SHANK3*	chr22	51153371	nonsynonymous SNV	SHANK3:NM_033517:exon19:c.G2119A:p.A707T	mt	SH3 And Multiple Ankyrin Repeat Domains 3	regulation of excitatory glutamatergic synapses

eSA205	*GRIA1*	chr5	152871837	nonsynonymous SNV	GRIA1:NM_001258021:exon1:c.T79C:p.F27L;GRIA1:NM_001258022:exon1:c.T79C:p.F27L	pt	Glutamate Ionotropic Receptor AMPA Type Subunit 1	glutamate receptor
eSA205	*GRID2*	chr4	93511396	nonsynonymous SNV	GRID2:NM_001286838:exon2:c.C203T:p.T68M;GRID2:NM_001510:exon2:c.C203T:p.T68M	mt	Glutamate Ionotropic Receptor Delta Type Subunit 2	glutamate receptor
eSA205	*GRIN2A*	chr16	9858173	nonsynonymous SNV	GRIN2A:NM_001134407:exon13:c.C3228A:p.N1076K;GRIN2A:NM_001134408:exon13:c.C3228A:p.N1076K;GRIN2A:NM_000833:exon14:c.C3228A:p.N1076K	pt	Glutamate Ionotropic Receptor NMDA Type Subunit 2A	glutamate receptor
eSA219	*TPH2*	chr12	72366465	stop variant	NM_173353:exon6:c.775C>T:p.Q259X	de novo	Tryptophan hydroxylase 2	serotonin synthesis
eSA219	*ALDH5A1*	chr6	24505099	stopgain	NM_170740:exon4:c.612G>A:p.Trp204X	mt	Succinic semialdehyde dehydrogenase	GABA catabolism
eSA219	*ALDH5A1*	chr6	24503590	nonsynonymous SNV	NM_170740:exon3:c.538C>T:p.His180Tyr	pt	Succinic semialdehyde dehydrogenase	GABA catabolism
eSA219	*ITPR3*	chr6	33651870	nonsynonymous SNV	NM_002224:exon36:c.4862T>C:p.Leu1621Pro	pt	Inositol 1,4,5-trisphosphate receptor	realease of neurotransmitters
eSA2015	*CPLX3*	chr15	75120439	stopgain	CPLX3:NM_001030005:exon2:c.C229T:p.R77X	pt	complexin 3	realease of neurotransmitters

eSA2015	*ALDH5A1*	chr6	24533761	stopgain	ALDH5A1:NM_001080:exon10:c.C1429T:p.Q477X;ALDH5A1:NM_170740:exon11:c.C1468T:p.Q490X	mt	Succinic semialdehyde dehydrogenase	GABA catabolism
eSA208B	*CACNA1B*	chr9	140972591	nonsynonymous SNV	CACNA1B:NM_000718:exon35:c.G4975A:p.E1659K;CACNA1B:NM_001243812:exon35:c.G4975A:p.E1659K	de novo	Calcium Voltage-Gated Channel Subunit Alpha1 B	realease of neurotransmitters

Legend: mt= maternal, pt= paternal

Possible significative variants and interactions between genes

In case eSA205 we detected, in addition to the 6 variants involved in the neurotransmitter pathways (Table 5), three other interesting variants: a frameshift deletion on *SYNGAP1* (NM_006772:exon19:c.3963_3964del:p.P1321fs), a nonsynonymous on *CSMD1* (NM_033225:exon58:c.G8935A:p.G2979S*)* and another on *PHB* (NM_001281496:exon3:c.G128T:p.R43L; NM_001281497:exon3:c.G128T:p.R43L;NM_002634:exon3:c.G128T:p.R43L;NM_001281715:exon4:c.G128T:p.R43L).

About the first gene I have already discussed above (see Syndromic variants) and this variant can have a possible significance.

CSMD1, CUB and Sushi multiple domains 1, is a gene with SFARI score 3, that is predominantly expressed in the brain by neurons, and is enriched at synapses where it regulates complement pathway activation.

The Gly residue at position 2979 is very highly conserved (conservation = 0.7 from 135 aligned protein seqs). So, it is very likely to be important for the protein's function. Potentially damaging heterozygous missense variants in the CSMD1 gene were identified in ASD.

PHB is a mitochondrial scaffolding protein, and its role as a key regulator of neuronal survival is emerging.

PHB shows a global expression pattern throughout cortical and subcortical regions. This gene is highly expressed in the thalamus, which receives the input from the basal ganglia and its frontal cortical circuitry, brain regions implicated in motor activity and repetitive behaviours.

I show you 3 patients who share some clinical and genetic aspects as emerged from the (eSA213, eSA219 and eSA2015).

Patient eSA219 was diagnosed with ASD and showed cognitive and motor impairments, stereotypies, hyperactivity, sleep and gastrointestinal disturbances (Figure 1).

After the bioinformatic analysis of WES data, an unreported maternally inherited missense variant affecting the E3 ubiquitin ligase (HECT) domain of *HUWE1* (NM_031407:exon79: c.12365G>A:p.R4122H), and a *de novo* unreported stop variant in *TPH2* (NM_173353:exon6: c.775C>T:p.Q259X) (see Table 4 and 5) were identified. The analysis unveiled the presence of additional variants in the following genes, *ALDH5A1* and *ITPR3* (see Table 5). This case was described in a published case report[64].

Figure 1 Clinical hystory of patient eSA219

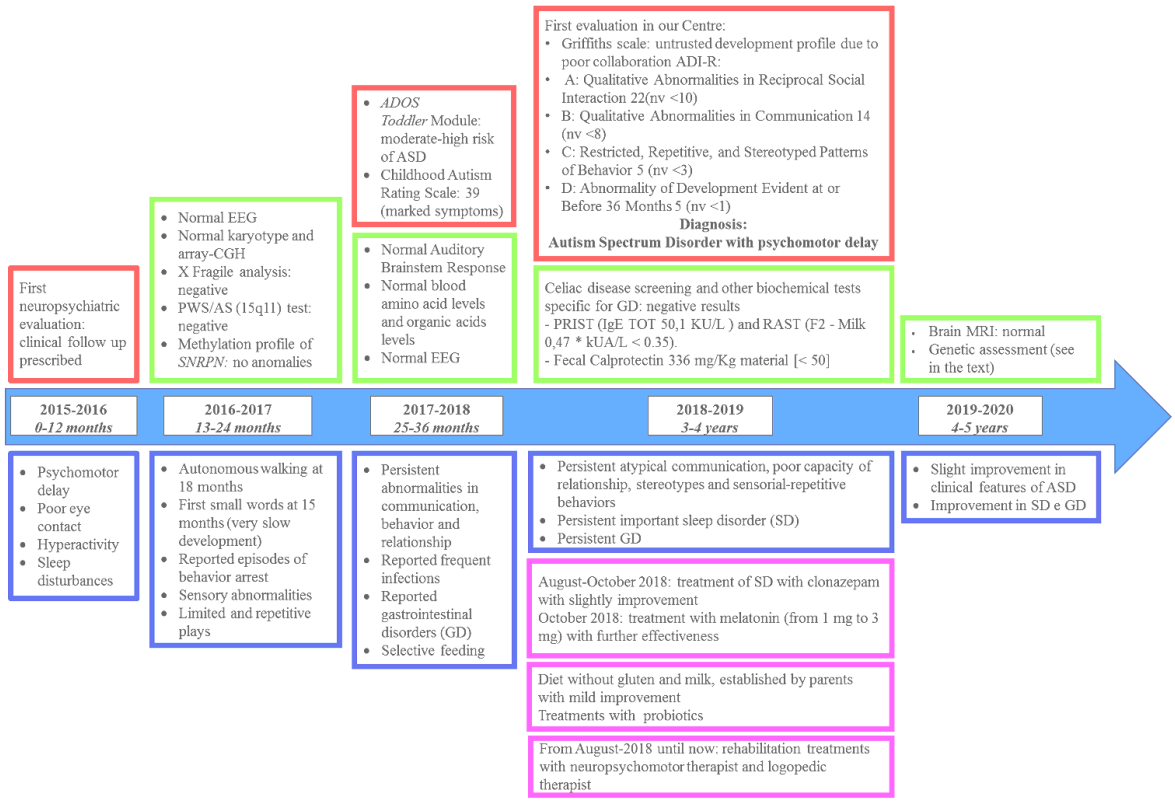


Figure legend:

Signs and symptoms

Clinical evaluations

Laboratory tests

Treatments

PWS/AS= Prader Willi Syndrome/Angelman Syndrome

GD= gastro-intestinal disorders

SD= sleep disorders

ADOS= Autism Diagnostic Observation Schedule

ADI-R= Autism Diagnostic Interview-Revised

ASD= Autism Spectrum Disorder

PRIST= Paper Radio Immuno Sorbent Test

RAST= Radio Allergo Sorbent Test

MRI= Magnetic Resonance Imaging

nv= normal value

Patient eSA213 was diagnosed with ASD, at three years old, by the help of Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R). He had developmental delay, language impairment, hyperactivity, attentional lability, stereotypes.

More in details, the clinical assessment of the patient through specific tests showed a moderate cognitive impairment with low adaptive level for chronological age, prevailing on communication, personal and social autonomies, and socialization domains. During the evaluation we found anxiety and attentional disturbances. The parents reported mainly hyperactivity and attention deficit.

By using specific tests (Sleep-CGI-S, Sleep-CGI-I) the patient was diagnosed with a mild sleep disorder. Sleep-CGI-S showed a mild sleep disorder with mild to moderate night waking. Using the Criteria of Rome IV, no gastrointestinal disorders were diagnosed.

He was carrying 2 maternally inherited variants, one hemizygous in *BCOR* (see Table 4) and one heterozygous in *MYO9B*, genes associated to cognitive and behaviour impairment. A third heterozygous paternally inherited variant affected *HTR1E*, a serotonin receptor hypothesized to play a role in autism-like behaviour and sleep disturbance (see Table 5).

In the third case (eSA2015), another male patient, the clinical assessment through specific tests showed a borderline cognitive level when he was 8-year-old.

At the last evaluation with WISC IV, the intelligence quotient is placed below the average (TIQ 66), with discrepancy between the various Indices: score in the norm in Perceptual Reasoning (PRI 85) and in Verbal Comprehension (VCI 84), while low scores for the working memory (WMI 58) and for the processing speed (PSI 62). The profile appears disharmonious but interpretable. Also, within the Indices there is a difference between the scores at the various subtests as often happens for autism spectrum disorders. In anamnesis he presented an episode of myositis and frequent bone fractures. ABC-movement 2 test did not reveal a noticeable movement disorder, but parents report difficulties in movement and coordination that affect self-esteem and social interaction.

By using specific tests (Sleep-CGI-S, Sleep-CGI-I, SDSC) the patient was diagnosed with a moderate sleep disorder. Sleep-CGI-S showed a bedtime resistance with moderate sleep onset delay. At SDSC the results confirmed disturbance of the wake / sleep transition, restless sleep with sweating and impaired breathing. He was treated with melatonin.

He had a functional gastrointestinal disorder: 2 criteria for constipation and 2 criteria for diarrhoea according to Criteria of Rome IV.

He had 2 maternally inherited variants, one hemizygous in *ZC4H2* (see Table 4), associated to developmental delay, and one heterozygous in *ALDH5A1*, associated to behaviour and sleep impairment. A further paternally inherited variant affected *CPLX3*, involved in neurotransmitter release, hypothesized to be implicated in neurodevelopmental delay (see Table 5).

Other two of our patients are brother (eSA208B) and sister (eSA208A). As we can see in Table 1 and 2, the sister showed a worse clinical picture than the brother. She had important communication and behaviour problems with multiple episodes of clinical regression. The brother showed an improving phenotype in the evolutive age. Both shared variants in genes that may be involved in the pathogenesis of their disorder.

They had in common a heterozygous variant on *DHCR7*, gene that caused Smith Lemli Opitz Syndrome in homozygous conditions. In literature there are studies about mice and human subjects having autistic traits in heterozygosity[65]. They also shared a variant on *GORASP2 that* encodes a member of the Golgi reassembly stacking protein that plays a role in Golgi ribbon formation. A de novo missense mutation in *GORASP2* has been reported in some ASD patients, and SNPs within this gene have been associated with cognitive abilities in GWAS studies[66].

In the male subjects we found different variants, such as *CACNA1B* (Table 5), in the female we we hypothesize that the worst phenotype is due to a variant in *TSC2*. This is a gene of Tuberous Sclerosis Complex, involving on cellular proliferation via mTOR regulation axon guidance.

6.4. DISCUSSION

ASD is a heterogeneous and complex condition characterized by problems on communication and social interaction, presence of repetitive behaviours, and restricted interests. It is associated with a range of comorbid conditions and there is a high grade of overlap with the other Neurodevelopmental disorders.

In our current series there is a clear prevalence of male cases, in accordance with the frequency of the disorder, but also of variants affecting the sex X chromosome.

The X chromosome is a sub-centric chromosome in which there are almost 1098 genes.

It is characterized by homology regions with the Y chromosome, those are located at telomere level and defined pseudo-autosomal regions (PAR1 and PAR2) where homologous recombination events occur.

Recent studies suggest that functional reasons may have influenced the gene content of this chromosome, with some classes of genes overrepresented and others underrepresented: the genes expressed in the muscle and brain are particularly abundant compared to other classes[67].

During evolution, the X chromosome has accumulated many genes involved in cognitive functions. In fact, about a third of known X-linked genes are involved in mental retardation[68]. In a minority of cases the Fragile X syndrome is associated with autism, while the Rett syndrome has some phenotypic characteristics common to disorders of the autism spectrum. Furthermore, mutations in two genes encoding neuroligines, *NLGN3* in Xq13 and *NLGN4* in Xp22.3 has been associated in some cases of ASD[69].

In the patients we analysed, the presence of genes important for neurotransmitter circuits and their functioning also emerged.

Numerous studies have investigated the involvement of various neurotransmitter systems in ASD, sometimes with insignificant results. The only exception is the serotonergic system. Some of these studies have indeed confirmed that a certain proportion of autistic individuals (about 1/3) have increased levels of serotonin (5-HT) in their blood and/or urine. This seems likely due to increased platelet storage rather than synthesis abnormalities[70]. This could reflect a low availability of serotonin in the CNS, as evidenced by PET studies, and justify the efficacy of treatment with transporter or selective serotonin re-uptake inhibitors.

In literature we found that depleting tryptophan (an essential precursor for 5-HT biosynthesis) in adult subjects worsens behaviour and affective features.

Serotonin is known to affect sleep, mood, arousal, aggression, impulsivity, and affiliation, all of which are important to autism spectrum disorders. In fact, sleep disturbances have been observed in a large part of autistic individuals. Genetic data obtained from other neuropsychiatric conditions, some sharing behavioural features with ASD (mood, anxiety, obsessive-compulsive disorders), suggest involvement of multiple genes in the serotonergic pathway. Numerous studies have also indicated a possible role for melatonin, an epiphyseal neuro-hormone, the precursor of which is serotonin. Some studies have shown the high prevalence of sleep disturbances and altered circadian rhythms in autistic individuals, suggesting abnormalities in the physiology of melatonin[71]. Furthermore, it seems that oral administration of this hormone to autistic sufferers can improve sleep-wake rhythms. Further analysis showed a substantial reduction in nocturnal melatonin levels produced in autistic individuals compared to controls of equal age and sex[72].

In the patients I report here, as can be seen in Table 2, Figure 1 and in the description of the cases, a sleep disturbance with effective use of melatonin was found too[64].

Regard to studies on neurotransmitters in the ASD, elevated levels of circulating GABA and its essential precursor glutamate have been observed in children with this disorder[73 74 75].

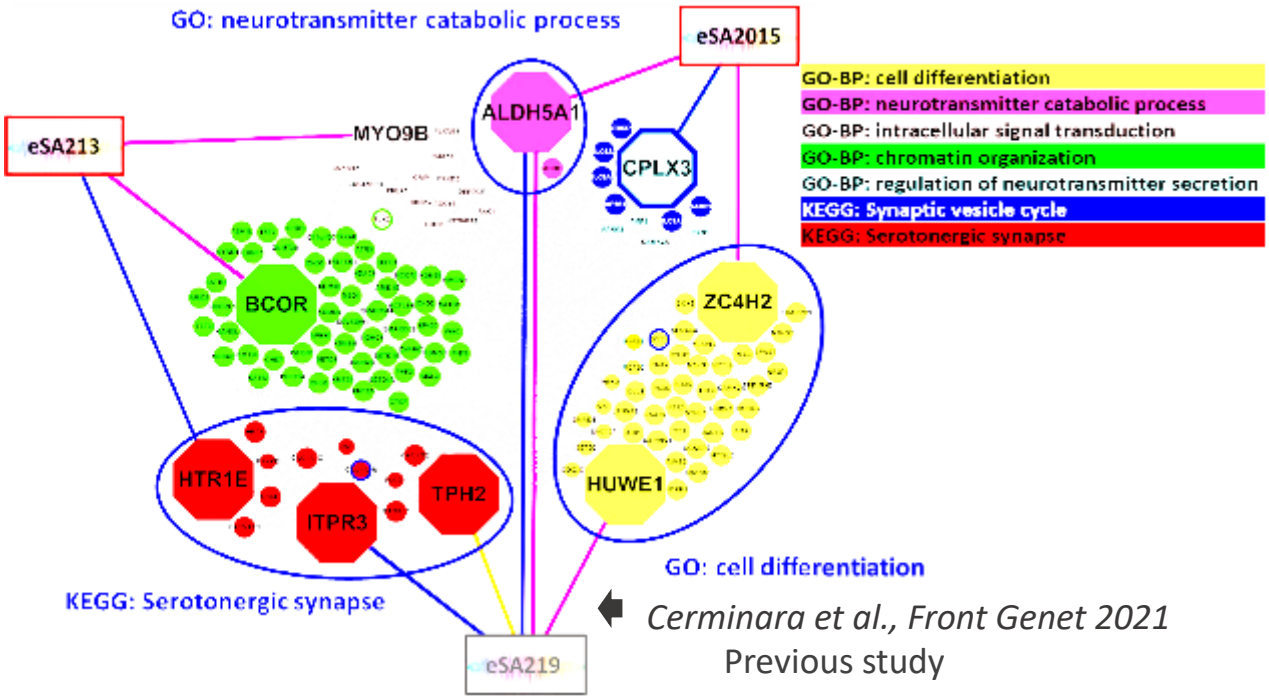
In patient eSA205 we detected 6 uncommon heterozygotic variants all embedded inside genes strongly associated to ASD. While the impact of each of these variants is likely very mild, we believe that an increased burden of variants affecting genes essential for the physiology of the glutamatergic synapse signalling pathway could be detrimental to the point of being causative for the observed ASD phenotype. In the same patient, after that, we found variant in other genes possible responsible, such as *CSMD1*.

Our findings in genes present on the X chromosome and involved in neurotransmitter pathways seem to us to be potential candidates in the development of the observed phenotypes.

Implicated genes in our patients revealed enrichment in ASD-associated biological processes and pathways. As our previously described patient[64], the other subjects presented complex ASD phenotype with other NDDs, sleep and systemic disorders.

As we can see in Figure 2 all variants, found in eSA213, eSA219 and eSA2015, collectively could have a role in the biological processes involved in the pathogenesis of ASD thus contributing to the different features of the patient's phenotype.

Figure 2 Network analysis



The network depicts the protein-protein interaction (gray lines) and was constructed using enriched BP terms clustering the candidate genes (octagon shape) and the SFARI genes (circle shape). The network has a clear distinction into 6 clusters representing the top GO-BPs. Protein-protein interactions were revealed by STRING analysis.

6.5. CONCLUSION

About 10% of patients from our ASD cohort showed rare deleterious variants in multiple genes that seem to fully explain their complex phenotype.

The analysis and the study we performed in order to prioritize the detected variants according to their putative relevance to the proband phenotype allowed us to pinpoint both clinical cases where the root of the phenotype is a single high penetrance variant and cases where an increased burden of low penetrance inherited variants could concur to cause the phenotype.

Only 5% of our ASD patients presented a syndromic picture. These were subjects with a more severe, encephalopathic phenotype.

In the majority of subjects, the contribution of X-linked genes and others involved in neurotransmitter pathways appeared important in defining the phenotype.

By analysing an extensive number of autism spectrum disorders cases, we believe that this workflow could help us to reconstruct the genomics at the basis of complex ASD phenotypes.

7. CONCLUSIONS AND FUTURE WORKS

In this chapter the main outcomes of this research work will be summarized, and I will outline future research directions. This work has given an overview of the methods of genetic study of patients affected by NDDs and ASD, already included in the diagnostic routine and / or used as new research approaches.

Also in the study of CGH arrays, usually used for a first diagnostic approach in these disorders, I tried to highlight how a careful re-analysis of stand-by data in previous years, thanks to the most recent publications and a study that brings together the clinical data and genetic knowledge of the results, with the help of new bioinformatics tools can be innovative and help to better clarify a patient's phenotype.

Thanks to this type of work, the first of the objectives I had set for my PhD project has been achieved.

This allowed to identify, as we can see in chapter 4 and 5, a large number of variants previously classified as VOUS and which, after our re-analysis, assume the role of pathogenetic CNVs involving genes recently associated with ID / ASDs.

Together with these, we have been able to identify potentially pathogenic CNVs involving genes not yet associated with NDDs but which, due to their function or belonging to certain pathways, could be significant.

The study of the interactions between the genes involved also allowed us to better understand situations of incomplete penetrance, variable expressivity, as well as to hypothesize the reason for a variable spectrum of severity in a given phenotype. In fact, we have been able to theorize in some complex patients the existence, in the context of oligogenic mechanisms, of double hit models that could cause or modify the clinical picture.

More and more often, in the daily clinic work, we realize how important it is to evaluate the clinical differences even within the same type of disorder. There is a symptoms spectrum, which DSM 5 has partly taken into consideration, which makes each subject an individual in itself, also for the therapeutic response, and which seems to me important to try to genetically explain. About Autism Spectrum Disorders, Kanner, in his 1969 article on differential diagnosis, highlighted the tendency to pigeonhole patients into a category rather than really understand them—"that children had not read the right books" when it came to diagnosis[76].

With this in mind, it was important to evaluate each patient with ASD before enrolling him for the WES analysis and to analyse the found variants, enhancing the clinical features, also with a reverse phenotyping approach, doing any other clinical investigations and diagnostic tests.

The results we have obtained lead us to advance hypotheses on the involvement of some neurotransmitters, such as serotonin, in these disorders and would also explain the associated sleep or gastrointestinal disorders in some of them.

If our theories are valid, these patient-calibrated results could lead to better use of drugs and ad hoc therapeutic strategies.

A limit of our project was the smallness of the sample, by collecting more patients who share the same genes and / or gene pathways, it will be easier to complete the study with the functional part, both in vitro and in the animal.

These results provide preliminary findings for further study and illustrate the potential usefulness and desire to continue NDDs projects in a strong collaboration between neuropsychiatrists and geneticists.

As genotyping throughput increases and costs decrease, NGS studies of large autism samples will soon offer the potential to detect new genes and mechanisms that can explain the different phenotypes.

8. PUBLICATIONS

- Alternating Hemiplegia of Childhood: Genotype-Phenotype Correlations in a Cohort of 39 Italian Patients. Cordani R, Stagnaro M, **Pisciotta L**, Tiziano FD, Calevo MG, Nobili L; I.B.AHC Consortium, De Grandis E. Front Neurol. 2021 Apr 8;12:658451. doi: 10.3389/fneur.2021.658451.
- CASK related disorder: Epilepsy and developmental outcome. Giacomini T, Nuovo S, Zanni G, Mancardi MM, Cusmai R, Pepi C, Bertini E, Valente EM, Battini R, Ferrari A, Romaniello R, Zucca C, Borgatti R, Uccella S, Severino M, Striano P, Pistorio A, Prato G, De Grandis E, Nobili L, **Pisciotta L**. Eur J Paediatr Neurol. 2021 Mar;31:61-69. doi:10.1016/j.ejpn.2021.02.006.
- Photoparoxysmal response in ADCK3 autosomal recessive ataxia: a case report and literature review. Uccella S, **Pisciotta L**, Severino M, Bertini E, Giacomini T, Zanni G, Prato G, De Grandis E, Nobili L, Mancardi MM. Epileptic Disord. 2021 Feb 1;23(1):153-160. doi: 10.1684/epd.2021.1243.
- Case Report: Whole Exome Sequencing Revealed Disease-Causing Variants in Two Genes in a Patient with Autism Spectrum Disorder, Intellectual Disability, Hyperactivity, Sleep and Gastrointestinal Disturbances. Cerminara M, Spirito G, **Pisciotta L**, Squillario M, Servetti M, Divizia MT, Lerone M, Berloco B, Boeri S, Nobili L, Vozzi D, Sanges R, Gustincich S, Puliti A. Front Genet. 2021 Feb 18;12:625564. doi: 10.3389/fgene.2021.625564.
- Cardiac phenotype in ATP1A3-related syndromes: A multicenter cohort study. Balestrini S, Mikati MA, Álvarez-García-Rovés R, Carboni M, Hunanyan AS, Kherallah B, McLean M, Prange L, De Grandis E, Gagliardi A, **Pisciotta L**, Stagnaro M, Veneselli E, Campistol J, Fons C, Pias-Peleteiro L, Brashear A, Miller C, Samões R, Brankovic V, Padiath QS, Potic A, Pilch J, Vezyroglou A, Bye AME, Davis AM, Ryan MM, Semsarian C, Hollingsworth G, Scheffer IE, Granata T, Nardocci N, Ragona F, Arzimanoglou A, Panagiotakaki E, Carrilho I, Zucca C, Novy J, Dzieżyc K, Parowicz M, Mazurkiewicz-Bełdzińska M, Weckhuysen S, Pons R, Groppa S, Sinden DS, Pitt GS, Tinker A, Ashworth M, Michalak Z, Thom M, Cross JH, Vavassori R, Kaski JP, Sisodiya SM. Neurology. 2020 Nov 24;95(21):e2866-e2879. doi: 10.1212/WNL.0000000000010794.
- Targeted re-sequencing in malformations of cortical development: genotype-phenotype correlations. Accogli A, Severino M, Riva A, Madia F, Balagura G, Iacomino M, Carlini B, Baldassari S, Giacomini T, Croci C, **Pisciotta L**, Messana T, Boni A, Russo A, Bilo L, Tonziello R, Coppola A, Filla A, Mecarelli O, Casalone R, Pisani F, Falsaperla R, Marino S, Parisi P, Ferretti A, Elia M, Luchetti A,

Milani D, Vanadia F, Silvestri L, Rebessi E, Parente E, Vatti G, Mancardi MM, Nobili L, Capra V, Salpietro V, Striano P, Zara F. Seizure. 2020 Aug;80:145-152. doi: 10.1016/j.seizure.2020.05.023.

- Optic Atrophy and Generalized Chorea in a Patient Harboring an OPA10/RTN4IP1 Pathogenic Variant. Giacomini T, Gamucci A, **Pisciotta L**, Nesti C, Fiorillo C, Doccini S, Morana G, Nobili L, Santorelli FM, Mancardi MM, De Grandis E. Neuropediatrics. 2020 Dec;51(6):425-429. doi: 10.1055/s-0040-1708539.
- RARS1-related leukodystrophy: When hypomyelination meets brain atrophy Accogli A, **Pisciotta L**, Figà-Talamanca L, Di Rocco M, Severino M. Med Res Innov. 2020 July. doi: 10.15761/MRI.1000178
- Schimke immuno-osseous dysplasia, two new cases with peculiar EEG pattern. Prato G, De Grandis E, Mancardi MM, Cordani R, Giacomini T, **Pisciotta L**, Uccella S, Severino M, Tortora D, Pavanello M, Bertamino M, Verrina E, Caridi G, Di Rocco M, Nobili L. Brain Dev. 2020 May;42(5):408-413. doi: 10.1016/j.braindev.2020.01.008.
- White matter and cerebellar involvement in alternating hemiplegia of childhood. Severino M*, **Pisciotta L***, Tortora D, Toselli B, Stagnaro M, Cordani R, Morana G, Zicca A, Kotzeva S, Zanaboni C, Montobbio G, Rossi A, De Grandis E; IBAHC Consortium. J Neurol. 2020 May;267(5):1300-1311. doi: 10.1007/s00415-020-09698-3.

* These authors equally contributed to this work.

- Age and sex prevalence estimate of Joubert syndrome in Italy. Nuovo S, Bacigalupo I, Ginevrino M, Battini R, Bertini E, Borgatti R, Casella A, Micalizzi A, Nardella M, Romaniello R, Serpieri V, Zanni G, Valente EM, Vanacore N; **JS Italian Study Group***. Neurology. 2020 Feb 25;94(8):e797-e801. doi: 10.1212/WNL.0000000000008996.

*I am part of JS Italian Study Group.

- Epileptic Encephalopathy, Myoclonus-Dystonia, and Premature Pubarche in Siblings with a Novel C-Terminal Truncating Mutation in ATRX Gene. Giacomini T, Vari MS, Janis S, Prato G, **Pisciotta L**, Rocchi A, Michelucci A, Di Rocco M, Gandullia P, Mattioli G, Sacco O, Morana G, Mancardi MM. Neuropediatrics 2019 Oct;50(5):327-331. doi: 10.1055/s-0039-1692141.
- Severe early-onset developmental and epileptic encephalopathy (DEE) associated with novel compound heterozygous pathogenic variants in SLC25A22: Case report and literature review. Giacomini T, **Pisciotta L**, Prato G, Meola I, Zara F, Fiorillo C, Baratto S, Severino M, De Grandis E, Mancardi MM. Seizure. 2019 Aug;70:56-58. doi: 10.1016/j.seizure.2019.06.029.
- Personality profile and health-related quality of life in adults with previous continuous spikewaves during slow sleep syndrome. Lenci G, Calevo MG, Gaggero R, Prato G, **Pisciotta L**, De Grandis E, Mancardi MM, Baglietto MG, Vigano' M, Veneselli E. Brain Dev. 2019 Jun;41(6):522-530. doi: 10.1016/j.braindev.2019.02.004.
- Intragenic duplication of KCNQ5 gene results in aberrant splicing leading to a premature termination codon in a patient with intellectual disability. Rosti G, Tassano E, Bossi S, Divizia MT, Ronchetto P, Servetti M, Lerone M, **Pisciotta L**, Mancardi MM, Veneselli E, Puliti A. Eur J Med Genet. 2019 Sep;62(9):103555. doi: 10.1016/j.ejmg.2018.10.007.
- Biallelic Mutations in ADPRHL2, Encoding ADP-Ribosylhydrolase 3, Lead to a Degenerative Pediatric Stress-Induced Epileptic Ataxia Syndrome. Ghosh SG, Becker K, Huang H, Salazar TD, Chai G, Salpietro V, Al-Gazali L, Waisfisz Q, Wang H, Vaux KK, Stanley V, Manole A, Akpulat U, Weiss MM, Efthymiou S, Hanna MG, Minetti C, Striano P, **Pisciotta L**, De Grandis E, Altmüller J, Nürnberg P, Thiele H, Yis U, Okur TD, Polat AI, Amiri N, Doosti M, Karimani EG, Toosi MB, Haddad

G, Karakaya M, Wirth B, van Hagen JM, Wolf NI, Maroofian R, Houlden H, Cirak S, Gleeson JG. Am J Hum Genet. 2018 Nov 1;103(5):826. doi: 10.1016/j.ajhg.2018.10.002.

- CNNM2 homozygous mutations cause severe refractory hypomagnesemia, epileptic encephalopathy and brain malformations. Accogli A, Scala M, Calcagno A, Napoli F, Di Iorgi N, Arrigo S, Mancardi MM, Prato G, **Pisciotta L**, Nagel M, Severino M, Capra V. Eur J Med Genet. 2019 Mar;62(3):198-203. doi: 10.1016/j.ejmg.2018.07.014.
- Epileptic Encephalopathy in Adams-Oliver Syndrome Associated to a New DOCK6 Mutation: A Peculiar Behavioral Phenotype. **Pisciotta L**, Capra V, Accogli A, Giacomini T, Prato G, Tavares P, Pinto-Basto J, Morana G, Mancardi MM. Neuropediatrics. 2018 Jun;49(3):217-221. doi: 10.1055/s-0038-1639372
- ATP1A3 spectrum disorders: A video-documented history of 7 genetically confirmed early onset cases. Stagnaro M, **Pisciotta L**, Gherzi M, Di Rocco M, Gurrieri F, Parrini E, Prato G, Veneselli E, De Grandis E. Eur J Paediatr Neurol. 2018 Mar;22(2):264-271. doi: 10.1016/j.ejpn.2018.01.010

Submitted articles:

- Alternating Hemiplegia of Childhood in a child harboring a novel TBC1D24 mutation: case report and literature review. Cordani R* , **Pisciotta L*** , Mancardi MM , Stagnaro M, Prato G, Giacomini G , Morana G, Walsh P, Ghia T, Nobili L , De Grandis E.

Manuscript ID is neuped-12-2020-2815-SC.R1

* These authors equally contributed to this work.

9. REFERENCES

[1] American Psychiatric Association., *Diagnostic and Statistical Manual of Mental Disorders: DSM-III, 3th Ed*, 1980.

[2] American Psychiatric Association., *Diagnostic and Statistical Manual of Mental Disorders: DSM-5™, 5th Ed* (Arlington, VA, US: American Psychiatric Publishing, Inc., 2013) <https://doi.org/10.1176/appi.books.9780890425596>.

[3] World Health Organization, *International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision)*, 2020.

[4] Craig J. Newschaffer, Matthew D. Falb, and James G. Gurney, 'National Autism Prevalence Trends from United States Special Education Data', *Pediatrics*, 115.3 (2005), e277-282 <https://doi.org/10.1542/peds.2004-1958>.

[5] M. Prior, 'Is There an Increase in the Prevalence of Autism Spectrum Disorders?', *Journal of Paediatrics and Child Health*, 39.2 (2003), 81–82 <https://doi.org/10.1046/j.1440-1754.2003.00097.x>.

[6] M. Rutter, 'Incidence of Autism Spectrum Disorders: Changes over Time and Their Meaning', *Acta Paediatrica (Oslo, Norway: 1992)*, 94.1 (2005), 2–15 <https://doi.org/10.1111/j.1651-2227.2005.tb01779.x>.

[7] Centers for Disease Control and Prevention, 'Morbidity and Mortality Weekly Report', . . *Increasing Prevalence of Parent-Reported Attention-Deficit/Hyperactivity Disorder among Children --- United States, 2003 and 2007*, 59.44, 1439–43.

[8] Matthew J. Maenner and others, 'Potential Impact of *DSM-5* Criteria on Autism Spectrum Disorder Prevalence Estimates', *JAMA Psychiatry*, 71.3 (2014), 292 <https://doi.org/10.1001/jamapsychiatry.2013.3893>.

[9] American Psychiatric Association., *Diagnostic and Statistical Manual of Mental Disorders (4th Ed., Text Rev.)*, 2020 <https://doi.org/10.1176/appi.books.9780890423349.>.

[10] Maenner MJ, Shaw KA, Baio J, et al., 'Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016.', 69.No. SS-4, 1–12 <http://dx.doi.org/10.15585/mmwr.ss6904a1>.

[11] Roy McConkey, 'The Rise in the Numbers of Pupils Identified by Schools with Autism Spectrum Disorder (ASD): A Comparison of the Four Countries in the United Kingdom', *Support for Learning*, 35.2 (2020), 132–43.

[12] Anita Thapar, Miriam Cooper, and Michael Rutter, 'Neurodevelopmental Disorders', *The Lancet Psychiatry*, 4.4 (2017), 339–46 <https://doi.org/10.1016/S2215-0366(16)30376-5>.

[13] Tania Das Banerjee, Frank Middleton, and Stephen V. Faraone, 'Environmental Risk Factors for Attention-Deficit Hyperactivity Disorder', *Acta Paediatrica (Oslo, Norway: 1992)*, 96.9 (2007), 1269–74 <https://doi.org/10.1111/j.1651-2227.2007.00430.x>.

[14] Adnan T. Bhutta and others, 'Cognitive and Behavioral Outcomes of School-Aged Children Who Were Born Preterm: A Meta-Analysis', *JAMA*, 288.6 (2002), 728–37 <https://doi.org/10.1001/jama.288.6.728>.

[15] Melissa Herrmann, Katherine King, and Michael Weitzman, 'Prenatal Tobacco Smoke and Postnatal Secondhand Smoke Exposure and Child Neurodevelopment', *Current Opinion in Pediatrics*, 20.2 (2008), 184–90 <https://doi.org/10.1097/MOP.0b013e3282f56165>.

[16] D. K. Daily, H. H. Ardinger, and G. E. Holmes, 'Identification and Evaluation of Mental Retardation', *American Family Physician*, 61.4 (2000), 1059–67, 1070.

[17] J. Flint and A. O. Wilkie, 'The Genetics of Mental Retardation', *British Medical Bulletin*, 52.3 (1996), 453–64 <https://doi.org/10.1093/oxfordjournals.bmb.a011559>.

[18] Patricia N. Pastor and Cynthia A. Reuben, 'Diagnosed Attention Deficit Hyperactivity Disorder and Learning Disability: United States, 2004-2006', *Vital and Health Statistics. Series 10, Data from the National Health Survey*, 237, 2008, 1–14.

[19] K. Langley and others, 'Maternal Smoking during Pregnancy as an Environmental Risk Factor for Attention Deficit Hyperactivity Disorder Behaviour. A Review', *Minerva Pediatrica*, 57.6 (2005), 359–71.

[20] Institute of Medicine (US) Committee on Understanding Premature Birth and Assuring Healthy Outcomes, *Preterm Birth: Causes, Consequences, and Prevention*, ed. by Richard E. Behrman and Adrienne Stith Butler,

The National Academies Collection: Reports Funded by National Institutes of Health (Washington (DC): National Academies Press (US), 2007) <http://www.ncbi.nlm.nih.gov/books/NBK11362/> [accessed 24 April 2021].

[21] Joel Nigg, Molly Nikolas, and S. Alexandra Burt, 'Measured Gene-by-Environment Interaction in Relation to Attention-Deficit/Hyperactivity Disorder', *Journal of the American Academy of Child and Adolescent Psychiatry*, 49.9 (2010), 863–73 <https://doi.org/10.1016/j.jaac.2010.01.025>.

[22] Beena Johnson, 'Learning Disabilities in Children: Epidemiology, Risk Factors and Importance of Early Intervention', *BMH Med J*, 2017, 31–37.

[23] Scott M. Myers, Chris Plauché Johnson, and American Academy of Pediatrics Council on Children With Disabilities, 'Management of Children with Autism Spectrum Disorders', *Pediatrics*, 120.5 (2007), 1162–82 <https://doi.org/10.1542/peds.2007-2362>.

[24] Macey L. Murray and others, 'Pharmacological Treatments Prescribed to People with Autism Spectrum Disorder (ASD) in Primary Health Care', *Psychopharmacology*, 231.6 (2014), 1011–21 <https://doi.org/10.1007/s00213-013-3140-7>.

[25] Beverly Pringle and others, 'Diagnostic History and Treatment of School-Aged Children with Autism Spectrum Disorder and Special Health Care Needs', *NCHS Data Brief*, 97, 2012, 1–8.

[26] Mark D. Shen and others, 'Early Brain Enlargement and Elevated Extra-Axial Fluid in Infants Who Develop Autism Spectrum Disorder', *Brain: A Journal of Neurology*, 136.Pt 9 (2013), 2825–35 <https://doi.org/10.1093/brain/awt166>.

[27] Ruth A. Carper and others, 'Cerebral Lobes in Autism: Early Hyperplasia and Abnormal Age Effects', *NeuroImage*, 16.4 (2002), 1038–51 <https://doi.org/10.1006/nimg.2002.1099>.

[28] Heather Cody Hazlett, Michele Poe, and others, 'Magnetic Resonance Imaging and Head Circumference Study of Brain Size in Autism: Birth through Age 2 Years', *Archives of General Psychiatry*, 62.12 (2005), 1366–76 <https://doi.org/10.1001/archpsyc.62.12.1366>.

[29] Cynthia M. Schumann and others, 'Longitudinal Magnetic Resonance Imaging Study of Cortical Development through Early Childhood in Autism', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30.12 (2010), 4419–27 <https://doi.org/10.1523/JNEUROSCI.5714-09.2010>.

[30] Heather Cody Hazlett, Michele D. Poe, and others, 'Early Brain Overgrowth in Autism Associated with an Increase in Cortical Surface Area before Age 2 Years', *Archives of General Psychiatry*, 68.5 (2011), 467–76 <https://doi.org/10.1001/archgenpsychiatry.2011.39>.

[31] Frank H. Duffy and Heidelise Als, 'A Stable Pattern of EEG Spectral Coherence Distinguishes Children with Autism from Neuro-Typical Controls - a Large Case Control Study', *BMC Medicine*, 10 (2012), 64 <https://doi.org/10.1186/1741-7015-10-64>.

[32] Shanshan Hong and others, 'Detecting Abnormalities of Corpus Callosum Connectivity in Autism Using Magnetic Resonance Imaging and Diffusion Tensor Tractography', *Psychiatry Research*, 194.3 (2011), 333–39 <https://doi.org/10.1016/j.pscychresns.2011.03.009>.

[33] Jason J. Wolff and others, 'Differences in White Matter Fiber Tract Development Present from 6 to 24 Months in Infants with Autism', *The American Journal of Psychiatry*, 169.6 (2012), 589–600 <https://doi.org/10.1176/appi.ajp.2011.11091447>.

[34] A. Di Martino and others, 'The Autism Brain Imaging Data Exchange: Towards a Large-Scale Evaluation of the Intrinsic Brain Architecture in Autism', *Molecular Psychiatry*, 19.6 (2014), 659–67 <https://doi.org/10.1038/mp.2013.78>.

[35] Colleen P. Chen and others, 'Diagnostic Classification of Intrinsic Functional Connectivity Highlights Somatosensory, Default Mode, and Visual Regions in Autism', *NeuroImage. Clinical*, 8 (2015), 238–45 <https://doi.org/10.1016/j.nicl.2015.04.002>.

[36] Angela E. Abbott and others, 'Patterns of Atypical Functional Connectivity and Behavioral Links in Autism Differ Between Default, Salience, and Executive Networks', *Cerebral Cortex (New York, N.Y.: 1991)*, 26.10 (2016), 4034–45 <https://doi.org/10.1093/cercor/bhv191>.

[37] S. Jill James and others, 'Metabolic Endophenotype and Related Genotypes Are Associated with Oxidative Stress in Children with Autism', *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, 141B.8 (2006), 947–56 <https://doi.org/10.1002/ajmg.b.30366>.

[38] Olga A. Sedelnikova and others, 'Role of Oxidatively Induced DNA Lesions in Human Pathogenesis', *Mutation Research*, 704.1–3 (2010), 152–59 <https://doi.org/10.1016/j.mrrev.2009.12.005>.

[39] Janie F. Shelton, Daniel J. Tancredi, and Irva Hertz-Picciotto, 'Independent and Dependent Contributions of Advanced Maternal and Paternal Ages to Autism Risk', *Autism Research: Official Journal of the International Society for Autism Research*, 3.1 (2010), 30–39 <https://doi.org/10.1002/aur.116>.

[40] J. F. Crow, 'The Origins, Patterns and Implications of Human Spontaneous Mutation', *Nature Reviews. Genetics*, 1.1 (2000), 40–47 <https://doi.org/10.1038/35049558>.

[41] Deborah J. Morris-Rosendahl and Marc-Antoine Crocq, 'Neurodevelopmental Disorders—the History and Future of a Diagnostic Concept', *Dialogues in Clinical Neuroscience*, 22.1 (2020), 65–72 <https://doi.org/10.31887/DCNS.2020.22.1/macrocq>.

[42] L Ghirardi and others, 'The Familial Co-Aggregation of ASD and ADHD: A Register-Based Cohort Study', *Molecular Psychiatry*, 23.2 (2018), 257–62 <https://doi.org/10.1038/mp.2017.17>.

[43] John N. Constantino and others, 'Sibling Recurrence and the Genetic Epidemiology of Autism', *American Journal of Psychiatry*, 167.11 (2010), 1349–56 <https://doi.org/10.1176/appi.ajp.2010.09101470>.

[44] Sally Ozonoff and others, 'Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study', *Pediatrics*, 128.3 (2011), e488-495 <https://doi.org/10.1542/peds.2010-2825>.

[45] Rebecca E. Rosenberg and others, 'Characteristics and Concordance of Autism Spectrum Disorders Among 277 Twin Pairs', *Archives of Pediatrics & Adolescent Medicine*, 163.10 (2009), 907 <https://doi.org/10.1001/archpediatrics.2009.98>.

[46] Hiroko Taniai and others, 'Genetic Influences on the Broad Spectrum of Autism: Study of Proband-ascertained Twins', *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147B.6 (2008), 844–49.

[47] Silvia De Rubeis and others, 'Synaptic, Transcriptional and Chromatin Genes Disrupted in Autism', *Nature*, 515.7526 (2014), 209–15 <https://doi.org/10.1038/nature13772>.

[48] Dheeraj Malhotra and Jonathan Sebat, 'CNVs: Harbingers of a Rare Variant Revolution in Psychiatric Genetics', *Cell*, 148.6 (2012), 1223–41 <https://doi.org/10.1016/j.cell.2012.02.039>.

[49] David T. Miller and others, 'Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies', *American Journal of Human Genetics*, 86.5 (2010), 749–64 <https://doi.org/10.1016/j.ajhg.2010.04.006>.

[50] Lira Mamanova and others, 'Target-Enrichment Strategies for next-Generation Sequencing', *Nature Methods*, 7.2 (2010), 111–18 <https://doi.org/10.1038/nmeth.1419>.

[51] Marcos Morey and others, 'A Glimpse into Past, Present, and Future DNA Sequencing', *Molecular Genetics and Metabolism*, 110.1–2 (2013), 3–24 <https://doi.org/10.1016/j.ymgme.2013.04.024>.

[52] Henrike O. Heyne and others, 'De Novo Variants in Neurodevelopmental Disorders with Epilepsy', *Nature Genetics*, 50.7 (2018), 1048–53 <https://doi.org/10.1038/s41588-018-0143-7>.

[53] Jeremy F. McRae and others, 'Prevalence and Architecture of de Novo Mutations in Developmental Disorders', *Nature*, 542.7642 (2017), 433–38 <https://doi.org/10.1038/nature21062>.

[54] Thomas G. Schulze and Francis J. McMahon, 'Defining the Phenotype in Human Genetic Studies: Forward Genetics and Reverse Phenotyping', *Human Heredity*, 58.3–4 (2004), 131–38 <https://doi.org/10.1159/000083539>.

[55] Holly A. Stessman, Raphael Bernier, and Evan E. Eichler, 'A Genotype-First Approach to Defining the Subtypes of a Complex Disease', *Cell*, 156.5 (2014), 872–77 <https://doi.org/10.1016/j.cell.2014.02.002>.

[56] Brian J. O'Roak and others, 'Sporadic Autism Exomes Reveal a Highly Interconnected Protein Network of de Novo Mutations', *Nature*, 485.7397 (2012), 246–50 <https://doi.org/10.1038/nature10989>.

[57] Heng Li and Richard Durbin, 'Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform', *Bioinformatics (Oxford, England)*, 26.5 (2010), 589–95 <https://doi.org/10.1093/bioinformatics/btp698>.

[58] Kai Wang, Mingyao Li, and Hakon Hakonarson, 'ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data', *Nucleic Acids Research*, 38.16 (2010), e164 <https://doi.org/10.1093/nar/gkq603>.

[59] Bradley P. Coe and others, 'Refining Analyses of Copy Number Variation Identifies Specific Genes Associated with Developmental Delay', *Nature Genetics*, 46.10 (2014), 1063–71 <https://doi.org/10.1038/ng.3092>.

[60] Giulia Rosti and others, 'Intragenic Duplication of KCNQ5 Gene Results in Aberrant Splicing Leading to a Premature Termination Codon in a Patient with Intellectual Disability', *European Journal of Medical Genetics*, 62.9 (2019), 103555 <https://doi.org/10.1016/j.ejmg.2018.10.007>.

[61] Arnold Munnich and others, 'Impact of On-Site Clinical Genetics Consultations on Diagnostic Rate in Children and Young Adults with Autism Spectrum Disorder', *Molecular Autism*, 10.1 (2019), 33 <https://doi.org/10.1186/s13229-019-0284-2>.

[62] Katy Phelan, R. Curtis Rogers, and Luigi Boccuto, 'Phelan-McDermid Syndrome', in *GeneReviews®*, ed. by Margaret P. Adam and others (Seattle (WA): University of Washington, Seattle, 1993) <http://www.ncbi.nlm.nih.gov/books/NBK1198/> [accessed 19 June 2021].

[63] James P. Clement and others, 'Pathogenic SYNGAP1 Mutations Impair Cognitive Development by Disrupting Maturation of Dendritic Spine Synapses', *Cell*, 151.4 (2012), 709–23 <https://doi.org/10.1016/j.cell.2012.08.045>.

[64] Maria Cerminara and others, 'Case Report: Whole Exome Sequencing Revealed Disease-Causing Variants in Two Genes in a Patient With Autism Spectrum Disorder, Intellectual Disability, Hyperactivity, Sleep and Gastrointestinal Disturbances', *Frontiers in Genetics*, 12 (2021), 625564 <https://doi.org/10.3389/fgene.2021.625564>.

[65] Avi Saskin and others, 'Prevalence of Four Mendelian Disorders Associated with Autism in 2392 Affected Families', *Journal of Human Genetics*, 62.6 (2017), 657–59 <https://doi.org/10.1038/jhg.2017.16>.

[66] Yasser Al-Sarraj and others, 'Family-Based Genome-Wide Association Study of Autism Spectrum Disorder in Middle Eastern Families', *Genes*, 12.5 (2021), 761 <https://doi.org/10.3390/genes12050761>.

[67] Di Kim Nguyen and Christine M. Disteche, 'Dosage Compensation of the Active X Chromosome in Mammals', *Nature Genetics*, 38.1 (2006), 47–53 <https://doi.org/10.1038/ng1705>.

[68] David H. Skuse, 'X-Linked Genes and Mental Functioning', *Human Molecular Genetics*, 14 Spec No 1 (2005), R27-32 <https://doi.org/10.1093/hmg/ddi112>.

[69] Stéphane Jamain and others, 'Mutations of the X-Linked Genes Encoding Neuroligins NLGN3 and NLGN4 Are Associated with Autism', *Nature Genetics*, 34.1 (2003), 27–29 <https://doi.org/10.1038/ng1136>.

[70] Dubravka Hranilović and others, 'Hyperserotonemia in Autism: Activity of 5HT-Associated Platelet Proteins', *Journal of Neural Transmission*, 116.4 (2009), 493–501 <https://doi.org/10.1007/s00702-009-0192-2>.

[71] Beth A. Malow, 'Sleep Disorders, Epilepsy, and Autism', *Mental Retardation and Developmental Disabilities Research Reviews*, 10.2 (2004), 122–25 <https://doi.org/10.1002/mrdd.20023>.

[72] Sylvie Tordjman and others, 'Nocturnal Excretion of 6-Sulphatoxymelatonin in Children and Adolescents with Autistic Disorder', *Biological Psychiatry*, 57.2 (2005), 134–38 <https://doi.org/10.1016/j.biopsych.2004.11.003>.

[73] Sarah Aldred and others, 'Plasma Amino Acid Levels in Children with Autism and Their Families', *Journal of Autism and Developmental Disorders*, 33.1 (2003), 93–97 <https://doi.org/10.1023/a:1022238706604>.

[74] Dirk Dhossche and others, 'Elevated Plasma Gamma-Aminobutyric Acid (GABA) Levels in Autistic Youngsters: Stimulus for a GABA Hypothesis of Autism', *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 8.8 (2002), PR1-6.

[75] H. Moreno-Fuenmayor and others, 'Plasma Excitatory Amino Acids in Autism', *Investigacion Clinica*, 37.2 (1996), 113–28.

[76] James Harris, 'Leo Kanner and Autism: A 75-Year Perspective', *International Review of Psychiatry (Abingdon, England)*, 30 (2018), 1–15 <https://doi.org/10.1080/09540261.2018.1455646>.

ACKNOWLEDGEMENTS

When I will be thinking about my PhD in the future, I will think of the years spent at the Gaslini Institute, as a newly specialist, enthusiastic about the new scientific discoveries that could help me improve the lives of many of my little patients.
I want to thank many people that have helped me professionally and personally over the last three years.
First of all, I would like to thank Prof. Edvige Veneselli for believing in me when I had to prepare for the PhD competition. Then I thank Paola Lanteri and Savina Severino for supporting me in writing the first project. The following thanks go to those who have supported me over the years, Elisa De Grandis, Margherita Mancardi, Thea Giacomini and Prof. Lino Nobili. In particular Elisa for her brilliant suggestions and continuous help, particularly in writing and reviewing papers and this thesis, and her willingness to do so at any time. Thea was always there to encourage me when I had to endure a difficult moment. For this I also thank my other friends Marcella, Ramona, Irene, Bianca, Bernadette, Carolina, Alessandra and Sara.

My deepest gratitude goes to Prof. Aldamaria Puliti for having been my reference point during my PhD. She was always helpful and gave me the opportunity of performing my research work at a high level with the collaboration of her laboratory, especially Martina and Maria who I have always found to give me a hand.
The contribution of Silvia Boeri has been precious in the last year, who I hope will continue in the ASD project.
Special thanks go to the other people in my Unit, who have become friends during these years.
Last but not least, I wish to thank Vincenzo, my mum and my dad for their help in every circumstance and for always believing in me.