University of Siena – Department of Medical Biotechnologies

Doctorate in Genetics, Oncology and Clinical Medicine (GenOMeC)

XXXV cycle (2019-2022)

Coordinator: Prof. Francesca Ariani

Exploring genetic bases of Intellectual disability and Autism: from Exome Sequencing on

Scientific disciplinary sector: Medical Genetics (MED/03)

Tutor — Prof. Alessandra Renieri

PhD Candidate — Dr. Lucia Pia Bruno

Index

Abstract.....pag. 4
1. Introduction .....pag. 5
1.1 Unlocking the biology of autism pag.6
1.2 ID/ ASD heterogeneity and complexity…..... ……pag.7
1.3 Whole Exome sequencing: positives and drawbacks……..…….....pag.9
1.4 WES as a resource for paediatric patients …….....pag.10
1.5 Gender differences.....pag.11
1.6. Site- directed mutagenesis and Neuroligin 4 X-linked…….....pag.12
2. Results……..…….……..pag.16
2.1 Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations…….……..……..……...pag16
2.2 New Candidates for Autism/Intellectual Disability Identified by Whole-Exome Sequencing……..…pag.32
2.3 *SPTBN5*, encoding the βV-spectrin protein, leads to a syndrome of intellectual disability, developmental delay, and seizures.…….……..…….……....pag.44
2.4 Natural history of KBG syndrome in a large European cohort…….……..…….pag.53
2.5 Additional collaborations and findings……pag.66
2.6 Site-directed mutagenesis on *NLGN4X*…….…….....…pag.67
3. Materials and Methods pag.74
3.1. WES…….…..pag.74
3.1.1. Selection of Patients and DNA Samples Preparation…..……...……..……...……...…pag.74
3.1.2. Libraries preparation and analysis……..……..…….pag.74
3.1.3. Filtering and Variant Prioritization……..……..…….……pag.75
3.1.4 Variant's identification……..……..…….....……..pag.75
3.1.4. Sanger Sequencing……...……pag.76
3.2 Splicing analysis……..……..……..pag.77
3.2.1 RNA Extraction and Reverse Transcriptase-PCR (RT-PCR)……...……pag.77
3.2.2 Transcript Characterization……..……...……..…..…pag.77
3.3 Site-specific mutagenesis……...……..……...pag.77
3.3.1. Immunofluorescence staining and Duolink studies on cells…..……..……...……pag.79
3.4 Western-blot procedure……..……..……pag.79
4. Discussion .....…pag.81
5. Limitation of the study.....……pag.86
6.Concluding remarks.....…….……....pag.86
7. Bibliography..... .pag.87

« Une des raisons les plus fortes qui conduisent les hommes à l'art et à la science c'est l'évasion de la vie quotidienne avec sa crudité douloureuse et son sombre manque d'espoir, de l'esclavage de ses désirs toujours changeants. »

Albert Einstein

Abstract

Intellectual disability and autism spectrum disorder are various conditions with features often including status epilepticus, attention deficit hyperactivity disorder, craniofacial dysmorphisms, symptoms overlapping to many neurodevelopmental disorders. A net discernment between these conditions appears tricky, and a clinical evolution for many pathologies has been observed from childhood to adulthood presenting symptoms common to different pathologies (e.g. KBG syndrome, Cornelia de Lange, Noonan syndrome…). Whole-Exome sequencing has recently been recommended as a first-tier diagnostic tool in the investigation of the genetic defects of neurodevelopmental disorders, being a valid tool to cope with the wide clinical and locus heterogeneity underlying the pathogenesis of intellectual disability and autism. Its application in the research field has elucidated novel altered pathways as well as it has raised the possibility to investigate that part of the genome still unknown. Moreover, thanks to the support of experimental evidence, mutation specific mechanisms have been unrevealed for the genes of intellectual disability and autism. We introduce a site-specific mutagenesis approach on the *NLGN4X* gene, which highlights a likely gain of function mechanism for a series of missense variants.

1.Introduction

Neurodevelopmental disorders (NDDs) are a wide group of conditions affecting brain development and various aspects of daily functioning. They include autism spectrum disorder (ASD), intellectual disability (ID), motor disability, seizures, learning disabilities, and attention deficit hyperactivity disorder (ADHD)[1]. In the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), NDDs are considered a group of conditions with onset in the developmental period, involving deficits that produce impairments of functioning [2]. At least 30% of NDDs are thought to have a genetic basis [3].

ID and ASD are the most common conditions overall affecting 3-5 % of the population. ID is a condition characterised by below-average intellectual functioning together with significant limitations in adaptive functioning. ID is generally assessed through an IQ test and attributed to individuals with a result <70[4]. It encompasses a clinically and genetically heterogeneous group of neurodevelopmental disorders often associated with ASD. ASD comprises neurodevelopmental conditions characterised by deficient social interactions, poor or absent communication, repetitive behaviours, and apparently limited interests[5].

ID and ASD can be non-syndromic or syndromic when peculiar facies, typical clinical traits, and/or growth abnormalities are reported[6]. Prenatal and perinatal events, such as drug and toxins exposure during pregnancy, are correlated to the ASD and ID onset. Examples of risk factors consistently associated with an ASD diagnosis are an advanced paternal age, low birth weight, birth defects, and perinatal hypoxia and respiratory stress. ASD risk factors overlap with those reported for ID. Among the significantly associated ID factor risks we can mention advanced maternal age, low maternal education, maternal alcohol use, maternal tobacco use, maternal diabetes, maternal hypertension, maternal epilepsy and maternal asthma, preterm birth, male gender, and low birth weight [7], [8].

Currently >2,000 genes are related to ID and ASD. They are implicated in many important biological processes such as cell cycle regulation, DNA methylation and repair, and damage response, chromatin remodelling transcription and translation[4]. About 12,000 genes have not yet been associated with human diseases, but we expect novel ID/ASD genes will be found among these as the collection and analysis of genetic data move forward[9]. The integration of the results from genomic and transcriptome data, as well as network analysis will further help the identification of novel ID/ASD candidate genes[10].

Combining clinical and molecular diagnosis is fundamental to deepen our knowledge of the pathogenic mechanisms underlying these disorders of neurodevelopment to target tailored treatments[11]. Single-nucleotide variants (SNVs), indels, and copy number variations (CNVs) have been identified as key variant types causing ID/ASD[12].

In this thesis, we take advantage of the Whole Exome Sequencing (WES) application to untangle the genetic bases of ID/ASD. However, the detection of novel ID/ASD candidate genes via WES is only the first step of the identification process of novel ID/ASD genes which demands subsequent functional investigation of their downstream effect on their related pathways[13].

1.1 Unlocking the biology of autism

ASDs are enigmatic diseases comprising classical idiopathic autism, Asperger's syndrome, Rett's syndrome, and pervasive developmental disorder. Other syndromic disorders, such as Down's syndrome, fragile X mental retardation, and tuberous sclerosis are often related with ASD[14]. The incidence of ASD has apparently increased recently, as a prevalence of 2.2-2.7% has been reported among children[15] It is widely recognized that this increase is due to the different diagnostic criteria used.

There is no definitive pharmacotherapy for the treatment of core symptoms of ASD. Therapies generally aim to reduce behavioural, educational, and immunological symptoms. Antipsychotic drugs only target secondary symptoms such as irritability, aggression, anxiety, or self-injurious behaviours[16].

Identifying the genes that contribute to autism is the most direct route to understanding the biology of the condition. Risk of autism has been linked to de novo copy number variants - deletions or duplications of large DNA regions. However, owing to the multigenic nature of these regions, determination of the causative gene is difficult.

WES has permitted the detection of de novo or inherited mutations in the coding regions as well as new candidate genes for ASD. An analysis of 238 families screened via WES (with one child with non-inherited ASD) from the Simons Simplex Collection revealed an enrichment of de novo SNVs in the affected individuals [17].

Modelling via computer simulation was engineered to calculate the likelihood that a SNV in the same gene in different individuals was responsible of their ID/ASD, as many de novo SNVs arise in siblings.

It results that de novo SNVs in brain-expressed genes were associated with ASD; these variants conferred risk of the disorder in at least 14% of the individuals studied[17].

A second WES analysis carried out on 189 trios (parents and ASD children) from the Simons Simplex Collection, found that 39% of 120 severe de novo mutations belong to highly interconnected protein networks, related to the β-catenin and chromatin remodelling pathways.

Noteworthy, de novo mutations were fourfold more likely to be inherited from the father than from the mother, with a correlation between paternal age and the number of de novo SNVs in their offspring[18].

Most of the ASD-related de novo SNVs identified in the three studies were found in independent genes, indicating the genetic heterogeneity of ASD. By use of statistical modelling, Sanders et al. showed that the presence of two or more de novo loss-of-function SNVs in the same gene is sufficient to demonstrate a genetic association with ASD[19]. After having proved a gene's pathogenicity, it remains to demonstrate its biological role and how it can be addressed to develop new future treatments.

1.2 ID/ASD heterogeneity and complexity

Heterogeneity of symptoms in patients with ID are frequently observed, including different core signs [20]. Often, ID and ASD are present in the same patient: 70% of ASD patients also show ID, and 40% of ID patients display ASD[21]. Many studies showed that 22% to 83% of children with ASD have clinically symptoms that satisfy the DSM-IV criteria for ADHD, and vice versa, 30% to 65% of children with ADHD have ASD symptoms. A meta-analysis has evaluated that about 70% of patients with ASD experience at least one additional psychiatric disorder whereas epilepsy was observed in about 30% of cases[22], [23].

ID complexity severely hampers the diagnostic ascertainment, etiological exploration, and treatment assessment[24]. A misjudgement of a variant can originate not only from the underestimation of its molecular impact as for noncoding ones, but, as the above-mentioned reasons may suggest, also from the lack of correspondence between the expected and the observed clinical phenotype[25]. Intrafamilial and interfamilial clinical variability, partial

medical descriptions in the literature, genetic modifiers, and other variants possibly modulating the symptoms make the genetic diagnosis even more difficult. Moreover, some conditions simply are not caused by genetic alterations[11]. Currently, approximately 30% of the ID /ASD patients remain without a molecular diagnosis.

Guidelines for genetic testing exist for ASD and for global developmental delay (GDD). These recommendations support the diagnostic decisions that, in certain cases, can be complicated by the GDDs complex genetic heterogeneity and require iterative genetic testing. The 2010 guideline from the American College of Medical Genetics and Genomics (ACMG) suggests that chromosomal microarray (CMA) and fragile X (FXS) testing should be first-tier tests for individuals with unexplained GDD/ID and/or ASD (except for females with ASD and normal cognition, for which FXS testing is not recommended). There are also considerations for single-gene testing of *MECP2* and *PTEN*[26]. Arteche-López et al recommend WES as a "first-tier test for the genetic diagnosis of ASD, when there is no suspicion of fragile X syndrome. Out of a diagnostic rate of 12, 8%, WES outperforms CMA (diagnostic rate 12,8%) permitting to diagnosis for mostly of the cases (75%) [27].

Both tests adopt an unbiased approach useful for conditions with heterogeneous genetics as NDDs are. Considering the diagnostic yield WES is recommended as first-tier test for NDDs[28]. If WES is nondiagnostic, the evaluation of CMA is suggested. If CMA is nondiagnostic, further tests could be considered as periodic clinical revaluation, reanalysis of WES data, allowing to enhance the molecular diagnostic yield by 10–16%, FXS testing, metabolic testing and/or mitochondrial DNA (mtDNA) sequencing, and karyotyping to assess for balanced chromosomal rearrangements [28].

The WES higher diagnostic yield for ID results from the capability to detect the most common (PVs) that are found in the population (single-nucleotide variants SNVs) representing many genetic loci for ID/ASD. This helps to deal with the high degree genetic and clinical heterogeneity of ID/ASD [20].

1.3 Whole Exome sequencing: benefits and drawbacks

The exome represents less than 1% of the human genome but contains 85% of known disease-causing variants [29]. Whole-exome sequencing (WES) is an application of the next generation sequencing technology exploring the variants in the coding regions of the genes. The advent of WES has enlightened monogenic forms of Global developmental delay (GDD) not detectable by CMA, FXS testing, or single-gene sequencing[4]. A relevant contribution came from studies involving genetic testing of biological parents or several affected family members[31].

WES is a powerful method to rapidly detect the cause of most genetic diseases and it is becoming more and more cost-effective and a reduction of the costs of the 76% in healthcare has been hypothesised[31]. It also enables to effectively perfectionate initial medical diagnosis of NDDs patients with an incomplete medical history and to dissect complex phenotypes misinterpreted as a unique clinical condition as a combination of multiple monogenic, digenic or oligogenic diseases; this eventually leads to simplifying patient management [32]. For these reasons, WES has largely been adopted in clinics in the last years.

Despite these indisputable positives, WES has some drawbacks that are being implemented: it do not detect chromosomal copy number variants (CNVs); it has problems in recognizing certain rearrangements or translocations; it do not identify epigenetic changes nor trinucleotide repeat expansion; it presents various technical limitations as the underestimation of positives because of low-covered regions and miscalling of false positives due to homopolymers; the WES sensitivity for the detection of mosaic events is lower compared to panel testing; which offers a higher read depth and sequence coverage[33]. Some of these limits could be overcome by combining WES to chromosomal microarray-CGH (CMA); the latter typically detects CNVs and regions of homozygosity, but not monogenic disorders[34].

However, the main challenge remains the interpretation of the huge amount of data generated through sequencing. Exclusion of causative variants could in fact be due to the selection of the de novo variants in a proband from an unaffected family, underestimating the incomplete penetrance and the variable expressivity of some clinical conditions[35]. Additionally, the variants classification should be subject to a periodic reanalysis going in

parallel with the increasing knowledge of the phenotypes and of the effects caused by the molecular changes. Indeed, it has been estimated that a likely pathogenic variant (LPV) has about 10% false positive rate as the cause of an individual's clinical presentation.

In the last few years WES has become the first-choice method to investigate genetic diseases. First, it gave us new opportunities to discover unknown disease-causing genes. Moreover, its effectiveness was clear in the field of syndromic intellectual disability, often characterised by phenotypic and genetic heterogeneity. WES is also an optimum method in cases of non-syndromic ID in which the lack of diagnostic handles makes obtaining a genetic diagnosis with traditional genetic tests arduous and sometimes enables personalised pharmacological treatments[36].

Yields of WES range between 30% to 50% for patients with mild to severe NDD/ID[36]. The diagnostic yield of WES analysis in a cohort of moderate–severe non syndromic ID patients was 49.2%. 31% for isolated NDDs, and >50% for the NDDs plus associated conditions. Several WES studies have focused on individuals with varying presentations of NDDs identifying genetic causes in as many as 61% of cases. Higher diagnostic rates have been observed in studies involving the genetic testing of parents or comprising other affected family members. The diagnostic rate for DNMs, although lower (11%), justifies the application of the WES for studying the causes of sporadic ID[30]. Recent literature showed us WES identification of two or more Mendelian diseases in the same patient with a complex phenotype[38]. It lets us think that with the natural evolution of the technique it will be even more common.

1.4 WES as resource for paediatric patients

Between 2009 and 2017 in the United States of America a relevant increase in the prevalence of children between 3 and 17 years old was observed for different conditions belonging to the spectrum of NDDs: developmental disability (16.2%–17.8%), ADHD (8.5%–9.5%,), ASD (1.1%–2.5%), and ID (0.9%–1.2%)[39].

Factors such as infection during pregnancy, birth complications, heredity, poison exposure, cancer, abuse, and low socioeconomic status are related to NDDs[40]. Mild symptoms were underestimated, and positive treatment was failed during the onset, therefore, it took a long time to reach the final diagnosis. For patients who could not receive an accurate diagnosis,

Trio-WES was performed. WES helps the assessment of significantly more conclusive diagnosis (29.3%) than the standard care pathway (7.3%) in paediatric neurology without incurring higher costs. Molecular diagnosis is crucial to improve outcome, prevent complications, and children's developmental outcome could benefit from early intervention. In a recent study, WES was performed for 17 children with a mean age of 5.6 years of age, with DD/ID reaching an overall diagnostic yield of 58.8% (10/17). Eight de novo variants were detected, including five SNVs/Indels and three CNVs. In addition, CNV analysis revealed three pathogenic CNVs, increasing the diagnostic yield by 17.6% (3/17), which agrees with a recent study according to the incorporation of exome-based CNV calling increased the diagnostic rate of trio-WES by 18.92% (14/74) in NDDs patients. The use of WES together with CNVs analysis leads to a positive outcome in about 54.05% (40/74) of the cases, with improvements in terms of health outcome, reduced morbidity and raising the possibility of life-saving treatments when performed in the early infancy[41], [42].

A study on a total of 102 paediatric patients diagnosed with ID and epilepsy with unknown causes highlighted the advantage in the utilisation of both CMA and WES as unbiased approaches for genetically heterogeneous conditions. The overall diagnostic yield of genetic aberrations was 33.3% (34/102), which comprised 50.0% with diagnostic CNVs and 50.0% with diagnostic SNVs. The diagnostic yield appeared to correlate positively with ID severity[43].

WES application appears promising when supported by a diagnostic suspicion and resolutive for paediatric patients with immature clinical traits, not clear markers of one syndrome rather than another [44]. Elucidating a molecular diagnosis is that it can modify clinical management for individuals and families affected by ID/ASD for example by entering a surveillance path for disease-related conditions, by avoiding repeated rounds of investigation, and/or by addressing further tests for associated medical conditions[4].

1.5 Gender differences

DSM-5 mentions male-to-female ratios of 4:1 for the diagnosis of ASD, 2:1 for the diagnosis of ADHD in children, and 1.6:1 and 1.2:1 for mild and severe ID, respectively; this increased percentage of males is often due to complex interactions between genetics, hormones, and environmental factors[2]. Classically, two theories have been trying to explain the different

male/female ratio observed in the NDDs: the female protective effect and the extreme male brain theory[45].

The female protective effect theory states that the differences in the female/ male percentages of the individuals affected with NDD can be explained by the genetic differences characterising males and females. Females have a second X chromosome, which compensate for the effect of "hits" that alone are sufficient for a male patient to develop symptoms. For example, in Fragile X syndrome, which is a genetic condition resulting in an expansion of the *FMR1* gene on the X chromosome, males usually have ID as their only X chromosome is affected; whereas females are less severely or not obviously affected because of the protective effect of their second unaffected X chromosome. Furthermore, the female brain matures faster than the male one and this could expose males to increasing risk for damage [46].

A study by Jacquemont et al. on thousands of individuals affected with NDDs strengthened the validity of the female protective effect. The distribution of deleterious autosomal variants (CNVs and SNVs) in male and female patients with an ascertained diagnosis of NDD proved a systematic excess of deleterious variants in females diagnosed with NDD. Also, mothers who underwent no medical investigation showed an increased mutational burden compared to fathers of probands with a NDD diagnosis. This evidence is in support of a "female protective model", for which female patients would require a higher mutational burden to reach the threshold needed to develop NDD symptoms than the male patients [47].

The Extreme Male Brain theory developed by Baron-Cohen asserts that autistic people have a hypermasculinized brain that is better at systemizing rather than empathy [48].

The study found typical females had more empathising brain types, males had more systemizing brain types and self-reported autistic people reported more systemizing or 'masculinized' responses and below average empathy. However, there was a similar sex pattern between typical and autistic people with females still showing higher rates of empathy and lower rates of systemizing, regardless of whether they were autistic. That autistic people, in particular females, had more masculine responses linked back to possible androgen exposure in utero (the androgen theory of autism) and other biological differences between the sexes [49]. Maternal and childhood stress have also been related to the development of neurodevelopmental disorders. Placental levels of the X-linked gene O-linked *N*-acetylglucosamine transferase (OGT) are important in neurodevelopmental programming and metabolic regulation. Within the placenta, OGT levels are higher in females than males as it

escapes X inactivation (where one copy of the X chromosome is inactivated). In case of maternal stress, OGT is reduced in both males and females and male placentas exposed to prenatal stress have lower levels of OGT for a sustained period. The role of OGT includes stabilising the protein EZHZ, which leads to an increase in the histone repressive mark H3K27me3; thus, the low levels of OGT in males would place chromatin in a more reactive state compared to females exposing them to a higher sensitivity to insults[50].

1.6 Site-directed mutagenesis on neuroligin 4X-linked

Experimental evidence is required to shed light on the different mechanisms of action of different variants. Site-directed mutagenesis, firstly introduced in 1978 by Michael Smith and colleagues, creates a specific mutation being able to characterise gene and protein structure–function relationships, protein–protein interactions, binding domains of proteins, or active sites of enzymes for the last three decades. A nucleotide sequence of interest is experimentally altered using synthetic oligonucleotides. The most used approach is to use an oligonucleotide that is complementary to part of a single-stranded DNA template but containing an internal mismatch to generate the mutation. In addition to single point mutations, site-directed mutagenesis may also be used to construct multiple mutations, insertions, or deletions[51]. We employed this technique to explore the role of variants in *NLGN4X,* which plays a role both in the synaptic transmission and development. ID/ASD conditions are frequently observed in combination with an epileptic status; this can let one think that an imbalance between excitatory and inhibitory transmission could be involved in ASD[52].

Neuroligin 4 X-linked (*NLGN4X)* is a member of the family of the neuroligins. Neuroligins are post-synaptic adhesion molecules needed for the functionality and for the maturation of synapses interacting with pre-synaptic molecules called neurexins. Both these cell-adhesion molecules mediate the synapses-signalling and shape the properties of the neural networks. Considering that the inputs and outputs at the level of the cerebral circuit depend on the features of the individual synapses as well as on the overall synaptic connectivity they could represent a molecular hotspot in the pathogenesis of NDDs [52].

NLGN4X has been found adopting a reverse transcription PCR approach on the human brain RNA, which highlighted the presence of the three human homologs of the three rat neuroligins and a cDNA corresponding to the fourth neuroligin. *NLGN4X* is expressed in the brain, particularly in the frontal cortex; its knockdown is causative of gene expression and

morphological alterations, up-regulation, and down-regulation of several biological pathways [53].

Neuroligin 4 X-linked is an X-linked postsynaptic scaffolding protein, with functional role in excitatory synapse development and maintenance. There are 5 homologous neuroligin genes in the human genome, *NLGN1, NLGN2, NLGN3, NLGN4X,* and *NLGN4Y*, encoding for neuronal cell-surface proteins with the 70% of amino acid identity and the same predicted domain structure with an N-terminal signal peptide, the esterase-like domain with two sites of alternative splicing, a linker region, and a cytosolic portion[54]. Currently, all the neuroligins have been related to NDDs. The role of *NLGN4X* in the ASD has been questioned not being definable as a common cause of it, but currently mutations in *NLGN4X* have been described in more than 50 ASD patients [55].

NLGN4X promotes neuritogenesis in immature human neurons at the level of the leading edge of growth cones. Super-resolution microscopy revealed that *NLGN4X* clustering induced growth cone enlargement and influenced actin filament organisation. These morphological effects were not observed by ASD-patients with *NLGN4X* variants.

Many PV types in *NLGN4X* have been reported. Deletions, nonsense, and frameshift variants that have been postulated to lead to likely LoF mechanisms. The effect of missense variants, instead, seems to be mutation specific, as displayed by the following examples: -R87W causes the retention of NL4 in the endoplasmic reticulum (ER) and its subsequent rapid degradation; it inhibits its surface expression as a consequence of folding impairment (LoF); -R704C enhances the binding affinity of the neuroligin for synaptic receptor increasing the number of excitatory synapses and thus acting through a GoF mechanism; -R101Q impairs the *NLGN4X* maturation, decreases its surface trafficking and inhibits its silencing of excitatory synapses, increasing the excitatory transmission (LoF).

It is relevant to know the *NLGN4X*-mutation specific mechanisms of action as pharmacological approaches have already been applied to rescue the phenotype caused by LoF *NLGN4X* variants. Treatment with the chemical chaperone 4-PBA (4-Phenylbutyric acid) recovered the abnormal cell surface level and synaptogenic activity caused by misfolding and ER retention [56].

Currently, *NLGN4X* has been associated with several non-syndromic neuropsychiatric disorders such as ID, ASD, anxiety, ADHD, and Tourette's syndrome. *NLGN4X*'s phenotype

was considered fully penetrant until female patients mildly affected, or similarly affected as males, were documented. In one case, this was explained via random X-inactivation[57].
Nevertheless, as part of the family of neuroligins, *NLGN4X* could be implied in the syndromic ASD pathogenesis. *NLGN3* KO in mice is causative of a reduction in the miniature postsynaptic current amplitudes (mEPSC) and of an occlusion of long-term depression mediated by metabotropic glutamate receptor (mGluR). An unbalanced function of mGluR has been observed in the Tuberous Sclerosis and in the X- Fragile syndromes.
Moreover, studies on the neuromuscular junction (NMJ) in *Caenorhabditis elegans* demonstrated that in mutant worm lacking the function of a specific microRNA, a retrograde signal coming from the muscle inhibits the acetylcholine release by axonal terminal of motor neurons. In the wt instead, a retrograde signal occurs when miR-21, a specific muscular mi-RNA is inactivated. This signal depends on the interaction between NLGI-1 and NRX-1 at the level of the muscle[58].
To date, the effect of missense variants and the full phenotypic spectrum in females is not well understood. We collect a series of missense variants in *NLGN4X* to deepen our knowledge regarding their functional mechanism.

2. Results

2.1 Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations

In this section, we describe a study on 200 families with at least one proband with an unexplained diagnosis of ASD/ID-related phenotype. Considering the high disease heterogeneity and the difficulties in making an early diagnosis for ID/ASD patients, we confirm the WES to be an effective diagnostic tool for devising tailored treatments in patients with ID and/or ASD, particularly those with additional clinical features.

Finally, *CACNA2D1* and *GPR14* have been proposed as novel candidate genes needing further studies to characterise their contribution and to show how their haploinsufficiency is linked to ID/ASD.

Concluding, WES was able to avoid patients' "diagnostic odyssey" for a significant fraction of families.

brain sciences

MDPI

Article

Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations

Floriana Valentino [1,2,†], Lucia Pia Bruno [1,2,†], Gabriella Doddato [1,2], Annarita Giliberti [1,2], Rossella Tita [3], Sara Resciniti [1,2], Chiara Fallerini [1,2], Mirella Bruttini [1,3], Caterina Lo Rizzo [3], Maria Antonietta Mencarelli [3], Francesca Mari [1,2,3], Anna Maria Pinto [3], Francesca Fava [1,2,3], Margherita Baldassarri [1,2], Alessandra Fabbiani [1,2,3], Vittoria Lamacchia [1,2,3], Elisa Benetti [2], Kristina Zguro [2], Simone Furini [2], Alessandra Renieri [1,2,3] and Francesca Ariani [1,2,3,*]

[1] Medical Genetics, University of Siena, 53100 Siena, Italy; floriana.valentino@dbm.unisi.it (F.V.); lucia.bruno@dbm.unisi.it (L.P.B.); gabriella.doddato@dbm.unisi.it (G.D.); giliberti@student.unisi.it (A.G.); sara.resciniti@dbm.unisi.it (S.R.); fallerini2@unisi.it (C.F.); mirella.bruttini@dbm.unisi.it (M.B.); francesca.mari@dbm.unisi.it (F.M.); francesca.fava@dbm.unisi.it (F.F.); margherita.baldassarri@dbm.unisi.it (M.B.); alessandra.fabbiani@dbm.unisi.it (A.F.); vittoria.lamacchia@dbm.unisi.it (V.L.); alessandra.renieri@dbm.unisi.it (A.R.)

[2] Med Biotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, 53100 Siena, Italy; elisa.benetti@dbm.unisi.it (E.B.); kristina.zguro@student.unisi.it (K.Z.); simone.furini@dbm.unisi.it (S.F.)

[3] Genetica Medica, Azienda Ospedaliera Universitaria Senese, 53100 Siena, Italy; rossella.tita@dbm.unisi.it (R.T.); lorizzo2@unisi.it (C.L.R.); mariaantonietta.mencarelli@unisi.it (M.A.M.); annamaria.pinto@dbm.unisi.it (A.M.P.)

* Correspondence: francesca.ariani@unisi.it; Tel.: +39-0577-233-303; Fax: +39-0577-233-325

† Co-first authors.

Abstract: Intellectual disability (ID) and autism spectrum disorder (ASD) belong to neurodevelopmental disorders and occur in ~1% of the general population. Due to disease heterogeneity, identifying the etiology of ID and ASD remains challenging. Exome sequencing (ES) offers the opportunity to rapidly identify variants associated with these two entities that often co-exist. Here, we performed ES in a cohort of 200 patients: 84 with isolated ID and 116 with ID and ASD. We identified 41 pathogenic variants with a detection rate of 22% (43/200): 39% in ID patients (33/84) and 9% in ID/ASD patients (10/116). Most of the causative genes are genes responsible for well-established genetic syndromes that have not been recognized for atypical phenotypic presentations. Two genes emerged as new candidates: *CACNA2D1* and *GPR14*. In conclusion, this study reinforces the importance of ES in the diagnosis of ID/ASD and underlines that "reverse phenotyping" is fundamental to enlarge the phenotypic spectra associated with specific genes.

Keywords: exome sequencing; intellectual disability; autism spectrum disorder

check for updates

Citation: Valentino, F.; Bruno, L.P.; Doddato, G.; Giliberti, A.; Tita, R.; Resciniti, S.; Fallerini, C.; Bruttini, M.; Lo Rizzo, C.; Mencarelli, M.A.; et al. Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations. *Brain Sci.* **2021**, *11*, 936. https://doi.org/10.3390/brainsci11070936

Academic Editor: Corrado Romano

Received: 10 June 2021
Accepted: 13 July 2021
Published: 16 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Intellectual disability (ID) is characterised by significant limitations in intellectual functioning (reasoning, learning, problem-solving) and adaptive behaviour (conceptual, social, and practical skills) that originate before the age of 18 [1]. Affecting 1–3% of the world's population, ID represents an important socio-economic problem in healthcare [2,3]. ID is characterized by limitations in cognitive functions that manifest as an intelligence quotient (IQ) below 70. ID may be "isolated" or "syndromic" when patients have peculiar facies, specific physical signs and/or an abnormal growth pattern [4].

Autism spectrum disorder (ASD) is characterized by deficient social interactions, poor or absent communication, repetitive behaviours, and apparently limited interests [5]. ASD generally becomes apparent after the first year of life and it has been reported in an increasing number (2.2–2.7%) of children, with boys four times more likely to be affected

than girls [6,7]. From a clinical point of view, the autism spectrum disorder is subdivided into "syndromic" when a co-occurrence between autism (ASD) and dysmorphic features, also other somatic or neurobehavioral abnormalities could be observed [8].

ID and ASD often co-exist and identifying the etiology of the two conditions remains challenging due to disease heterogeneity. Accurate clinical, as well as molecular diagnoses are essential for a deeper understanding of the pathogenesis of these conditions and for devising tailored treatments [4]. Until the advent of the first next generation sequencing (NGS) platforms, a large fraction of cases remained not diagnosed, with many families undergoing a "diagnostic odyssey" [9]. The introduction of exome or genome sequencing (ES/GS) has significantly improved diagnostic rates in individuals with suspected ID/ASD genetic disorders refractory to conventional diagnostic testing [10].

In the present study, ES was generated for a total of 200 individuals (84 ID and 116 ID/ASD patients). Pathogenic or likely pathogenic (P/LP) variants were found in 43 individuals (22%), with 45 variants of uncertain significance in an additional 20% (40/200). Our data strongly support the value of large-scale sequencing, especially ES within proband-parent trios, as an effective first-choice diagnostic tool.

2. Materials and Methods

2.1. Selection of Patients and DNA Samples' Preparation

Genetic counselling was carried out to evaluate each patient's personal and familial history. Parents provided and signed a written informed consent at the Medical Genetics department of the University of Siena, Italy, for exome sequencing analysis, clinical data usage, and the use of DNA samples from the tested individuals for both research and diagnosis purposes. We analysed a total of 200 patients affected by ID and ID/ASD (84 with ID and 116 with ID and ASD) collected from January 2019 until the end of March 2021.

Genomic DNA from the parents was isolated from EDTA peripheral blood samples using MagCore HF16 (Diatech Lab Line, Jesi, Ancona, Italy) according to the manufacturer's instructions.

2.2. Exome Sequencing

Sample preparation was performed following the Illumina DNA Prep with Enrichment manufacturer protocol. A bead-based transposome complex is used to perform tagmentation, a process that fragments the genomic DNA and then tags it with adapter sequences in one step. After saturation with input DNA, the bead-based transposome complex fragments a set number of DNA molecules. This fragmentation provides flexibility to use a wide DNA input range to generate normalized libraries with a consistent tight fragment size distribution. Then a limited-cycle PCR adds adapter sequences to the ends of a DNA fragment. A subsequent target enrichment workflow is then applied. Following pooling, the double-stranded DNA libraries are denatured and biotinylated. Illumina Exome Panel v1.2 (CEX) probes are hybridized to the denatured library fragments. Then Streptavidin Magnetic Beads (SMB) capture the targeted library fragments within the regions of interest. Then the indexed libraries are eluted from beads and further amplified before sequencing. The exome sequencing analysis was performed on the Illumina *NovaSeq6000 System* (Illumina San Diego, CA, USA) according to the NovaSeq6000 System Guide. Reads were mapped against the hg19 reference genome using the Burrow-Wheeler aligner BWA [11]. Variant calling was obtained using an in-house pipeline which takes advantage of the GATK Best Practices workflow [12].

Prioritization of the variants was obtained excluding polymorphisms (minor allele frequency, MAF <0.01), synonymous variants, variants classified as benign or likely benign. Frameshift, stopgain, and splice site variants were prioritized as pathogenic. Missense variants were predicted to be damaging by CADD-Phred prediction tools. The potential impact of variants on splicing was evaluated using Alamut® Visual software—version

2.11–0 (Interac-tive Biosoftware, Rouen, France), which employs five different algorithms: SpliceSiteFinder-like, MaxEntScan, NNSPLICE, GeneSplicer, and HumanSplicingFinder.

The following public databases were used for the interpretation of the variants: ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/, accessed on 9 June 2021), LOVD (https://databases.lovd.nl/shared/genes, accessed on 9 June 2021), the Human Genome Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/index.php3, accessed on 9 June 2021).

3. Results

3.1. Clinical Characteristics of Patients

We enrolled 200 families (574 individuals total) with at least one proband with an unexplained diagnosis of ASD/ID-related phenotype. In particular, 181 families had one affected proband and 4 families had two affected probands. ES was sequenced to an average depth of 100×, respectively, with ~94% of bases covered ≥ 20×. The study population had a mean age of 15 years and was 64% male (128/200). We divided into four groups on the basis of the age (0–10; 11–18; 19–30; 31–49) and gender (Figure 1). All individuals displayed ID, 58% (116/200) associated with ASD. The totality of the patients had been subjected to genetic testing prior to enrolment in this study. Our ID and ID/ASD cohort was characterized by the presence of additional associated clinical findings for 90% of the patients (179/200). These included epilepsy (n = 41/200, 20%), hypotonia (n = 10/200, 5%), and MRI abnormalities observed in 13 patients. Craniofacial dysmorphisms (n = 106/200, 53%) were found in 53% of the cases. The clinical descriptions of the 200 patients are summarized in Table S1.

Figure 1. Demographic analysis of the cohort, filtered by age and gender.

3.2. P/LP Variants Identified by ES

ES was performed in 200 probands with ASD and/or ID and 41 different pathogenic/likely pathogenic (P/LP) variants were identified with a detection rate of 22% (Table S2). Clinical features of patients with pathogenic variants in disease genes were described in Table 1.

Variants were classified based on frequency, mutation category, literature, and databases such as ClinVar. Affected individuals were categorized based on the number of parents that were sequenced along with the proband(s): proband-parent trios (184); duos with one parent (6); and proband-only singletons (10). Most P/LP variants were missense variants (15/200; 8%), while 5% (9/200) were nonsense, 7% (13/200) frameshift and 2% (4/200) splicing variants. Most (62%) P/LP variations occurred de novo, while 29% of individuals inherited P/LP variants as heterozygotes or homozygotes. Four out of 200 (2%) participants who harboured a P/LP result were sequenced with one or no biological parent and thus have unknown inheritance.

Mutations in the following genes were found in patients suffering from ID: ADNP, AP4M1, ATP1A3, BSCL2, CACNA1A, CSNK2B, CTU2, DEPDC5, DYRK1A, EFTUD2, HK1, IQSEC2, KANSL1, KIF1A, KMT2A, MBOAT7, MED13L, MMACHC, POGZ, PTPN11, RHOBTB2, SHANK3, SPG7, SPTBN2, TBCE, TUBA1A, WDR45, WFS1. In patients presenting autism and ID we found pathogenic variants in: BCOR, DDX3X, FGFR3, KCNQ3, SHANK3, SYNGAP1, TREX1, UPF3B, WFS1 (Table S2).

Two de novo P/LP variants were found in the new candidate genes: *CACNA2D1* and *GPR14* (Table S3), and the corresponding clinical pictures were reported in Table 2.

3.3. Uncertain Variants Identified by ES

We further reported 45 uncertain variants including in this number the variants that are currently considered to have an uncertain significance in the databases and other missense variants that have not been previously described in the scientific literature (Table S4). The effect on the encoded mutated proteins has been predicted using CADD (combined depletion annotation depletion). In our study the majority of uncertain and of P/LP variants fall in genes that play a role in the axon guidance and in the neurodevelopment processes (https://reactome.org/PathwayBrowser/#/, accessed on 6 July 2021).

Table 1. Clinical features of patients with pathogenic variants in disease genes.

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs	OMIM Phenotype (n°)
#1	*ADNP*	c.2314dup	p.(Thr772Asnfs *16)	Male	7	ID	Yes (Plagiocephaly, prominent forehead, sparse eyebrows, broad nasal bridge, small ears, uplifted earlobes, conical teeth)	Bilateral clinodactyly of the 5th toe, supernumerary nipple, epilepsy, heteroaggressive behaviour	Helsmoortel-van der Aa syndrome (#615873)
#2	*AP4M1*	c.1257_1282del c.1317G > A	p.(Val421Alafs *98) p.(Trp439 *)	Female	18	ID	Yes (microcephaly, triangular face, widow's peak, thick eyebrows, large nose, flat philtrum, thin lips, high vaulted and narrow palate)	Short stature, clinodactyly of the fifth finger, laryngomalacia, splenium thinning, epilepsy, hearing loss, flat foot, stereotypies, epilepsy, speech defect	Spastic paraplegia 50, autosomal recessive (#612936)
#3	*ATP1A3*	c.2116G > A	p.(Gly706Arg)	Female	22	ID	No	Epilepsy, ataxia, limited speech	Alternating hemiplegia of childhood 2 (#614820)/CAPOS syndrome (#601338)/Dystonia-12 (#128235)
#4	*BCOR*	c.776C > T	p.(Ser259Leu)	Male	14	ID and ASD	Yes (low hairline, periorbital swelling, big mouth, full lips)	Stereotypies	Microphthalmia, syndromic, 2 (#300166)
#5	*BSCL2*	c.856C > T	p.(Arg286 *)	Female	22	ID	Yes (elongated facies, prominent ears, high nasal bridge, open mouth and protruded tongue)	Dystonia, myoclonia, drug-resistant epilepsy, language delay, psychomotor delay, bruxism, hypertrichosis	Encephalopathy, progressive, with or without lipodystrophy (#615924)
#6	*BSCL2*	c.856C > T	p.(Arg286 *)	Female	16	ID	Yes (facies with coarse features, synophria, bulbous nose, large ears, large mouth)	Dystonia, myoclonia, hypertrichosis, language delay, drug-resistant epilepsy, psychomotor delay	Encephalopathy, progressive, with or without lipodystrophy (#615924)
#7	*CACNA1A*	c.4055G > A	p.(Arg1352Gln)	Male	14	ID	Yes (Craniosynostosis, long filter, long eyelashes, thin upper lip)	Stereotypies, alternating hemiplegia, language delay	Developmental and epileptic encephalopathy 42 (#617106); Episodic ataxia, type 2 (#108500); Migraine, familial hemiplegic, 1 (#141500); Migraine, familial hemiplegic, 1, with progressive cerebellar ataxia (#141500); Spinocerebellar ataxia 6 (#183086)

Table 1. *Cont.*

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs	OMIM Phenotype (n°)
#8	*CSNK2B*	c.170del	p.(Glu57Glyfs *15)	Female	20	ID	Yes (Open mouth, protruding tongue, narrow palate, dental anomalies)	Epilepsy, hypotonia, developmental delay, absent language	Poirier-Bienvenu neurodevelopmental syndrome (#618732)
#9	*CTU2*	c.881C > A	p.(Ser294 *)	Female	22	ID	Yes (broad eyebrows, low anterior and posterior hairline, forehead hypertrichosis, prominent ears, eye asymmetry, large nose with wide nostrils, large mouth, dental anomalies)	Upper limb hypotonia, lower limb hypertonia, recurrent infections, thinning corpus callosum, growth retardation, pectus excavatum, clinodactyly of the 5th fingers.	Microcephaly, facial dysmorphism, renal agenesis, and ambiguous genitalia syndrome (#617057)
#10	*DEPDC5*	c.1453C > T	p.(Arg485 *)	Male	11	ID	Yes (low hairline, flattened nasal bridge, bulbous nasal tip, anteverted nostrils, full lips, small and widely spaced teeth, anteverted ears)	Epilepsy, language delay	Epilepsy, familial focal, with variable foci 1 (#604364)
#11	*DDX3X*	c.976C > T	p.(Arg326Cys)	Female	12	ID and ASD	Yes (Microcephaly, long face, smooth and long philtrum, strabismus, up slanting palpebral fissures)	Developmental delay, stereotypic hand movements, hypotonia, bruxism, sialorrhea, corpus callosum hypoplasia	Intellectual developmental disorder, X-linked, syndrome, Snijders Blok type (#300958)
#12	*DYRK1A*	c.1669C > T	p.(Gln557 *)	Male	16	ID	Yes (Microcephaly, narrow forehead, frontal bossing, depressed nasal bridge, short philtrum, prognathism)	Hypotonia, epilepsy, stereotypies, absent language	Mental retardation, autosomal dominant 7 (#614104)
#13	*EFTUD2*	c.702 + 1G > A	NA	Female	22	ID	Yes (prominent columella, sloping forehead, high and scattered eyebrows, deeply set eyes, nose with anteverted nostrils, narrow palate, prominent incisors, prominent ears)	Hypertrichosis, small hands and feet, joint hyperextensivity	Mandibulofacial dysostosis, Guion-Almeida type (#603892)
#14	*FGFR3*	c.749C > G	p.(Pro250Arg)	Female	4	ID and ASD	Yes (craniosynostosis, flat forehead, down slanting palpebral fissures, high nasal bridge)	Language delay, hyperchromic and hypochromic spots	Muenke syndrome (#602849) CATSHL syndrome (#610474)

Table 1. *Cont.*

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs	OMIM Phenotype (n°)
#15	*HK1*	c.1367C > T	p.(Thr456Met)	Male	16	ID	Yes (proptosis, long eyelashes, synophria, narrow palate, dental anomalies, small mouth, and full lips)	Bilateral cryptorchidism, psychomotor delay, poor vision, ventricular system dilation, immature hippocampal structures, spastic paraparesis, epilepsy, hypothonia, absent language	Neurodevelopmental disorder with visual defects and brain anomalies (#618547)
#16	*IQSEC2*	c.3780del	p.(Gln1261Serfs *136)	Female	22	ID	No	Strabismus, cerebral atrophy, sialorrhea, motor problems, bruxism	Mental retardation, X-linked 1/78 (#309530)
#17	*KANSL1*	c.985_986del	p.(Leu329Glufs *22)	Female	6	ID	Yes (cleft palate, deeply set eyes, high nasal bridge, bulbous nasal tip)	Language delay, epilepsy, aggressiveness	Koolen-De Vries syndrome (#610443)
#18	*KCNQ3*	c.688C > T	p.(Arg230Cys)	Male	39	ID and ASD	Yes (triangular facies, high nasal bridge, bulbous nasal tip, open mouth, micrognathia, narrow and downturned eyelids)	Stereotypies, aggressiveness, bladder anomalies	Seizures, benign neonatal, 2 (#121201)
#19	*KIF1A*	c.37C > T	p.(Arg13Cys)	Female	16	ID	No	Spastic paraparesis, behaviour disorder, slight enlargement of the interfolial spaces of the cerebellar hemispheres, hypertone	Spastic paraplegia 30, autosomal dominant (#610357)
#20	*KIF1A*	c.914C > T	p.(Pro305Leu)	Female	10	ID	Yes (sparse eyebrows, small nose, thin upper lip, dental anomalies, chubby cheeks)	Language delay, cerebellar and worm atrophy, psychomotor delay, brain abnormalities, hypertrichosis, bilateral clinodactyly of the second/third/fourth/fifth toes	Spastic paraplegia 30, autosomal dominant (#610357)
#21	*KIF1A*	c.914C > T	p.(Pro305Leu)	Male	49	ID	No	Ataxia, spastic paraparesis, angioma, nystagmus, epilepsy	Spastic paraplegia 30, autosomal dominant (#610357)
#22	*KMT2A*	c.5256delA	p.(Ala1753Profs*70)	Male	16	ID	Yes (thick eyebrows, narrow and short eyelids, depressed helix, nose with bulbous tip, thick lips)	Hypertrichosis of the legs, arms and lumbar region, large hands, short fifth ray of the foot	Wiedemann-Steiner syndrome (#605130)

Table 1. *Cont.*

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs	OMIM Phenotype (n°)
#23	*MBOAT7*	c.477C > G	p.(Tyr159 *)	Female	2	ID	No	Motor stereotypies, epilepsy, thinning of the corpus callosum, ventricular enlargement	Mental retardation, autosomal recessive 57 (#617188)
#24	*MED13L*	c.72 + 1G > T	NA	Female	49	ID	Yes (sparse eyebrows, hypertelorism, narrow eyelids, gingival hypertrophy)	Syndactyly of the second and third toes, small feet, hypertrophy of the limbs and truncal obesity, strabismus	Mental retardation and distinctive facial features with or without cardiac defects (#616789)
#25	*MMACHC*	c.440G > C	p.(Gly147Ala)	Male	12	ID	Yes (synophria, horizontal eyebrows, wide nasal tip, anteverted nostrils, long filter, dental anomalies)	Spastic paraparesis, language delay, polyneuropathy	Methylmalonic aciduria and homocystinuria, cblC type (#277400)
#26	*POGZ*	c.1180_1181del	p.(Met394Valfs *9)	Female	6	ID	Yes (microcephaly, deeply set eyes, nose with bulbous tip, anteverted nostrils, full lips)	Hyperactivity, blepharophimosis, brachydactyly, nail hypoplasia, kidney abnormalities, language delay	White-Sutton syndrome (#616364)
#27	*POGZ*	c.1180_1181del	p.(Met394Valfs *9)	Male	12	ID	Yes (narrow bitemporal diameter, narrow and upward eyelid rims, deep philtrum, progatism, exaggerated Cupid's bow, buccal rim pointing downwards, uplifted ear lobe)	Hypotonia, obesity	White-Sutton syndrome (#616364)
#28	*POGZ*	c.1180_1181del	p.(Met394Valfs *9)	Female	44	ID	No	Microcephaly, brachydactyly, nail hypoplasia	White-Sutton syndrome (#616364)
#29	*PTPN11*	c.1471C > A	p.(Pro491Thr)	Male	2	ID	Yes (high forehead, low-set ears with large, downward-pointing auricles, down-slanting eyelids, broad nasal tip, long and thick filter, exaggerated Cupid's bow)	Café au lait spots in the thoracic and lumbar region, nail hypoplasia of the fifth toe, neurodevelopmental delay, cryptorchidism	Noonan syndrome 1 (#163950)

Table 1. *Cont.*

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs	OMIM Phenotype (n°)
#30	*RHOBTB2*	c.1382G > A	p.(Arg461His)	Male	18	ID	Yes (microcephaly, thick eyebrows, deeply set eyes, square chin)	epilepsy, thinning corpus callosum, language delay, cold extremities, cerebellar atrophy, muscle hypotrophy	Developmental and epileptic encephalopathy 64 (# 618004)
#31	*SHANK3*	c.2717_2718dup	p.(Ser907Alafs *3)	Male	45	ID and ASD	Yes (alopecia in the fronto-temporal region, depressed ocular region, thin upper lip)	Absent language, obsessive behaviour, tapered fingers with widening of the intermediate interphalangeal joints, nail dystrophy, clinodactyly of the third toe, stereotypies, mouth chewing automatisms	Phelan-McDermid syndrome (#606232)
#32	*SHANK3*	c.2313 + 1G > A	NA	Male	21	ID and ASD	No	epilepsy, psychomotor delay	Phelan-McDermid syndrome (#606232)
#33	*SHANK3*	c.3250_3253del	p.(Leu1084Cysfs *9)	Male	18	ID	No	Language delay, psychomotor delay, manual stereotypies, increased tolerance to pain	Phelan-McDermid syndrome (#606232)
#34	*SPG7*	c.233T > A	p.(Leu78 *)	Female	10	ID	Yes (high forehead, spaced teeth, pursed lips attitude)	Epilepsy, neurodevelopmental regression, stereotypies, syndactyly	Spastic paraplegia 7, autosomal recessive (#607259)
#35	*SPTBN2*	c.1310G > A	p.(Arg437Gln)	Male	17	ID	Yes (sloping forehead, prognathism)	Strabismus, hyperchromic spot in the thoracic area, cerebellar atrophy, language delay, ataxia	Spinocerebellar ataxia 5 (#600224)
#36	*SYNGAP1*	c.2337–1G > A	NA	female	26	ID and ASD	No	Stereotypies, self-harm, cold extremities, sphincter control not acquired, language regression	Mental retardation, autosomal dominant 5 (#612621)
#37	*TBCE*	c.464T > A c.134dupA	p.(Ile155Asn) p.(Arg46Glufs *5)	Male	21	ID	Yes (bulbous nasal tip, full lips, spaced teeth)	Scoliosis, sialorrhea, optic atrophy, psychomotor delay, epilepsy, spastic paraparesis, absent language.	Encephalopathy, progressive, with amyotrophy and optic atrophy (#617207)

Table 1. *Cont.*

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs	OMIM Phenotype (n°)
#38	*TREX1*	c.558_573del	p.(Phe186Leufs *24)	Male	10	ID and ASD	No	Gastroesophageal reflux, sphincter control not acquired, stereotyped behaviour	Aicardi-Goutieres syndrome 1 dominant and recessive (#225750)
#39	*TUBA1A*	c.352G > A	p.(Val118Met)	Male	9	ID	Yes (advanced hairline, long and thick eyebrows, anteverted nostrils, long filter, thin upper lip)	Hypoplasia of the cerebellar vermis, thinned corpus callosum, cerebellar asymmetry, angioma, hypertrichosis, epilepsy, apraxia, ataxia, and psychomotor delay	Lissencephaly 3 (#611603)
#40	*UPF3B*	c.1288C > T	p.(Arg430 *)	Male	7	ID and ASD	Yes (Wide forehead, arched eyebrows, deeply set eyes, pointed chin)	Limited speech, neurodevelopmental delay	Mental retardation, X-linked syndromic 14 (#300676)
#41	*WDR45*	c.66del	p.(Cys23Alafs *15)	Female	20	ID	Yes (thick eyebrows, prominent upper arch, hyperemic gums, high palate)	Limited speech, scoliosis, locomotor impairment, manual stereotypies, tapered fingers	Neurodegeneration with brain iron accumulation 5 (#300894)
#42	*WFS1*	c.124C > T	p.(Arg42 *)	Male	6	ID and ASD	No	Language delay, oppositional and provocative behaviour	Wolfram-like syndrome, autosomal dominant (#614296)
#43	*WFS1*	c.1230_1233del	p.(Leu412Serfs *29)	Female	14	ID	Yes (microcephaly, synophria, long eyelashes, bulbous nasal tip)	Hypertrichosis, drug-resistant seizures, spastic tetraparesis, renal failure.	Wolfram-like syndrome, autosomal dominant (#614296)

* means change in a stop codon.

Table 2. Clinical features of patients with pathogenic variants in the new candidate genes.

Patient	Gene	Variant (HGVS)	Protein (HGVS)	Gender	Age (Years)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs
#44	CACNA2D1	c.659–2_659-1insT	NA	Female	14	ID	Yes (deeply set eyes, squared facies, bulbous nasal tip, full lips, horizontal eyebrows, enlarged nasal bridge, mouth with downward corners, anteverted nostrils)	Language delay, epilepsy
#45	GPR14 (before UTS2R)	c.844C > T	p.(Gln282 *)	Male	6	ID and ASD	Yes (long and large nose, broad nasal bridge, anteverted nostrils, deep philtrum)	Frequent infections

* means change in a stop codon.

4. Discussion

This study emphasizes the clinical diagnostic relevance of ES in patients with ID and/or autism with additional clinical features. In particular, in a cohort of 200 patients, we reached a diagnostic yield of 22% (43/200) with a higher rate in ID patients (33/84; 39%) with respect to ID and ASD patients (10/116; 9%). The diagnostic yield is lower with respect to other studies that employ ES in neurodevelopmental disorder (30–43%) [13]. ES has technological limitations, including the inability to detect noncoding variants, copy number variants (CNVs), epigenetic changes, and trinucleotide repeat expansion [14]. In our cohort, 30% of cases have not been screened for CNVs and this could have underestimated the presence of other pathogenic genetic alterations, in particular in patients with ASD. Until recently, whole genome chromosomal microarray was recommended as a first-tier clinical genetic test for detecting disease-causing CNVs in individuals with ASD [15–17].

Most of the previous diagnoses failed for the atypical phenotypic presentation of well-established genetic syndromes. The *DDX3X* mutated patient (#11) shows Rett-like spectrum features with typical hand-washing stereotypes and was initially screened for mutations in *MECP2*, *FOXG1* and *CDKL5* genes [18–21]. Differently, a hundred patients are reported in literature mutated in *DDX3X* with various clinical features including hypotonia, movement disorder, behavioural problems, corpus callosum hypoplasia and epilepsy [22,23]. Another atypical clinical picture was manifested by a patient bearing the mutation c.688C > T (p.(Arg230Cys)) in *KCNQ3*, who did not suffer from any status epilepticus but showing ID, autism, stereotypies, aggressiveness, bladder anomalies; he also presented craniofacial dysmorphisms (patient #18). *KCNQ3* pathogenic alterations are generally linked to the occurrence of seizures but recently patients with no EEG abnormalities have been described [24]. *KIF1A* gene was found altered in three unrelated cases with different phenotypic presentations (patients #19, #20, #21). *KIF1A* mutations cause NESCAV syndrome (NESCAVS), a neurodegenerative disorder characterized by global developmental delay, progressive spasticity, ID, speech delay, learning disabilities and/or behavioural abnormalities [25]. The mutation c.37C > T (p.(Arg13Cys)) in *KIF1A* was found in patient #19 with spastic paraparesis, behaviour disorder, slight enlargement of the interfolial spaces of the cerebellar hemispheres, hypertone and no craniofacial dysmorphisms. In particular the same mutation c.914C > T (p.(Pro305Leu)) was shown in two different patients (#20, #21). One of these, presented language delay, cerebellar and vermis atrophy, psychomotor delay, brain abnormalities, hypertrichosis, bilateral clinodactyly, and facial dysmorphisms (patient #20). The other showed ataxia, spastic paraparesis, angioma, nystagmus, and seizures (patient #21). Mutations in *POGZ* are associated with the White-Sutton syndrome, which is a neurodevelopmental disorder characterized by delayed psychomotor development and a characteristic constellation of dysmorphic facial features [26]. Additional features may include hypotonia, sensorineural hearing impairment, visual defects, joint laxity, and gastrointestinal difficulties [27]. The pathogenic variant c.1180_1181del (p.(Met394Valfs*9)) was carried by two siblings and their mother (#26, #27, #28). The sister exhibited craniofacial dysmorphisms and ID; she also showed hyperactivity, blepharophimosis, brachydactyly, nail hypoplasia, kidney abnormalities and language delay as additional clinical signs (patient #26). The brother was affected by ID, hypotonia, obesity and had some craniofacial dysmorphisms (patient #27). Their mother instead displayed ID, microcephaly, brachydactyly, and nail hypoplasia (patient #28). Another example, *SHANK3* was found altered in three unrelated patients (#31, #32, #33). Mutations in *SHANK3* cause Phelan-McDermid syndrome, a developmental disorder with variable features including neonatal hypotonia, global developmental delay, absent to severely delayed speech, autistic behaviour, and minor dysmorphic features [28–30]. One of the three mutated patients did not show neither autism nor dysmorphic features and was initially classified as a "non-syndromic" ID case.

The following new candidate genes for ID/ASD have emerged: *CACNA2D1* and *GPR14*. They all show de novo truncating variants (patients #44–#45). *CACNA2D1* encodes the alpha-2/delta subunit of skeletal muscle and brain voltage-dependent calcium

channels [31]. A genomic aberration affecting the *CACNA2D1* gene has been previously characterized in patients with epilepsy and ID, pinpointing the gene as an interesting candidate gene for these clinical features [32]. Mice-bearing point mutations in the *CACNA2D1* gene have an abnormal central nervous system synaptic transmission [33]. *GPR14* gene, encoding the orphan G protein-coupled receptor 14 for Urotensin II, is widely expressed in the brain and spinal cord [34]. We found 45 variants of unknown significance (VUS) in 40 patients. These variants are 22/84 (26%) in ID patients and 18/116 (16%) in ID/ASD patients. These variants are mostly missense 42/45 (93%) and CADD $\geq$ 25 in 20/42 (48%) of cases. With increased knowledge over time, exome reanalysis may change the clinical interpretation of a VUS. Thus, it is important to list all the VUS, analyse them periodically and write a report in case of changes to provide a timely response for patients and families. An accurate molecular diagnosis allows for precise genetic counselling and has the potential to change clinical management.

5. Conclusion

ES was able to avoid a sort of "diagnostic odyssey" for a significant fraction of families consisting in the step-by-step application of the traditional genetic methods. ES revealed atypical phenotypic presentations and new candidate genes for ID/ASD. Further studies are needed to better characterize the contribution of new candidates and to show how their haploinsufficiency can determine ID/ASD.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/brainsci11070936/s1, Table S1: Clinical features of the 200 ID and ID/ASD patients; Table S2: P/LP variants identified in disease genes; Table S3: P/LP variants identified in candidate disease genes; Table S4: Uncertain variants identified by ES

Author Contributions: F.V., L.P.B., G.D., A.G., A.R. and F.A. have made substantial contributions to conceptions and design and have been involved in drafting the paper. R.T., S.R., C.F., M.B. (Mirella Bruttini), E.B., K.Z., S.F. has made substantial contributions to acquisition and analysis of the data. M.A.M., F.M., A.M.P., F.F., M.B. (Margherita Baldassarri), C.L.R., A.F., V.L., made a clinical evaluation. All authors have given final approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding: This research received no external fundings.

Institutional Review Board Statement: Ethical review and approval were not required for the study on human participants in accordance with local legislation and institutional requirements.

Informed Consent Statement: The patients/participants provided their written informed consent to participate in this study.

Data Availability Statement: NGS data has been deposited in publicly accessible repositories. The data can be found here: http://nigdb.cineca.it/.

Acknowledgments: We are grateful to our patient for the cooperation. The "Cell lines and DNA bank of Rett Syndrome, X-linked mental retardation, and other genetic diseases," member of the Telethon Network of Genetic Biobanks (project nos. GTB12001 and GFB18001), funded by Telethon Italy, and of the EuroBioBank network provided us with specimens.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Tassé, M.J.; Luckasson, R.; Nygren, M. AAIDD proposed recommendations for ICD-11 and the condition previously known as mental retardation. *Intellect. Dev. Disabil.* **2013**, *51*, 127–131. [CrossRef] [PubMed]
2. Kolset, S.O. Intellectual disability and nutrition-related health. *EMBO Mol. Med.* **2020**, *12*. [CrossRef]
3. Maulik, P.K.; Mascarenhas, M.N.; Mathers, C.D.; Dua, T.; Saxena, S. Prevalence of intellectual disability: A meta-analysis of population-based studies. *Res. Dev. Disabil.* **2011**, *32*, 419–436. [CrossRef]
4. Chiurazzi, P.; Kiani, A.K.; Miertus, J.; Barati, S.; Manara, E.; Paolacci, S.; Stuppia, L.; Gurrieri, F.; Bertelli, M. Genetic analysis of intellectual disability and autism. *Acta Biomed.* **2020**, *91*, e2020003. [CrossRef] [PubMed]

5. Sharma, S.R.; Gonda, X.; Tarazi, F.I. Autism Spectrum Disorder: Classification, diagnosis and therapy. *Pharmacol. Ther.* **2018**, *190*, 91–104. [CrossRef] [PubMed]
6. Baio, J.; Wiggins, L.; Christensen, D.L.; Maenner, M.J.; Daniels, J.; Warren, Z.; Kurzius-Spencer, M.; Zahorodny, W.; Rosenberg, C.R.; White, T.; et al. Prevalence of autism spectrum disorder among children aged 8 Years—Autism and developmental disabilities monitoring network, 11 Sites, United States, 2014. *MMWR Surveill. Summ.* **2018**, *67*, 1–23. [CrossRef]
7. Zablotsky, B.; Black, L.I.; Maenner, M.J.; Schieve, L.A.; Danielson, M.L.; Bitsko, R.H.; Blumberg, S.J.; Kogan, M.D.; Boyle, C.A. Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics* **2019**, *144*, e20190811. [CrossRef] [PubMed]
8. Baranova, J.; Dragunas, G.; Botellho, M.C.S.; Ayub, A.L.P.; Bueno-Alves, R.; Alencar, R.R.; Papaiz, D.D.; Sogayar, M.C.; Ulrich, H.; Correa, R.G. *Autism Spectrum Disorder: Signaling Pathways and Prospective Therapeutic Targets*; Springer: Berlin/Heidelberg, Germany, 2021; Volume 41, ISBN 0123456789.
9. Thevenon, J.; Duffourd, Y.; Masurel-Paulet, A.; Lefebvre, M.; Feillet, F.; El Chehadeh-Djebbar, S.; St-Onge, J.; Steinmetz, A.; Huet, F.; Chouchane, M.; et al. Diagnostic odyssey in severe neurodevelopmental disorders: Toward clinical whole-exome sequencing as a first-line diagnostic test. *Clin. Genet.* **2016**, *89*, 700–707. [CrossRef] [PubMed]
10. Munnich, A.; Demily, C.; Frugère, L.; Duwime, C.; Malan, V.; Barcia, G.; Vidal, C.; Throo, E.; Besmond, C.; Hubert, L.; et al. Impact of on-site clinical genetics consultations on diagnostic rate in children and young adults with autism spectrum disorder. *Mol. Autism* **2019**, *10*, 1–10. [CrossRef]
11. Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **2010**, *26*, 589–595. [CrossRef] [PubMed]
12. Poplin, R.; Ruano-Rubio, V.; DePristo, M.; Fennell, T.; Carneiro, M.; Van der Auwera, G.; Kling, D.; Gauthier, L.; Levy-Moonshine, A.; Roazen, D.; et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* **2017**. [CrossRef]
13. Srivastava, S.; Love-Nichols, J.A.; Dies, K.A.; Ledbetter, D.H.; Martin, C.L.; Chung, W.K.; Firth, H.V.; Frazier, T.; Hansen, R.L.; Prock, L.; et al. Meta-analysis and multidisciplinary consensus statement: Exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet. Med.* **2019**, *21*, 2413–2421. [CrossRef] [PubMed]
14. Nolan, D.; Carlson, M. Whole exome sequencing in pediatric neurology patients: Clinical implications and estimated cost analysis. *J. Child Neurol.* **2016**, *31*, 887–894. [CrossRef]
15. Miller, D.T.; Adam, M.P.; Aradhya, S.; Biesecker, L.G.; Brothman, A.R.; Carter, N.P.; Church, D.M.; Crolla, J.A.; Eichler, E.E.; Epstein, C.J.; et al. Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* **2010**, *86*, 749–764. [CrossRef] [PubMed]
16. Battaglia, A.; Doccini, V.; Bernardini, L.; Novelli, A.; Loddo, S.; Capalbo, A.; Filippi, T.; Carey, J.C. Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *Eur. J. Paediatr. Neurol.* **2013**, *17*, 589–599. [CrossRef] [PubMed]
17. Volkmar, F.R.; McPartland, J.C. From Kanner to DSM-5: Autism as an evolving diagnostic concept. *Annu. Rev. Clin. Psychol.* **2014**, *10*, 193–212. [CrossRef] [PubMed]
18. Ariani, F.; Hayek, G.; Rondinella, D.; Artuso, R.; Mencarelli, M.A.; Spanhol-Rosseto, A.; Pollazzon, M.; Buoni, S.; Spiga, O.; Ricciardi, S.; et al. FOXG1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **2008**, *83*, 89–93. [CrossRef]
19. Renieri, A.; Meloni, I.; Longo, I.; Ariani, F.; Mari, F.; Pescucci, C.; Cambi, F. Rett syndrome: The complex nature of a monogenic disease. *J. Mol. Med.* **2003**, *81*, 346–354. [CrossRef] [PubMed]
20. Amir, R.E.; Van Den Veyver, I.B.; Wan, M.; Tran, C.Q.; Francke, U. Rett syndrome is caused by mutations in X-linked. *Nat. Genet.* **1999**, *23*, 185–188. [CrossRef]
21. Scala, E.; Ariani, F.; Mari, F.; Caselli, R.; Pescucci, C.; Longo, I.; Meloni, I.; Giachino, D.; Bruttini, M.; Hayek, G.; et al. CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. *J. Med. Genet.* **2005**, *2*, 103–107. [CrossRef] [PubMed]
22. Snijders Blok, L.; Madsen, E.; Juusola, J.; Gilissen, C.; Baralle, D.; Reijnders, M.R.F.; Venselaar, H.; Helsmoortel, C.; Cho, M.T.; Hoischen, A.; et al. Mutations in DDX3X are a common cause of unexplained intellectual disability with gender-specific effects on wnt signaling. *Am. J. Hum. Genet.* **2015**, *97*, 343–352. [CrossRef]
23. Wang, X.; Posey, J.E.; Rosenfeld, J.A.; Bacino, C.A.; Scaglia, F.; Immken, L.D.; Harris, J.M.; Hickey, S.E.; Mosher, T.M.; Slavotinek, A.; et al. Phenotypic expansion in DDX3X—A common cause of intellectual disability in females. *Ann. Clin. Transl. Neurol.* **2018**, *15*, 1277–1285. [CrossRef]
24. Gilling, M.; Rasmussen, H.B.; Calloe, K.; Sequeira, A.F.; Baretto, M.; Oliveira, G.; Almeida, J.; Lauritsen, M.B.; Ullmann, R.; Boonen, S.E.; et al. Dysfunction of the heteromeric KV7.3/KV7.5 potassium channel is associated with autism spectrum disorders. *Front. Genet.* **2013**, *4*, 54. [CrossRef]
25. Nemani, T.; Steel, D.; Kaliakatsos, M.; DeVile, C.; Ververi, A.; Scott, R.; Getov, S.; Sudhakar, S.; Male, A.; Mankad, K.; et al. KIF1A-related disorders in children: A wide spectrum of central and peripheral nervous system involvement. *J. Peripher. Nerv. Syst.* **2020**, *25*, 117–124. [CrossRef]
26. Garde, A.; Cornaton, J.; Sorlin, A.; Moutton, S.; Nicolas, C.; Juif, C.; Geneviève, D.; Perrin, L.; Khau-Van-Kien, P.; Smol, T.; et al. Neuropsychological study in 19 French patients with White-Sutton syndrome and POGZ mutations. *Clin. Genet.* **2021**. [CrossRef]

27. Assia Batzir, N.; Posey, J.E.; Song, X.; Akdemir, Z.C.; Rosenfeld, J.A.; Brown, C.W.; Chen, E.; Holtrop, S.G.; Mizerik, E.; Nieto Moreno, M.; et al. Phenotypic expansion of POGZ-related intellectual disability syndrome (White-Sutton syndrome). *Am. J. Med. Genet. Part A* **2020**, *182*, 38–52. [CrossRef] [PubMed]
28. Precht, K.S.; Lese, C.M.; Spiro, R.P.; Huttenlocher, P.R.; Johnston, K.M.; Baker, J.C.; Christian, S.L.; Kittikamron, K.; Ledbetter, D.H. Two 22q telomere deletions serendipitously detected by FISH. *J. Med. Genet.* **1998**, *35*, 939–942. [CrossRef] [PubMed]
29. Durand, C.M.; Betancur, C.; Boeckers, T.M.; Bockmann, J.; Chaste, P.; Fauchereau, F.; Nygren, G.; Rastam, M.; Gillberg, I.C.; Anckarsäter, H.; et al. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat. Genet.* **2007**, *39*, 25–27. [CrossRef] [PubMed]
30. Prasad, C.; Prasad, A.N.; Chodirker, B.N.; Lee, C.; Dawson, A.K.; Jocelyn, L.J.; Chudley, A.E. Genetic evaluation of pervasive developmental disorders: The terminal 22q13 deletion syndrome may represent a recognizable phenotype. *Clin. Genet.* **2000**, *57*, 103–109. [CrossRef] [PubMed]
31. Powers, P.A.; Scherer, S.W.; Tsui, L.C.; Gregg, R.G.; Hogan, K. Localization of the Gene Encoding the α^2/δ Subunit (CACNL2A) of the Human Skeletal Muscle Voltage-Dependent Ca^{2+} Channel to Chromosome 7q21-q22 by Somatic Cell Hybrid Analysis. *Genomics* **1994**, *19*, 192–193. [CrossRef]
32. Vergult, S.; Dheedene, A.; Meurs, A.; Faes, F.; Isidor, B.; Janssens, S.; Gautier, A.; Le Caignec, C.; Menten, B. Genomic aberrations of the CACNA2D1 gene in three patients with epilepsy and intellectual disability. *Eur. J. Hum. Genet.* **2015**, *23*, 628–632. [CrossRef] [PubMed]
33. Bian, F.; Li, Z.; Offord, J.; Davis, M.D.; McCormick, J.; Taylor, C.P.; Walker, L.C. Calcium channel alpha2-delta type 1 subunit is the major binding protein for pregabalin in neocortex, hippocampus, amygdala, and spinal cord: An ex vivo autoradiographic study in alpha2-delta type 1 genetically modified mice. *Brain Res.* **2006**, *1075*, 68–80. [CrossRef] [PubMed]
34. Mori, M.; Sugo, T.; Abe, M.; Shimomura, Y.; Kurihara, M.; Kitada, C.; Kikuchi, K.; Shintani, Y.; Kurokawa, T.; Onda, H.; et al. Urotensin II is the endogenous ligand of a G-protein-coupled orphan receptor, SENR (GPR14). *Biochem. Biophys. Res. Commun.* **1999**, *137*, 579–588. [CrossRef] [PubMed]

2.2 New Candidates for Autism/Intellectual Disability Identified by Whole-Exome Sequencing

In this chapter, we present a study concerning the identification of four novel ID/ASD candidate genes. We performed parent–offspring trio WES in a cohort of 60 mostly syndromic ID/ASD patients and we detected 8 PVs in genes already known to be associated with ID/ASD (*SYNGAP1, SMAD6, PACS1, SHANK3, KMT2A, KCNQ2, ACTB,* and *POGZ*). We found four de novo disruptive variants in four novel candidate ASD/ID genes: *MBP, PCDHA1, PCDH15, PDPR*.

Finally, we selected via bioinformatic tools VUS in ID/ASD known genes and in novel genes that alone or in combination can contribute to the phenotypes.

In conclusion, we confirmed that WES offers expanded diagnostic options for patients with ID/ASD who resulted negative to standard genetic testing.

International Journal of
Molecular Sciences

MDPI

Article

New Candidates for Autism/Intellectual Disability Identified by Whole-Exome Sequencing

Lucia Pia Bruno [1,2,†], Gabriella Doddato [1,2,†], Floriana Valentino [1,2], Margherita Baldassarri [1,2], Rossella Tita [3], Chiara Fallerini [1,2], Mirella Bruttini [1,3], Caterina Lo Rizzo [3], Maria Antonietta Mencarelli [3], Francesca Mari [1,2,3], Anna Maria Pinto [3], Francesca Fava [1,2,3], Alessandra Fabbiani [1,2,3], Vittoria Lamacchia [1,2,3], Anna Carrer [1,2,3], Valentina Caputo [1,2,3], Stefania Granata [1,2,3], Elisa Benetti [2], Kristina Zguro [2], Simone Furini [2], Alessandra Renieri [1,2,3] and Francesca Ariani [1,2,3,*]

[1] Medical Genetics, University of Siena, 53100 Siena, Italy; lucia.bruno@dbm.unisi.it (L.P.B.); gabriella.doddato@dbm.unisi.it (G.D.); floriana.valentino@dbm.unisi.it (F.V.); margherita.baldassarri@dbm.unisi.it (M.B.); chiara.fallerini@dbm.unisi.it (C.F.); mirella.bruttini@dbm.unisi.it (M.B.); francesca.mari@dbm.unisi.it (F.M.); francesca.fava@dbm.unisi.it (F.F.); alessandra.fabbiani@dbm.unisi.it (A.F.); vittoria.lamacchia@dbm.unisi.it (V.L.); anna.carrer@dbm.unisi.it (A.C.); valentina.caputo@dbm.unisi.it (V.C.); stefania.granata@dbm.unisi.it (S.G.); alessandra.renieri@dbm.unisi.it (A.R.)

[2] Med Biotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, 53100 Siena, Italy; elisa.benetti@dbm.unisi.it (E.B.); kristina.zguro@student.unisi.it (K.Z.); simone.furini@dbm.unisi.it (S.F.)

[3] Genetica Medica, Azienda Ospedaliera Universitaria Senese, 53100 Siena, Italy; rossella.tita@dbm.unisi.it (R.T.); lorizzo2@unisi.it (C.L.R.); mariaantonietta.mencarelli@dbm.unisi.it (M.A.M.); annamaria.pinto@dbm.unisi.it (A.M.P.)

* Correspondence: francesca.ariani@unisi.it; Tel.: +39-0577-233303

† Co-first authors.

check for updates

Citation: Bruno, L.P.; Doddato, G.; Valentino, F.; Baldassarri, M.; Tita, R.; Fallerini, C.; Bruttini, M.; Lo Rizzo, C.; Mencarelli, M.A.; Mari, F.; et al. New Candidates for Autism/Intellectual Disability Identified by Whole-Exome Sequencing. *Int. J. Mol. Sci.* **2021**, *22*, 13439. https://doi.org/10.3390/ijms222413439

Academic Editor: Ivan Y. Iourov

Received: 29 November 2021
Accepted: 10 December 2021
Published: 14 December 2021

Abstract: Intellectual disability (ID) is characterized by impairments in the cognitive processes and in the tasks of daily life. It encompasses a clinically and genetically heterogeneous group of neurodevelopmental disorders often associated with autism spectrum disorder (ASD). Social and communication abilities are strongly compromised in ASD. The prevalence of ID/ASD is 1–3%, and approximately 30% of the patients remain without a molecular diagnosis. Considering the extreme genetic locus heterogeneity, next-generation sequencing approaches have provided powerful tools for candidate gene identification. Molecular diagnosis is crucial to improve outcome, prevent complications, and hopefully start a therapeutic approach. Here, we performed parent–offspring trio whole-exome sequencing (WES) in a cohort of 60 mostly syndromic ID/ASD patients and we detected 8 pathogenic variants in genes already known to be associated with ID/ASD (*SYNGAP1*, *SMAD6*, *PACS1*, *SHANK3*, *KMT2A*, *KCNQ2*, *ACTB*, and *POGZ)*. We found four de novo disruptive variants of four novel candidate ASD/ID genes: *MBP*, *PCDHA1*, *PCDH15*, *PDPR*. We additionally selected via bioinformatic tools many variants in unknown genes that alone or in combination can contribute to the phenotype. In conclusion, our data confirm the efficacy of WES in detecting pathogenic variants of known and novel ID/ASD genes.

Keywords: autism; intellectual disability; whole-exome sequencing

1. Introduction

Intellectual disability (ID), previously known as "mental retardation", represents a major public health problem [1]. ID is a condition characterized by below-average intellectual functioning (IQ < 70) together with significant limitations in adaptive functioning [2]. ID can be "isolated" or "syndromic" when peculiar facies, typical clinical traits, and/or growth abnormalities are documented [3,4]. Often, a coexistence of ID and autism spectrum disorder (ASD) arises, with 70% of ASD patients also showing ID, and 40% of ID patients

displaying ASD [4,5]. ASD encompasses neurodevelopmental conditions characterized by deficient social interactions, poor or absent communication, repetitive behaviors, and apparently limited interests [3,6]. ASD and ID affect about 1–3% of the general population [7,8]. Genetic factors, fetal intrauterine environment, and external environmental factors play important roles in ID and ASD [9,10].

The genetic basis of ID and ASD is deeply heterogeneous, implicating more than 2000 OMIM genes, in turn, involved in different pathways and biological processes, such as those regulating synaptic plasticity, chromatin remodeling, gene transcription, and protein degradation [11,12]. Combining clinical and molecular diagnosis is fundamental to deepen our knowledge of the pathogenic mechanisms underlying these medical conditions and to develop personalized treatments [13]. Single-nucleotide variants (SNVs), indels, and copy number variations (CNVs) have been identified as key variant types causing ID/ASD [7,14].

Up to 50% of the ID/ASD patients remain without a molecular diagnosis [15]. Next-generation sequencing (NGS) technologies have greatly improved the chance of identifying known as well as novel responsible genes [15–17]. The use of Whole-Exome Sequencing (WES) together with CNVs analysis can identify a pathogenic variant in about 30% of patients [17]. The current diagnostic yields suggest the application of WES as a routine first-tier diagnostic test permitting an early diagnosis for ID/ASD patients, with improvement in terms of the quality of life of the affected families [18,19].

In the present study, WES was performed for a total of 60 trios with a diagnosis of an ASD/ID-related phenotype. We identified pathogenic variants in already known ID/ASD genes in eight families. We found four new candidate genes with de novo truncating variants (three frameshift deletions/duplications and one nonsense variant), i.e., *MBP*, *PCDHA1*, *PCDH15*, *PDPR*. Among the remaining cases, we selected missense variants for which bioinformatic tools suggested pathogenicity. We found that these unknown rare variants, alone or in combination with each other, contributed to the phenotype. In conclusion, our data confirm the efficacy of WES in detecting pathogenic variants in known and novel ID/ASD genes

2. Results

2.1. Study Cohort

In the present study, we enrolled 60 proband–parent trios with a diagnosis of ASD/ID-related phenotype. Probands were a total of 15 females and 45 males with a mean age of 10 years. Patients diagnosed with ID were 19, one patient had high-functioning ASD, and 40 patients had ASD and ID. Of the 60 patients, 12% presented epilepsy, 80% language delay, 66% psychomotor delay, and 73% craniofacial dysmorphisms. The clinical description of the patients is reported in Table S1.

2.2. WES and Variants Pathogenicity Assessment

WES was performed for 60 proband–parent trios, and a mean coverage of 94 reads for targeted sequenced regions was obtained. Patients were previously analyzed by array-CGH and were all negative. Pathogenic variants were selected according to variant frequency and category, co-segregation with the disease, literature data, and database classification (ClinVar database).

We selected novel candidate genes starting from truncating variants that could be assumed to disrupt gene function (Table 1). We found four cases with de novo truncating variants (three frameshift deletions/duplications and one nonsense variant) in genes not traditionally associated with ID/ASD: *MBP*, *PCDHA1*, *PCDH15*, and *PDPR* (Table 1) (ST2). For missense variants CADD phred values higher than 20 suggested pathogenicity (Table S3). We also performed structural predictions through the HOPE tool (Table S2). Potential splicing impact was predicted for the variants in the unknown genes *TYRO3*, *OPN4*, *CBX3*. We found eight pathogenic variants (two frameshift deletions/duplications, three nonsense, and three reported missense variants) in genes already known to be associated

with ID/ASD: *SYNGAP1*, *SMAD6*, *PACS1*, *SHANK3*, *KMT2A*, *KCNQ2*, *ACTB*, and *POGZ* (Table). Some variants were reported in the database (ClinVar database) or previously described in the scientific literature (Tables 1–3). VUS were detected in 28 novel genes and in 10 known ID/ASD genes (Table S3). VUS were found alone or in combination with each other (Table S3).

2.3. Clinical Features of Patients with Truncating Variants of ID/ASD Candidate Genes

Four patients showed rare de novo truncating variants of the novel candidate genes *MBP*, *PCDHA1*, *PCDH15*, *PDPR*. The clinical findings of each patient are described in Tables 3 and S1. Patient I was a 15-year-old female child with a diagnosis of ID. Pregnancy was complicated by intrauterine growth retardation. However, her karyotype was normal. She was born at 38 weeks of gestation. The growth parameters at birth were: length of 46 cm (3–10° percentile), weight of 2300 g (10–25° percentile), and Occipital Frontal Circumference (OFC) of 34 cm (50° percentile). At birth, she was diagnosed with interatrial septal defect (DIA) and interventricular septal defect (DIV). Since the first months of life, she suffered from esophageal reflux and feeding difficulties. She started independent walking at the age of 19 months, but her speech development was severely delayed. Sphincter control was acquired at 3 years. Recurrent airway infections emerged during childhood. Physical examination at the age of 13 years showed height of 144.5 cm (3–10° percentile), weight of 35 kg (3–10° percentile), and head circumference of 50.5 cm (−2.4 SD). She presented a triangular face with low anterior hairline, broad nasal bridge with bulbous nasal tip, M-shaped upper lip, and everted lower lip. Arachnodactyly of the hands and feet was noticed (Figure 1A,B). She was attending school with support and had poor reading and writing skills. She received physiotherapy and psychotherapy and showed good interaction with peers. Parents reported hand stereotypies, episodes of unprovoked laughter, hyperactivity, and sleep disorder. Her array CGH analysis did not reveal any chromosomal aberrations. WES analysis was thus performed, and we detected a de novo truncating variant c.138del (p.(Phe46fs*18)) of the *MBP* gene.

Patient II, a girl (14 years of age) diagnosed with ID and affected by epilepsy, carried a frameshift de novo variant c.5573_5576 (p.(Lys1859Asnfs * 2)) of *PCDH15* and a de novo splicing variant c.1074–11G > A of *OPN4*. She was attending school with a support teacher. She had facial dysmorphisms such as a square-shaped face, deeply set eyes, bilateral underfolded helix, and short and stocky neck (Figure 1C,D). She was born at term after a pregnancy with gestational diabetes. Parameters at birth were: length, 48 cm (25–50° percentile), weight, 2930 g (10–25° percentile), and APGAR, 9–10. No suction difficulties were observed. At 16 months of age, she started to present seizures for which she is still on therapy. She acquired autonomous deambulation at 18 months, and language and psychomotor delay were reported.

Patient III was a 9-year-old male presenting with ADHD (Attention-Deficit/Hyperactivity Disorder) and behavior disorder. Both parents are healthy, without a family history of neurodevelopmental disorder. He was diagnosed with ID and ASD. He was born at term, without any problems during pregnancy. He started walking and pronounced the first words at 1 year of age. Cerebral Magnetic Resonance Imaging (MRI) revealed reduced white matter, hypoplasia of vermis and trunk, tortuous optic nerves and vertebral arteries. EEG showed left-sided centro-temporal epileptic elements. Physical examination at the age of 9 years showed height of 136 cm (66° percentile), weight of 33 kg (72° percentile), and OFC of 54 cm (86° percentile). He presented simplified auricles and no further dysmorphic signs. Via WES we found a de novo frameshift mutation c.670_673dup (p.(Thr225Argfs*4)) in the *PCDHA1* gene.

Table 1. Molecular information for the four truncating variants of novel ID/ASD candidate genes.

Proband	Gene	Transcript (hg19)	Variant (HGVS)	Protein (HGVS)	MAF (gnomAD All)	MAF (gnomAD NFE)	dbSNP	ClinVar Classification	CADD	Transmission	Origin	Classification
I	*MBP*	NM_001025081.1	c.138del	p.(Phe46Leufs * 18)	NA	NA	NA	NA	NA	Autosomal dominant	De novo	Pathogenic
II	*PCDH15*	NM_033056.3	c.5573_5576dup	p.(Lys1859Asnfs * 2)	0.0068%	0.011%	rs770082088	NA	NA	Autosomal dominant	De novo	Pathogenic
III	*PCDHA1*	NM_018900.3	c.670_673dup	p.(Thr225Argfs * 4)	NA	NA	NA	NA	NA	Autosomal dominant	De novo	Pathogenic
IV	*PDPR*	NM_001322118.1	c.826C > T	p.(Gln276*)	0.00071%	0.0016%	NA	NA	22.7	Autosomal dominant	De novo	Pathogenic

* means change in a stop codon.

Table 2. Molecular information for the pathogenic variants of genes already known to be associated with ID/ASD.

Proband	Gene	Transcript (hg19)	Variant (HGVS)	Protein (HGVS)	MAF (gnomAD All)	MAF (gnomADNFE)	dbSNP	ClinVar Classification	CADD	Transmission	Origin	Classification	Reference
VIII	*ACTB*	NM_001101.4	c.583G > A	p.(Glu195Lys)	NA	NA	NA	Likely pathogenic	37	Autosomal dominant	De novo	Likely pathogenic	NA
IX	*KCNQ2*	NM_172107.2	c.628C > T	p.(Arg210Cys)	NA	NA	rs796052626	Pathogenic	27.3	Autosomal dominant	De novo	Pathogenic	[20]
X	*KMT2A*	NM_001197104.1	c.478C > T	p.(Arg160*)	NA	NA	NA	NA	36	Autosomal dominant	De novo	Pathogenic	NA
XI	*PACS1*	NM_018026.3	c.607C > T	p.(Arg203Trp)	0%	NA	rs398123009	Pathogenic	29.4	Autosomal dominant	De novo	Pathogenic	[21]
XII	*POGZ*	NM_145796.3	c.2716C > T	p.(Arg906 *)	NA	NA	rs869312833	Pathogenic	12.48	Autosomal dominant	De novo	Pathogenic	[22]
XIII	*SHANK3*	NM_001080420.1	c.1807_1811del	p.(Val604Leufs * 80)	NA	NA	NA	NA	NA	Autosomal dominant	De novo	Pathogenic	NA
XIV	*SMAD6*	NM_005585.4	c.137dup	p.(Tyr459Leufs * 106)	NA	NA	NA	NA	NA	Autosomal dominant	De novo	Pathogenic	NA
XV	*SYNGAP1*	NM_001130066.1	c.3670C > T	p.(Arg1224 *)	NA	NA	rs869312955	Pathogenic	36	Autosomal dominant	De novo	Pathogenic	[23]

Table 3. Genotype–phenotype correlations for the candidate genes.

Proband	Gene	Variant (HGVS)	Protein (HGVS)	A Gender	Age (Years Old)	ID/ASD	Craniofacial Dysmorphisms	Additional Clinical Signs
I	*MBP*	c.138del	p.(Phe46Leufs*18)	F	15	ID	Triangular facies, prominent ears, thin upper lip, absent eyebrows, broad nasal bridge, bulbous nasal tip, thin and sparse hair, everted lower lip, M-shaped upper lip, hairline anteriorly advanced.	Hyperactivity, language delay and aggressiveness, disturbed wake–sleep cycle, arachnodactyly of the hand and feet.
II	*PCDH15*	c.5573_5576dup	p.(Lys1859Asnfs * 2)	F	14	ID	Square-shaped face, deeply set eyes, bilateral underfolded helix, short and stocky neck.	Epilepsy, language and psychomotor delay.
III	*PCDHA1*	c.670_673dup	p.(Thr225Argfs * 4)	M	9	ASD and ID	Simplified auricles	ADHD, vermis and brain stem hypoplasia, tortuous course of the optic nerves, flat feet, hyperlaxity, psychomotor and language delay.
IV	*PDPR*	c.826C > T	p.(Gln276*)	M	11	ID	Deep-set eyes, wide nasal tip, thin upper lip, chin dimple, and macrodontia.	Cognitive impairment, repetitive behaviors, an altered sleep pattern with difficulty in falling asleep, isolationist tendencies, manual stereotypies, hypochromic stains, food selectiveness, language and psychomotor delay.

* means change in a stop codon.

We visited patient IV when he was 8 years and 10 months old. He was born at term by cesarean section. During the pregnancy, her mother suffered from a cytomegalovirus infection. His perinatal and postnatal period were normal. He pronounced his first words at the age of 12 months and walked independently at 13 months. Sphincter control was acquired at 38 months. He presented difficulty in social interactions, with a tendency to isolation. The patient showed difficulties in acquiring reading–writing skills and received psychomotor and speech therapy. He received a diagnosis of ID. Clinical examination revealed height of 125 cm (10–25° percentile), weight of 31 kg (50–75° percentile), and head circumference of 54 cm (75–90° percentile). His dysmorphic facial features included deeply set eyes, wide nasal tip, thin upper lip, chin dimple, and macrodontia (Figure 1E,F). Furthermore, hypochromic spots were observed on the back and upper limbs. A de novo stop-gain variant c.826C > T (p.(Gln276*)) of *PDPR* was revealed by WES.

Figure 1. *Cont.*

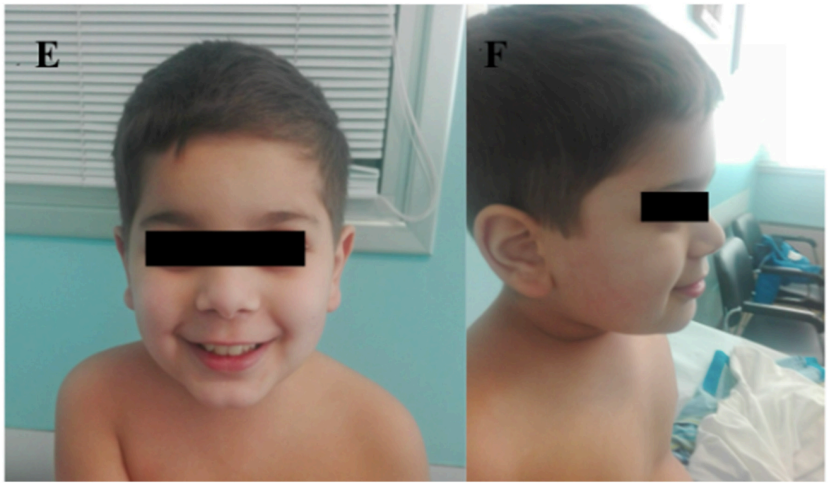


Figure 1. (**A**) Frontal view of proband I (*MBP* gene; p.(Phe46fs*18)) showing triangular facies, prominent ears, thin upper lip, absent eyebrows, broad nasal bridge, bulbous nasal tip, thin and sparse hair, everted lower lip, advanced hairline; (**B**) close-up of the arachnodactyly of the hand of proband I; (**C**) frontal view of proband II (*PCDH15* gene; p.(Lys1859Asnfs*)) displaying square-shaped face, deep-set eyes, bilateral underfolded helix, short and stocky neck; (**D**) lateral view of proband II; (**E**) frontal view of proband IV (*PDPR* gene; p.(Gln276*)) showing big and deep-set eyes, wide nasal tip, thin upper lip, chin dimple and macrodontia; (**F**) lateral view of proband IV.

3. Discussion

Identifying the etiology of ASD and ID poses an arduous challenge due to a relevant clinical heterogeneity and a high genetic heterogeneity of these medical conditions, even across single families [23,24]. WES demonstrated to effectively detect novel candidate genes potentially associated with these neurodevelopmental disorders [3]. In the present study, we performed WES in a cohort of 60 trios with a diagnosis of ASD/ID-related phenotype and we detected 8 pathogenic variants in genes already known to be associated with ID/ASD (*SYNGAP1*, *SMAD6*, *PACS1*, *SHANK3*, *KMT2A*, *KCNQ2*, *ACTB*, and *POGZ)*. We also identified four de novo disrupting variants in the following candidate genes: *MBP, PCDHA1, PCDH15,* and *PDPR*.

MBP encodes the principal constituent of myelin, needed for the maintenance of the compact multilamellar membrane structure of its mature form [25,26]. *MBP* is expressed in oligodendrocytes in the central nervous system and in Schwann cells in the peripheral nervous system [26]. Mice with mutations in the *MBP* gene developed a decreased myelination of the central nervous system, with tremors and convulsions, progressively leading to an early death [27]. A caudal-to-rostral gradient of transcription for *MBP* has been observed in the developing human brain, reflecting the process of myelination [28]. A variant in the *MBP* gene was found in patient I, who presented hand stereotypies, hyperactivity, sleep disorder, and speech delay.

The *PCDHA* gene cluster consists of 14 tandemly arranged genes that are expressed in the vertebrate brain and encode diverse membrane proteins involved in axonal projection, learning, and memory [29,30].

PCDHA1 was found mutated in a patient with ADHD, language and psychomotor delay, simplified auricles, and abnormal cerebral structures. Among the protocadherin superfamily members, *PCDH15* pathogenic variants cause Usher syndrome following an autosomal recessive inheritance pattern [31–33]. Although primarily recognized as a disease associated with deafness and blindness, more than 20% of Usher syndrome patients

display psychiatric symptoms, and comorbidities of Usher syndrome with various mental illnesses are well documented [34–37]. Moreover, rare heterozygous SNVs in PCDH15 were detected in a patient with schizophrenia and ASD [38]. Importantly, a 386 kb deletion in 10q21.1, including the first three exons of *PCDH15*, was reported in a subject with ASD [39]. We found a *PCDH15* frameshift variant in a patient with ID, epilepsy, and language, and psychomotor delay and we hypothesized that, in part, it might contribute to the phenotype.

The pyruvate dehydrogenase phosphatase regulatory subunit (*PDPR*) is variably expressed in the brain, with the highest levels in the corpus callosum and in the cerebellum [40,41]. *PDPR* has been recently proposed as a novel candidate gene for ID. A missense variant in homozygosity, in fact, was found in a patient with global developmental delay, Joubert-like symptoms, and MRI findings [42]. Here, we propose that the de novo *PDPR* stop-gain variant might contribute to the language delay and behavior disorder observed in patient IV.

Finally, we found VUS in ID/ASD known genes (*BRWD3*, *CNOT1*, *DNMT3A*, *FGF13*, *HUWE1*, *KMT2B*, *NLGN4X*, *PHF8*, *TAF1*, *DDX3X*) and in novel genes (*RGPD4*, *RIN3*, *SBNO2*, *RBMS1*, *SORBS1*, *CHRFAM7A*, *RSF1*, *AGPAT5*, *GALR3*, *NRXN2*, *TYRO3*, *MAP3K10*, *SLC7A8*, *LONP1*, *CBX3*, *DCAF11*, *GINS2*, *SLC8A1*, *CDC7*, *IPO7*, *POTEH*, *MAP3K5*, *SORBS1*, *POU3F2*, *KANSL3*, *IQSEC3*). These variants have been found alone or in combination. The association between de novo SNVs and ASD/ID has been demonstrated by different studies, firstly for disruptive variants [43]. Recently, an enrichment of multiple de novo variants in various genes found via large-scale exome sequencing has suggested the involvement of an oligogenic model in patients with ASD [44]. Our findings are in agreement with the theory of a multi-hit model in the pathogenesis of ASD and ID [45].

In conclusion, we confirmed that WES offers expanded diagnostic options for patients with ID/ASD who resulted negative to array-CGH analysis. However, these genes are candidates, and additional studies are required to assess their role in neurodevelopment, to determine if they are related to abnormal phenotypes according to a monogenic model or a polygenic model and if modifier genes and environmental factors can alter the associated phenotypes [46].

4. Materials and Methods

4.1. Selection of Patients and DNA Samples Preparation

Genetic counseling was carried out to evaluate each patient's personal and familial history. Parents provided and signed a written informed consent at the Medical Genetics department of the University of Siena for exome sequencing analysis, clinical data usage, and the use of DNA samples from the tested individuals for both research and diagnosis purposes. We analyzed 60 patients affected by ID and ID/ASD (19 with ID, 40 with ID and ASD, and 1 with isolated ASD) recruited from January 2019 until the end of August 2021.

Genomic DNA from the probands and parents was isolated from EDTA peripheral blood samples using MagCore HF16 (Diatech Lab Line, Jesi, Ancona, Italy) according to the manufacturer's instructions.

4.2. Whole-Exome Sequencing

Sample preparation was performed following the "Illumina DNA prep with enrichment" manufacturer's protocol. This protocol involves the use of a bead-based transposome complex to perform tagmentation, a process that fragments the genomic DNA and then tags it with adapter sequences in one step. After saturation with input DNA, the bead-based transposome complex fragments a set number of DNA molecules. This fragmentation provides flexibility to use a wide DNA input range to generate normalized libraries with a consistent tight fragment size distribution. Then, a limited-cycle PCR adds adapter sequences to the ends of a DNA fragment. A subsequent target enrichment workflow is applied. Following pooling, the double-stranded DNA libraries are denatured and biotinylated. Illumina Exome Panel v1.2 (CEX) probes are hybridized to the denatured library fragments. Then Streptavidin Magnetic Beads (SMB) capture the targeted library frag-

ments within the regions of interest. The indexed libraries are eluted from the beads and further amplified before sequencing. Whole-exome sequencing analysis was performed on the Illumina NovaSeq6000 System (Illumina San Diego, CA, USA) according to the NovaSeq6000 System Guide. Reads were mapped against the hg19 reference genome using the Burrow–Wheeler aligner (BWA) [47]. Variant calling was obtained using an in-house pipeline which takes advantage of the GATK Best Practices workflow [48].

4.3. Filtering and Variant Prioritization

All variants were screened according to frequency, location, mutation category, literature, and mutation database data (ClinVar database, LOVD database, HGMD database). Polymorphisms (minor allele frequency, MAF < 0.01) were excluded, and synonymous variants were assumed to be benign or likely benign. Missense variants were predicted to be damaging by CADD-Phred prediction tools for functional effect prediction. Frameshift, stop-gain, and splice site variants were prioritized as pathogenic. A prediction of damage for the unreported missense variants in the new candidate genes came from HOPE and Phyre2 tools. The potential impact of the variants on splicing was evaluated using Alamut® Visual software—version 2.11-0 (Interac-tive Biosoftware, Rouen, France), which employs five different algorithms: SpliceSiteFinder-like, MaxEntScan, NNSPLICE, GeneSplicer, and HumanSplicingFinder.

The following public databases were used for the interpretation of the variants: ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/, accessed 26 November 2021), LOVD (https://databases.lovd.nl/shared/genes, accessed 26 November 2021), Human Genome Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/index.php3, accessed 26 November 2021).

4.4. Sanger Sequencing

Pathogenic variants in new candidate genes were confirmed by Sanger sequencing. DNA samples were sequenced using the PE Big Dye Terminator Cycle Sequencing Kit on an ABI Prism 3130 analyzer (Applied Biosystems). The data were analyzed using the Sequencher version 4.9 software.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms222413439/s1.

Author Contributions: L.P.B. and G.D. made important contributions in the interpretation of the molecular results and drafted the manuscript. F.V., R.T., C.F., M.B. (Mirella Bruttini), E.B., K.Z. and S.F. performed the experiments and data acquisition. M.B. (Margherita Baldassarri), C.L.R., M.A.M., F.M., A.M.P., F.F., A.F., V.L., A.C., V.C., S.G. conducted genetic counselling to the family. A.R. and F.A. made substantial contributions to the conception and design of the study and reviewed the manuscript. All authors approved the final version of the manuscript for publication.

Funding: This research received no external fundings.

Institutional Review Board Statement: Ethical review and approval were not required for the study on human participants in accordance with local legislation and institutional requirements.

Informed Consent Statement: The patients/participants provided their written informed consent to participate in this study.

Data Availability Statement: NGS data has been deposited in publicly accessible repositories. The data can be found here: http://nigdb.cineca.it/.

Acknowledgments: We are grateful to our patients for their cooperation. The "Cell lines and DNA bank of Rett Syndrome, X-linked mental retardation, and other genetic diseases", member of the Telethon Network of Genetic Biobanks (project nos. GTB12001 and GFB18001), funded by Telethon Italy, and of the EuroBioBank network provided us with specimens.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. May, M.E.; Kennedy, C.H. Health and Problem Behavior Among People with Intellectual Disabilities. *Behav. Anal. Pract.* **2010**, *3*, 4–12. [CrossRef]
2. Bishop, S.L.; Farmer, C.; Thurm, A. Measurement of Nonverbal IQ in Autism Spectrum Disorder: Scores in Young Adulthood Compared to Early Childhood. *J. Autism Dev. Disord.* **2015**, *45*, 966–974. [CrossRef] [PubMed]
3. Chiurazzi, P.; Kiani, A.K.; Miertus, J.; Barati, S.; Manara, E.; Paolacci, S.; Stuppia, L.; Gurrieri, F.; Bertelli, M. Genetic Analysis of Intellectual Disability and Autism. *Acta Biomed.* **2020**, *9*, 91. [CrossRef]
4. Srivastava, A.K.; Schwartz, C.E. Intellectual Disability and Autism Spectrum Disorders: Causal Genes and Molecular Mechanisms. *Neurosci. Biobehav. Rev.* **2014**, *46*, 161–174. [CrossRef] [PubMed]
5. la Malfa, G.; Lassi, S.; Bertelli, M.; Salvini, R.; Placidi, G.F. Autism and Intellectual Disability: A Study of Prevalence on a Sample of the Italian Population. *J. Intellect. Disabil. Res.* **2004**, *48*, 262–267. [CrossRef] [PubMed]
6. Hodges, H.; Fealko, C.; Soares, N. Autism Spectrum Disorder: Definition, Epidemiology, Causes, and Clinical Evaluation. *Transl. Pediatrics* **2020**, *9*, S55. [CrossRef] [PubMed]
7. Lai, M.C.; Lombardo, M.V.; Baron-Cohen, S. Autism. *Lancet* **2014**, *383*, 896–910. [CrossRef]
8. Lord, C.; Elsabbagh, M.; Baird, G.; Veenstra-Vanderweele, J. Autism Spectrum Disorder. *Lancet* **2018**, *392*, 508–520. [CrossRef]
9. Moeschler, J.B.; Shevell, M.; Saul, R.A.; Chen, E.; Freedenberg, D.L.; Hamid, R.; Jones, M.C.; Stoler, J.M.; Tarini, B.A. Comprehensive Evaluation of the Child with Intellectual Disability or Global Developmental Delays. *Pediatrics* **2014**, *134*, e903–e918. [CrossRef] [PubMed]
10. Chaste, P.; Leboyer, M. Autism Risk Factors: Genes, Environment, and Gene-Environment Interactions. *Dialogues Clin. Neurosci.* **2012**, *14*, 281. [CrossRef] [PubMed]
11. Bourgeron, T. From the Genetic Architecture to Synaptic Plasticity in Autism Spectrum Disorder. *Nat. Rev. Neurosci.* **2015**, *16*, 551–563. [CrossRef] [PubMed]
12. Dulac, C. Brain Function and Chromatin Plasticity. *Nature* **2010**, *465*, 728–735. [CrossRef]
13. Savatt, J.M.; Myers, S.M. Genetic Testing in Neurodevelopmental Disorders. *Front. Pediatrics* **2021**, *9*, 52. [CrossRef] [PubMed]
14. Rylaarsdam, L.; Guemez-Gamboa, A. Genetic Causes and Modifiers of Autism Spectrum Disorder. *Front. Cell. Neurosci.* **2019**, *13*, 385. [CrossRef] [PubMed]
15. Bruel, A.L.; Vitobello, A.; Tran Mau-Them, F.; Nambot, S.; Sorlin, A.; Denommé-Pichon, A.S.; Delanne, J.; Moutton, S.; Callier, P.; Duffourd, Y.; et al. Next-Generation Sequencing Approaches and Challenges in the Diagnosis of Developmental Anomalies and Intellectual Disability. *Clin. Genet.* **2020**, *98*, 433–444. [CrossRef]
16. Harripaul, R.; Noor, A.; Ayub, M.; Vincent, J.B. The Use of Next-Generation Sequencing for Research and Diagnostics for Intellectual Disability. *Cold Spring Harb. Perspect. Med.* **2017**, *7*, a026864. [CrossRef] [PubMed]
17. Valentino, F.; Bruno, L.P.; Doddato, G.; Giliberti, A.; Tita, R.; Resciniti, S.; Fallerini, C.; Bruttini, M.; lo Rizzo, C.; Mencarelli, M.A.; et al. Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations. *Brain Sci.* **2021**, *11*, 936. [CrossRef]
18. Munnich, A.; Demily, C.; Frugère, L.; Duwime, C.; Malan, V.; Barcia, G.; Vidal, C.; Throo, E.; Besmond, C.; Hubert, L.; et al. Impact of On-Site Clinical Genetics Consultations on Diagnostic Rate in Children and Young Adults with Autism Spectrum Disorder. *Mol. Autism* **2019**, *10*, 1–10. [CrossRef] [PubMed]
19. Ontario Health (Quality). Genome-Wide Sequencing for Unexplained Developmental Disabilities or Multiple Congenital Anomalies: A Health Technology Assessment. *Ont. Health Technol. Assess. Ser.* **2020**, *20*, 1–178.
20. Hewson, S.; Puka, K.; Mercimek-Mahmutoglu, S. Variable expressivity of a likely pathogenic variant in KCNQ2 in a three-generation pedigree presenting with intellectual disability with childhood onset seizures. *Am. J. Med. Genet. A* **2017**, *173*, 2226–2230. [CrossRef]
21. Stessman, H.A.F.; Willemsen, M.H.; Fenckova, M.; Penn, O.; Hoischen, A.; Xiong, B.; Wang, T.; Hoekzema, K.; Vives, L.; Vogel, I.; et al. Disruption of POGZ Is Associated with Intellectual Disability and Autism Spectrum Disorders. *Am. J. Hum. Genet.* **2016**, *98*, 541–552. [CrossRef]
22. Helbig, K.L.; Farwell Hagman, K.D.; Shinde, D.N.; Mroske, C.; Powis, Z.; Li, S.; Tang, S.; Helbig, I. Diagnostic Exome Sequencing Provides a Molecular Diagnosis for a Significant Proportion of Patients with Epilepsy. *Genet. Med.* **2016**, *18*, 898–905. [CrossRef]
23. Masi, A.; DeMayo, M.M.; Glozier, N.; Guastella, A.J. An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. *Neurosci. Bull.* **2017**, *33*, 183–193. [CrossRef] [PubMed]
24. Lenroot, R.K.; Yeung, P.K. Heterogeneity within Autism Spectrum Disorders: What Have We Learned from Neuroimaging Studies? *Front. Hum. Neurosci.* **2013**, *7*, 733. [CrossRef] [PubMed]
25. Barbarese, E.; Barry, C.; Chou, C.J.; Goldstein, D.J.; Nakos, G.A.; Hyde-DeRuyscher, R.; Scheld, K.; Carson, J.H. Expression and Localization of Myelin Basic Protein in Oligodendrocytes and Transfected Fibroblasts. *J. Neurochem.* **1988**, *51*, 1737–1745. [CrossRef] [PubMed]
26. Haas, S.; Thatikunta, P.; Steplewski, A.; Johnson, E.M.; Khalili, K.; Amini, S. A 39-KD DNA-Binding Protein from Mouse Brain Stimulates Transcription of Myelin Basic Protein Gene in Oligodendrocytic Cells. *J. Cell Biol.* **1995**, *130*, 1171–1179. [CrossRef]
27. Roach, A.; Boylan, K.; Horvath, S.; Prusiner, S.B.; Hood, L.E. Characterization of Cloned CDNA Representing Rat Myelin Basic Protein: Absence of Expression in Brain of Shiverer Mutant Mice. *Cell* **1983**, *34*, 799–806. [CrossRef]

28. Kamholz, J.; Toffenetti, J.; Lazzarini, R.A. Organization and Expression of the Human Myelin Basic Protein Gene. *J. Neurosci. Res.* **1988**, *21*, 62–70. [CrossRef] [PubMed]
29. Ribich, S.; Tasic, B.; Maniatis, T. Identification of Long-Range Regulatory Elements in the Protocadherin-α Gene Cluster. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19719–19724. [CrossRef] [PubMed]
30. Hasegawa, S.; Hamada, S.; Kumode, Y.; Esumi, S.; Katori, S.; Fukuda, E.; Uchiyama, Y.; Hirabayashi, T.; Mombaerts, P.; Yagi, T. The Protocadherin-α Family Is Involved in Axonal Coalescence of Olfactory Sensory Neurons into Glomeruli of the Olfactory Bulb in Mouse. *Mol. Cell. Neurosci.* **2008**, *38*, 66–79. [CrossRef] [PubMed]
31. Fukuda, E.; Hamada, S.; Hasegawa, S.; Katori, S.; Sanbo, M.; Miyakawa, T.; Yamamoto, T.; Yamamoto, H.; Hirabayashi, T.; Yagi, T. Down-Regulation of Protocadherin-α A Isoforms in Mice Changes Contextual Fear Conditioning and Spatial Working Memory. *Eur. J. Neurosci.* **2008**, *28*, 1362–1376. [CrossRef] [PubMed]
32. Ahmed, Z.M.; Riazuddin, S.; Bernstein, S.L.; Ahmed, Z.; Khan, S.; Griffith, A.J.; Morell, R.J.; Friedman, T.B.; Riazuddin, S.; Wilcox, E.R. Mutations of the Protocadherin Gene PCDH15 Cause Usher Syndrome Type 1F. *Am. J. Hum. Genet.* **2001**, *69*, 25–34. [CrossRef] [PubMed]
33. Rebibo-Sabbah, A.; Nudelman, I.; Ahmed, Z.M.; Baasov, T.; Ben-Yosef, T. In Vitro and Ex Vivo Suppression by Aminoglycosides of PCDH15 Nonsense Mutations Underlying Type 1 Usher Syndrome. *Hum. Genet.* **2007**, *122*, 373–381. [CrossRef] [PubMed]
34. Ahmed, Z.M.; Riazuddin, S.; Aye, S.; Ali, R.A.; Venselaar, H.; Anwar, S.; Belyantseva, P.P.; Qasim, M.; Riazuddin, S.; Friedman, T.B. Gene Structure and Mutant Alleles of PCDH15: Nonsyndromic Deafness DFNB23 and Type 1 Usher Syndrome. *Hum. Genet.* **2008**, *124*, 215–223. [CrossRef] [PubMed]
35. Carvill, S. Sensory Impairments, Intellectual Disability and Psychiatry. *J. Intellect. Disabil. Res.* **2001**, *45*, 467–483. [CrossRef] [PubMed]
36. Domanico, D.; Fragiotta, S.; Trabucco, P.; Nebbioso, M.; Vingolo, E.M. Genetic Analysis for Two Italian Siblings with Usher Syndrome and Schizophrenia. *Case Rep. Ophthalmol. Med.* **2012**, *2012*, 380863. [CrossRef] [PubMed]
37. Rao, N.P.; DaNivas, V.; Venkatasubramanian, G.; Behere, R.V.; Gangadhar, B.N. Comorbid Bipolar Disorder and Usher Syndrome. *Prim. Care Companion J. Clin. Psychiatry* **2010**, *12*, PCC.09l00792. [CrossRef]
38. Rijavec, N.; Grubic, V.N. Usher Syndrome and Psychiatric Symptoms: A Challenge in Psychiatric Management. *Psychiatr. Danub.* **2009**, *21*, 68–71.
39. Ohara, O.; Nagase, T.; Mitsui, G.; Kohga, H.; Kikuno, R.; Hiraoka, S.; Takahashi, Y.; Kitajima, S.; Saga, Y.; Koseki, H. Characterization of Size-Fractionated CDNA Libraries Generated by the in Vitro Recombination-Assisted Method. *DNA Res.* **2002**, *9*, 47–57. [CrossRef]
40. Lawson, J.E.; Park, S.H.; Mattison, A.R.; Yan, J.; Reed, L.J. Cloning, Expression, and Properties of the Regulatory Subunit of Bovine Pyruvate Dehydrogenase Phosphatase. *J. Biol. Chem.* **1997**, *272*, 31625–31629. [CrossRef] [PubMed]
41. Alazami, A.M.; Patel, N.; Shamseldin, H.E.; Anazi, S.; Al-Dosari, M.S.; Alzahrani, F.; Hijazi, H.; Alshammari, M.; Aldahmesh, M.A.; Salih, M.A.; et al. Accelerating Novel Candidate Gene Discovery in Neurogenetic Disorders via Whole-Exome Sequencing of Prescreened Multiplex Consanguineous Families. *Cell Rep.* **2015**, *10*, 148–161. [CrossRef]
42. Sanders, S.J.; Murtha, M.T.; Gupta, A.R.; Murdoch, J.D.; Raubeson, M.J.; Willsey, A.J.; Ercan-Sencicek, A.G.; di Lullo, N.M.; Parikshak, N.N.; Stein, J.L.; et al. De Novo Mutations Revealed by Whole-Exome Sequencing Are Strongly Associated with Autism. *Nature* **2012**, *485*, 237–241. [CrossRef] [PubMed]
43. Du, Y.; Li, Z.; Liu, Z.; Zhang, N.; Wang, R.; Li, F.; Zhang, T.; Jiang, Y.; Zhi, X.; Wang, Z.; et al. Nonrandom Occurrence of Multiple de Novo Coding Variants in a Proband Indicates the Existence of an Oligogenic Model in Autism. *Genet. Med.* **2020**, *22*, 170–180. [CrossRef] [PubMed]
44. Wilfert, A.B.; Turner, T.N.; Murali, S.C.; Hsieh, P.H.; Sulovari, A.; Wang, T.; Coe, B.P.; Guo, H.; Hoekzema, K.; Bakken, T.E.; et al. Recent Ultra-Rare Inherited Variants Implicate New Autism Candidate Risk Genes. *Nat. Genet.* **2021**, *53*, 1125–1134. [CrossRef]
45. Coll-Tane, M.; Krebbers, A.; Castells-Nobau, A.; Zweier, C.; Schenck, A. Intellectual Disability and Autism Spectrum Disorders "on the Fly": Insights from Drosophila. *DMM Dis. Models Mech.* **2019**, *12*, dmm039180. [CrossRef] [PubMed]
46. Li, H.; Durbin, R. Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2010**, *26*, 589–595. [CrossRef] [PubMed]
47. Poplin, R.; Ruano-Rubio, V.; DePristo, M.A.; Fennell, T.J.; Carneiro, M.O.; van der Auwera, G.A.; Kling, D.E.; Gauthier, L.D.; Levy-Moonshine, A.; Roazen, D.; et al. Scaling Accurate Genetic Variant Discovery to Tens of Thousands of Samples. *bioRxiv* **2017**. [CrossRef]
48. Venselaar, H.; te Beek, T.A.H.; Kuipers, R.K.P.; Hekkelman, M.L.; Vriend, G. Protein Structure Analysis of Mutations Causing Inheritable Diseases. An e-Science Approach with Life Scientist Friendly Interfaces. *BMC Bioinform.* **2010**, *11*, 1–10. [CrossRef] [PubMed]

2.3 *SPTBN5*, encoding the βV-spectrin protein, leads to a syndrome of intellectual disability, developmental delay, and seizures

In this section, we illustrate four families with variants in *SPTBN5* associated with NDDs identified by WES.

Bioinformatic predictions and protein modelling by Pymol and AlphaFold revealed a detrimental effect provoked in the proteins encoded by the reported variants. AlphaFold generates 3D models exploiting convolutional neural networks trained on PDB structures, predicting the distances between pairs of amino acids and the angles between chemical bonds connecting those amino acids. AlphaFold appears useful especially in cases for which no structures for homologous proteins are experimentally determined like βV-spectrin, enabling an advance in protein-structure prediction.

The *SPTBN5*-clinical traits in our patients include ID (mild to severe), aggressive tendencies, accompanied by variable features such as craniofacial and physical dysmorphisms (clinodactyly, arachnodactyly, broad nasal bridge, triangular facies, thin upper lip, low set ears), ASD, and gastroesophageal reflux.

frontiers | Frontiers in Molecular Neuroscience

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fnmol.2022.877258

SPTBN5, Encoding the βV-Spectrin Protein, Leads to a Syndrome of Intellectual Disability, Developmental Delay, and Seizures

Amjad Khan[1,2†], Lucia Pia Bruno[2,3†], Fadhel Alomar[4], Muhammad Umair[5,6], Anna Maria Pinto[7], Abid Ali Khan[8], Alamzeb Khan[9], Saima[10], Alessandra Fabbiani[2,3], Kristina Zguro[3], Simone Furini[3], Maria Antonietta Mencarelli[7], Alessandra Renieri[2,3,7], Sara Resciniti[2,3], Karla A. Peña-Guerra[11], Francisco J. Guzmán-Vega[11], Stefan T. Arold[11,12], Francesca Ariani[2,3,7] and Shahid Niaz Khan[13]*

[1] Faculty of Science, Department of Biological Sciences (Zoology), University of Lakki Marwat, Lakki Marwat, Pakistan, [2] Medical Genetics, University of Siena, Siena, Italy, [3] Med Biotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, Siena, Italy, [4] Department of Pharmacology and Toxicology, College of Clinical Pharmacy, Imam Abdulrahman, Bin Faisal University, Dammam, Saudi Arabia, [5] Medical Genomics Research Department, King Abdullah International Medical Research Center (KAIMRC), King Saud bin Abdulaziz University for Health Sciences, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia, [6] Department of Life Sciences, School of Science, University of Management and Technology (UMT), Lahore, Pakistan, [7] Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy, [8] Faculty of Science, Department of Chemical Sciences, University of Lakki Marwat, Lakki Marwat, Pakistan, [9] Department of Pediatrics, Yale School of Medicine, Yal University, New Heaven, CT, United States, [10] Department of Biotechnology, Abdul Wali Khan University Mardan, Mardan, Pakistan, [11] Computational Bioscience Research Center (CBRC), Biological and Environmental Science and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, [12] Centre de Biologie Structurale (CBS), INSERM, CNRS, Université de Montpellier, Montpellier, France, [13] Department of Zoology, Kohat University of Science and Technology, Kohat, Pakistan

Whole exome sequencing has provided significant opportunities to discover novel candidate genes for intellectual disability and autism spectrum disorders. Variants in the spectrin genes *SPTAN1*, *SPTBN1*, *SPTBN2*, and *SPTBN4* have been associated with neurological disorders; however, *SPTBN5* gene-variants have not been associated with any human disorder. This is the first report that associates *SPTBN5* gene variants (ENSG00000137877: c.266A>C; p.His89Pro, c.9784G>A; p.Glu3262Lys, c.933C>G; p.Tyr311Ter, and c.8809A>T; p.Asn2937Tyr) causing neurodevelopmental phenotypes in four different families. The *SPTBN5*-associated clinical traits in our patients include intellectual disability (mild to severe), aggressive tendencies, accompanied by variable features such as craniofacial and physical dysmorphisms, autistic behavior, and gastroesophageal reflux. We also provide a review of the existing literature related to other spectrin genes, which highlights clinical features partially overlapping with *SPTBN5*.

Keywords: intellectual disability (ID), whole exome sequencing (WES), *SPTBN5*, heterozygous mutation, protein modeling 3

OPEN ACCESS

Edited by:
Maren Engelhardt, Johannes Kepler University of Linz, Austria

Reviewed by:
Alexei Surguchev, Yale University, United States
Matthew N. Rasband, Baylor College of Medicine, United States

***Correspondence:**
Amjad Khan
amjad@ulm.edu.pk

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Brain Disease Mechanisms, a section of the journal Frontiers in Molecular Neuroscience

Received: *16 February 2022*
Accepted: *07 April 2022*
Published: *17 June 2022*

Citation:
Khan A, Bruno LP, Alomar F, Umair M, Pinto AM, Khan AA, Khan A, Saima, Fabbiani A, Zguro K, Furini S, Mencarelli MA, Renieri A, Resciniti S, Peña-Guerra KA, Guzmán-Vega FJ, Arold ST, Ariani F and Khan SN (2022) SPTBN5, Encoding the βV-Spectrin Protein, Leads to a Syndrome of Intellectual Disability, Developmental Delay, and Seizures. Front. Mol. Neurosci. 15:877258. doi: 10.3389/fnmol.2022.877258

INTRODUCTION

Intellectual disability (ID) encompasses a heterogeneous group of neurodevelopmental disorders characterized by substantial intellectual and adaptive functioning limitations before the age of 18 (Musante and Ropers, 2014; Khan et al., 2020; Nøstvik et al., 2021). The overall incidence of ID varies from 1 to 3% in the general population and is commonly defined by an intelligence quotient (IQ) score of <70 (Ropers, 2010; Umair et al., 2020; Nøstvik et al., 2021). ID can be syndromic or non-syndromic and can be attributed to genetic and environmental factors (Umair et al., 2020; Shao et al., 2021). Prenatal and perinatal events, such as drug and toxins exposure during pregnancy, are currently correlated to autism (ASD) and ID (Bilder et al., 2013; Wang et al., 2017). To date, more than 700 causative genes have been reported to cause ID. They are implicated in many important biological processes such as cell cycle regulation, DNA methylation, DNA repair, damage response, chromatin remodeling transcription, and translational processes (Dulac, 2010; Bourgeron, 2015; Vissers et al., 2016; Deciphering Developmental Disorders, 2017).

The advent of next-generation sequencing (NGS) technologies, comprising whole exome and genome sequencing along with the use of different bioinformatics databases that promote sharing of information on genotype-phenotype correlation, has been a significant factor in the remarkable progress in unraveling the genetics of ID (Boycott et al., 2017; Khan et al., 2020; Umair et al., 2020).

The spectrin beta, non-erythrocytic 5 (*SPTBN5*) gene, alternatively called beta V spectrin, BSPECV, HUSPECV, and HUBSPECV is located on chromosome 15q15.1, having 68 exons encoding 3674 amino acid spectrin protein (Stabach and Morrow, 2000; **Figure 1E**). The spectrin protein is composed of calponin homology domains (CH), spectrin repeats, and pleckstrin homology domain (PH; **Figure 1D**). Spectrin is considered a central part of a ubiquitous complex system linking membrane proteins, lipids, and cytosolic factors with the significant cytoskeletal elements of the cell (Kennedy et al., 1994).

We present four different families with homozygous variants in the *SPTBN5* gene associated for the first time with a neurodevelopmental disorder. Furthermore, we interpret our findings in the light of what was previously reported regarding patients with mutations in the spectrin genes, delineating a phenotypic continuum for this family of genes.

MATERIALS AND METHODS

Consent Approval and Patient Recruitment

Family elders (parents) provided written informed consent for molecular analysis and publication of the clinical data for research and diagnosis purposes. The families reported in the present study originated from Pakistan and Italy. Family A was diagnosed at Khalifa Gul Nawaz Teaching Hospital, and the district headquarter hospital in Bannu Khyber Pakhtunkhwa, Pakistan. Families (B–D) were diagnosed at the Medical Genetics department of the University of Siena, Italy (**Figure 1A**).

Genetic counseling was performed to evaluate each patient's personal and familial history.

DNA Extraction and Quantification

Genomic DNA of both the affected and unaffected family members was isolated from peripheral blood using standard methods and quantified using the Nanodrop-2000 spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA).

Library Construction

Sample preparation was performed following the Illumina DNA prep with enrichment manufacturer protocol. A bead-based transposome complex was used to perform segmentation, which fragments the genomic DNA and then tags it with adapter sequences in one step. After saturation with input DNA, the bead-based transposome complex fragments a set number of DNA molecules. This fragmentation allows a wide DNA input range to generate normalized libraries with a consistent tight fragment size distribution. Then a limited-cycle PCR adds adapter sequences to the ends of a DNA fragment. A subsequent target enrichment workflow is then applied. Following pooling, the double-stranded DNA libraries are denatured and biotinylated. Next, Illumina Exome Panel v1.2 (CEX) probes are hybridized into the denatured library fragments. Then Streptavidin Magnetic Beads capture the targeted library fragments within the regions of interest. Then the indexed libraries are eluted from beads and further amplified before sequencing. The exome sequencing analysis was performed on the Illumina *NovaSeq6000 System* (Illumina San Diego, CA, USA) according to the NovaSeq6000 System Guide. Reads were mapped against the hg19 reference genome using the Burrow-Wheeler aligner BWA (Li and Durbin, 2010). Variant calling was obtained using an in-house pipeline which takes advantage of the GATK Best Practices workflow.

Whole Exome Sequencing and Data Analysis

Trio whole-exome sequencing (WES) was performed, including the proband and both parents. The caring clinicians gathered clinical and mutation details of the patients harboring the *SPTBN5* variants by getting in touch through Gene Matcher (http://www.genematcher.org). Exome sequencing and variant interpretation and analysis were as previously described (Monies et al., 2019; Shao et al., 2021). All the variants were screened according to the location, frequency, and type of variation. Variants were filtered with a minor allele frequency (MAF) cutoff of 1% in the Exome Variant Server (http://evs.gs.washington.edu/EVS/), GnomAD (https://gnomad.broadinstitute.org), and 1000 Genomes (http://www.1000genomes.org/). Sanger sequencing was performed using stranded methods (Khan et al., 2019; Umair et al., 2019).

Bioinformatics Analysis

The bioinformatics analysis focused on non-synonymous SNVs (missense, non-sense, splice-site, and frameshift) and was submitted to Polyphen-2 (http://genetics.bwh.

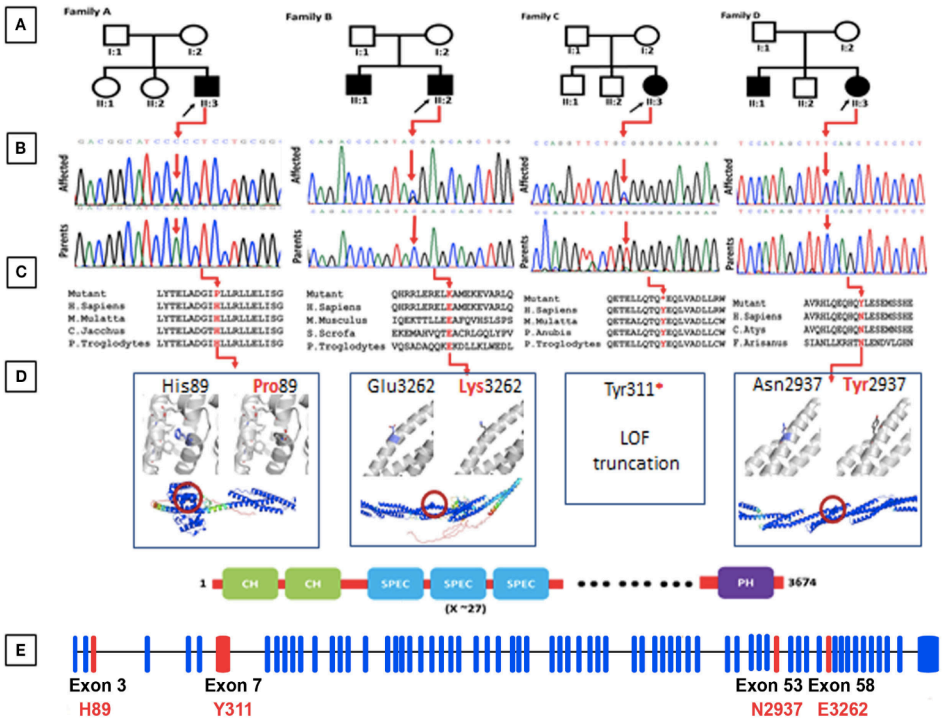


FIGURE 1 | (A) Representing pedigrees of all the four families (A–D). **(B)** Showing Sanger sequencing electro-grams of all the four families. **(C)** Conservation analysis of the mutations identified in the present studies along different species. **(D)** Structural models of the CH1 domain containing the His89Pro mutation, and the spectrin repeats containing the Asn2937Tyr and Glu3262Lys mutations. Three 1,000 segments containing the residues of interest were modeled with AlphaFold, and are shown in the bottom as cartoons, colored by pLDDT score. Residues with pLDDT<50 are colored red, representing regions with low confidence, and residues with pLDDT>90 are blue, showing high-confidence segments. The top insets show a zoom into the local region of each mutated residue (shown as blue sticks), and the comparison between the wild-type and mutated states. The residues predicted to interact with His89 are shown as gray sticks. **(E)** Exon organization of SPTBN5. Boxes are exons. Lines connecting the boxes are introns. Filled boxes are coding sequence, and empty, unfilled boxes are UTR (UnTranslated Region). Adapted from Ensembl (release 105) (Howe et al., 2021).

harvard.edu/pph2/), Sorting Intolerant from Tolerant (SIFT, http://sift.jcvi.org/), Protein Variation Effect Analyzer (PROVEAN, http://provean.jcvi.org), Mutation Taster (http://www.mutationtaster.org/), Varsome (https://varsome.com/), Mutation assessor (http://mutationassessor.org), and Combined Annotation Dependent Depletion (CADD, https://cadd.gs.washington.edu/) for functional effect prediction.

Protein Modeling

AlphaFold (Jumper et al., 2021) was used to produce high-quality structural models of three 1,000 residue segments of SPTBN5, containing the residues of interest. The average pLDDT score of the three models was higher than 80, and the residue pLDDT of His89, Asn2937, and Glu3262 and their local region was higher than 90, indicating a very confident 3D configuration. A configuration of four V100 GPUs and 16 CPU cores (provided by the KAUST IBEX cluster) was used for modeling. Models were manually inspected, and mutations were evaluated using the Pymol program (pymol.org). The American College of Medical Genetics and Genomics (ACMG) 2015 guidelines were used for the interpretation of variants (Li et al., 2020).

Primer Designing and Mutation Confirmation

Gene Runner (version 5.0.69 Beta, Hastings, NY, USA) software was used for Primers design. Sanger sequencing was performed using an ABI3730 Automated Sequencer to verify co-segregation of the identified variants with the disease phenotype (Thermo Fisher Scientific, Waltham, MA, USA). The Sanger sequencing results were examined and compared with the help of visual software such

TABLE 1 | Clinical features of affected subjects from families A–D.

Family parameters	**Families information**			
Parameters	**Family A**	**Family B**	**Family C**	**Family D**
Pedigree ID	II:3	II:2	II:3	II:3
Nationality	Pakistani	Italian	Italian	Italian
Gender	Male	Female	Male	Female
Current age (Years)	11 yrs	15 yrs	14 yrs	7 yrs
Family history	Sporadic	Familial	Sporadic	Familial
Disease onset (years)	First year of life	First year of life	First year of life	First year of life
Consanguinity	No	No	No	No
Gestation weeks (weeks)	38	39	38	38
Pregnancy event	Uneventful	Uneventful	Uneventful	Uneventful
Developmental Features				
Developmental delay	+	+	+	+
Language impairment	+	+		+
Learning disability	+	+	+	+
Sleep disorder	+	+	+	+
Head circumference	49cm	50.5 cm	46.5 cm	–
Height	52.1 cm	144.5 cm	150.5 cm	97 cm
Weight	45.9 kg	35 kg	34 kg	15 kg
Dysmorphic Features				
Low set ears	+	–	–	–
Nasal bridge	Broad	Broad	Depressed	Broad
Strabismus	+	–	–	–
Facial expression	Triangular	Triangular	Triangular	–
Thin upper lip	+	+	+	+
Skeletal anomalies				
Hands	Bilateral clinodactyly of the 5th little finger	Arachnodactyly, fusiforme fingers of the hands	Bilateral clinodactyly of 4th and 5th fingers of hands	Bilateral clinodactyly of 4th and 5th fingers of hands
Feet	Brachydactyly of the feet	Fusiforme fingers of the feet	Bilateral clinodactyly of 4th and 5th fingers of the feet	Bilateral clinodactyly of 4th and 5th fingers of hands
Behavioral Features				
Impairment social interaction	+	+	+	+
Feeding difficulty	+	+	–	+
Agression/hyperactivity	+	+	+	+
Neurological Features				
Intellectual disability	+	+	+	+
Anxiety/psychiatric	+	+	+	+
Seizure	+	+	+	–
Amnesia	+	+	+	+
Feeding difficulty	+	+	+	+
Karyotype	Normal	Normal	Normal	Normal

+, *present*; –, *absent*.

as Chromas Lite (http://technelysium.com.au/wp/) and Codon Code Aligner (https://www.codoncode.com/aligner/).

RESULTS

Clinical Report

This study recruited four families affected by ID from Pakistan and Italy. A clinical neurologist examined all the affected individuals. Additional clinical information on the affected individuals is summarized in **Table 1**.

Family A

Family A, of Pakistani origin, has a single affected individual (II:3), a 11-year-old male with ID had a developmental delay born from non-consanguineous parents. The pregnancy and delivery events were unremarkable. The parents of the affected

individuals are normal and have no neurological symptoms related to cognition. Patient II-1 was examined at the age of 10 and 11 years, respectively. Clinical examination revealed early-onset epilepsy and seizures at 2–3 years of age, anxiety, severe chronic constipation, aggressive behavior, feeding difficulties, and harming himself by beating his head against the wall. Physical examination revealed a prominent metopic ridge, strabismus, epicanthus, flat philtrum, stuttering, and relatively thin upper lips with mild wide spacing of teeth. He had a normal 46 XY karyotype. At the last examination (11 years of age), his weight, height, and Occipital Frontal Circumference (OFC) were 45.9 kg, 52.1 cm, and 49 cm, respectively. No cardiac, respiratory, skeletal, or skin anomalies were observed. His vision and hearing were seen as normal (**Figure 1A**).

Family B

We reported a 15-year-old girl with ID born from non-consanguineous parents regarding the Italian family B. She was born after pregnancy with threats of miscarriage; a growth deficit was signaled during the last month. She was diagnosed with drug-induced autoimmunity at birth and interatrial and interventricular defects (DIA and DIV). She came to our attention when she was 3-year-old, and since her first months of life, she presented recurrent infections in the upper respiratory vias and feeding difficulties with gastroesophageal reflux. The anamnestic collection reported a paternal cousin who acquired severe ID after meningitis. Physical examination at the age of 15 years showed a height of 144.5 cm, a weight of 35 kg, and a head circumference of 50.5 cm. She suffered from ID, hyperactivity, language/speech delay and aggressiveness, and a disturbing wake-sleep cycle. Arachnodactyly of the hands and feet was noticed. She showed triangular facies, prominent ears, thin upper lip, absent eyebrows, broad nasal bridge, bulbous nasal tip, thin and sparse hair, fusiform fingers of the hands and the feet, everted lower lip, M-shaped upper lip, hairline anteriorly advanced. Her array-CGH exams gave negative results (**Figure 1A**).

Family C

Family C includes a 14-year-old Italian boy with a history of physical dysmorphisms and behavioral and nonverbal learning disorders. He was born at 38 weeks with a cesarian section. Her anamnestic collection presented a brother with language and behavior delay, a maternal cousin with language delay, and a paternal cousin with seizures at pediatric age. His weight and height were 34 kg and 150.5 cm, and OFC was 46.5 cm, respectively. He showed psychomotor delay, cognitive delay, and manifested aggressiveness through himself and the others. His physical examination reveals thin upper lips, bilateral clinodactyly of the 3rd−4th and 5th fingers, depressed nasal bridge, turricephaly, high anterior hairline, hypertelorism, nasal voice, prominent forehead suture. His karyotype, array-CGH, and X-fragile exams were normal (**Figure 1A**).

Family D

Family D of Italian origin had an affected 6-year-old female child diagnosed with autism and ID. Her weight was 15 kg, height 97 cm, and chest circumference (CC) 50 cm. Her anamnestic collection highlighted language delay in the maternal line and autistic disorders in the paternal line. She was born at term and suffered from gastroesophageal reflux and feeding difficulties. Her weaning and her wake-sleep cycle were normal. However, she manifested aggressiveness, language delay, and expression disorder. Her ADOS score was 16 (moderate ASD), and her auditory evoked potentials were negative. In addition, her array-CGH exam was negative (**Figure 1A**).

Identification of Candidate *SPTBN5* Variants

By applying an iterative filtering strategy based on variant frequency, functional consequences on coding sequence, and heredity, *de novo* variants in the *SPTBN5* gene [NM_016642] were considered the best candidate. These include c.266A>C; [p.(His89Pro)] in Family A, c.9784G>A; [p.(Gly3262Lys)] in Family B, c.933C>G; [p.(Tyr311*)] family C, and c.8809A>T; [p.(Asn2937Tyr)] in family D. These candidate variants were confirmed by Sanger sequencing in all the available family members (**Figure 1B**).

Predicted Molecular Effect of the *SPTBN5* Variants

The *SPTBN5* encodes a 3674 long amino acids protein. It consists of two N-terminal calponin homology (CH) domains, actin-binding domains that cross-link actin filaments into bundles and networks (Richards et al., 2015). The α-helical CH domains are common in actin-binding proteins and play important regulatory roles in cytoskeletal dynamics and signaling. Most of the region C-terminal to the CH domains, almost 90% of the protein, is formed by spectrin repeats. The spectrin repeats form a tandem arrangement of helical coiled coils, where each repeat forms a three-helix bundle. Through tension-induced unfurling of the three-helix repeats, these domains allow spectrin to expand and contract, conferring flexibility to the protein (Borrego-Diaz et al., 2006). The spectrin repeats are also self-association and tetramer formation with alpha spectrins (Nicolas et al., 1998; Stradal et al., 1998). Finally, there is a pleckstrin homology domain (PH) at the C-terminal end. The PH domains bind to phosphatidylinositol lipids, allowing the recruitment of proteins to cellular membranes (Mayer et al., 1993).

His89Pro

His89 is exposed and located at the end of an alpha helix in the first CH domain, surrounded by the residues forming the linker to the adjacent helix (**Figure 1D**). AlphaFold predicts several interactions of His89 with its surrounding residues, namely Ile64, Phe68, Ile75, Ile77, and Glu83. The substitution for a proline could destabilize the helical structure leading to an earlier break, affecting the 3D packing and stability of this globular domain. This mutation was predicted as probably damaging by Polyphen2 (score = 0.998; Adzhubei et al., 2010), tolerated by SIFT (score = 0.31; Sim et al., 2012), and as deleterious by PROVEAN (score = −5.450; Choi et al., 2012).

TABLE 2 | *In silico* prediction analysis of *SPTBN5* variants in the present families.

Paramètres	Family A	Family B	Family C	Family D
Gene name		***SPTBN5***		
Other names		*BSPECV; HUSPECV; HUBSPECV*		
Chromosome location		15q15.1		
MIM number		605916		
Ensemble ID		ENSG00000137877		
Total exon		68		
cDNA change	c.266A>C	c.9784G>A	c.933C>G	c.8809A>T
Protien change	p.His89Pro	p.Glu3262Lys	p.Tyr311Ter	p.Asn2937Tyr
Variant exonic location	Exon 3	Exon 58	Exon 7	53
Variant chromosome location	Chr15:42185210	Chr15:41853778	Chr15:41886322	Chr15: 41856598
Variant type	Non synonymous	Non synonymous	Non synonymous	Non synonymous
SIFT	0.226/Tolerable	0.127/Tolerable	–	0.023/Damaging
Polyphen-2_HDIV	0.998/Probably_damaging	0.003/Benign	–	0.989/Probably_damaging
Polyphen-2_HVAR	0.939/Probably_damaging	0.014/Benign	–	0.832/Possibly_damaging
FATHMM	−3.54/ Damaging	0.73/ Tolerable	–	0.75/Tolerable
Mutation taster	1.000/Polymorphism	0.983/Polymorphism	1/Disease_causing	0.959/Disease_causing
PROVEAN	−5.45/Damaging	0.25/Tolerable	–	−4.54/Damaging
MetaSVM	0.088/Damaging	−1.014/Tolerable	–	−0.633/Tolerable
CADD	21.2/Damaging	22.5/Damaging	35/Damaging	26.6/Damaging
FATHMM_MKL	0.559/Damaging	0.513/Damaging	0.449/Tolerable	0.990/Damaging
GERP++	2.82/Conserved	1.48/conserved	−0.984/Nonconserved	5.01/Conserved
GnomAD_exome All	0.000004124	0.0001	0.000008175	0.0045

Tyr311*

The *Tyr311** mutation eliminates all the spectrin repeats and the C-terminal PH domain. It is expected to render the protein non-functional for activities that require interactions, providing flexibility or support through these domains. This mutation was predicted to be deleterious by PROVEAN (score = −14.963).

Asn2937Tyr and Glu3262Lys

Asn2937 and Glu3262 are located on the exterior of one of the repeat helices of a spectrin repeat, with Glu3262 being close to the C-terminus. Neither Asn2937Tyr nor Glu3262Lys are expected to severely disrupt the structural stability of their spectrin repeats as the side chains of both point toward the solvent and AlphaFold does not predict contacts with neighboring residues. However, both mutations change the local physicochemical surface characteristics in shape (Asn2937Tyr) and charge (Glu3262Lys). Thus, both variants could weaken or disrupt interactions with other molecules or affect the dynamics of the spectrin repeat under tension. Asn2937Tyr was predicted as probably damaging by Polyphen2 (score = 0.986), tolerated by SIFT (score = 0.25), and as deleterious by PROVEAN (score = −4.539). Glu3262Lys was predicted as benign by Polyphen2 (score = 0.002), tolerated by SIFT (score=0.82), and neutral by PROVEAN (score = 0.250).

In conclusion, *Tyr311** has by far the most deleterious effect by only leaving the CH domains intact. By affecting the CH domain stability and surface, His89Pro might weaken acting associations and hence make these interactions more prone to disruption under tension. In comparison, Asn2937Tyr and Glu3262Lys present mutations that are relatively benign for the tertiary structure of an individual molecule. However, their effect on dynamics and/or interactions is expected to become additive and potentially disruptive within spectrin networks. The results of the *in silico* analysis have been summarized in **Table 2**.

DISCUSSION

Spectrins constitute a family of cytoskeletal proteins interacting with actin filaments, microtubules, and intermediate filaments, and connecting the cytoskeleton to the plasma membrane (Ortiz-Gonzalez and Wierenga, 2020; Rosenfeld et al., 2021). *SPTBN5* forms extensive networks by forming tetramers and higher oligomers through homologous associations and associations with alpha spectrins (Stabach and Morrow, 2000). In this study, we report, for the first time, four different families of *SPTBN5* deficiency with features such as ID, developmental delay, seizures, aggressive behavior, variable dysmorphic features, and limb malformations. The molecular analysis using NGS identified four heterozygous variants {c.266A>C; [p.(His89Pro)] in family A, c.9784G>A; [p.(Glu3262Lys)] in family B, c.933C>G; [p.(Tyr311*)] family C, and c.8809A>T; [p.(Asn2937Tyr)] in family D} in *SPTBN5* gene. These variants are located in highly conserved positions and are predicted to be detrimental based on various *in silico* analyses (**Figure 1C**). Additional functional evidence is needed to clarify how the *SPTBN5* haploinsufficiency affects brain malformation.

SPTBN5 and other spectrins are expressed ubiquitously in the body, including the brain, eye, kidney, heart, gastrointestinal tract, and musculoskeletal tissue (https://www.proteinatlas.org/); thus, it is not surprising that the disease phenotype involves multiple organs (Uhlén et al., 2015). In addition, mutations in various members of the spectrin gene family are associated with erythroid cell disorders (*SPTA1*, *SPTB*) and neurological disorders (*SPTAN1, SPTBN1, SPTBN2,* and *SPTBN4*); however, no human genotype-phenotype correlation has been established for *SPTBN5* to date (Rosenfeld et al., 2021). Here we report *SPTBN5* variants that are predicted to severely truncate the protein (Tyr311*), weaken its actin association (His89Pro), or disrupt the integrity of spectrin networks (Asn2937Tyr and Glu3262Lys).

A comparison of the *SPTBN5* phenotype to the published *SPTBN1* and *SPTBN4* phenotype shows common features of ID, DD, and aggression reported in all four families for both diseases (Ortiz-Gonzalez and Wierenga, 2020; Rosenfeld et al., 2021). In addition, individuals with *SPTBN5* variants had low rates of seizures as compared to *SPTBN1 SPTBN4*, and individuals with *SPTBN1* and *SPTBN4* variants had cerebellar or cerebral atrophy that was not detected. In the *SPTBN5* cohort. However, abnormalities of the corpus callosum were found in both cohorts. In addition, our results further support the contribution of inherited pathogenic variants in candidate genes to ASD and ID and reinforce the theory of a multi-hit model, according to the coexistence between ultra-rare inherited variants and *de novo* mutations has been observed in ASD trios (Krumm et al., 2015; Wilfert et al., 2021).

In summary, we implicate *SPTBN5* as a gene whose disruption leads to human neurodevelopmental disease. However, open questions remain regarding the variability of the phenotype and the role of *SPTBN5* in multiple organ systems.

DATA AVAILABILITY STATEMENT

The datasets during and/or analyzed during the current study will be available from the corresponding author on reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Lakki Marwat, Pakistan and UMT, Lahore, Pakistan [Ref.#: DLSBBC-2022-04]. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AmK, LB, MU, AbK, and AlK: conceptualization, methodology, and writing—original draft preparation. KP-G, FG-V, and SA: structural modeling and manuscript writing. AR, AbK, FA, AP, and SN: conception and design of the study and reviewing the manuscript. AzK, Saima, AF, KZ, SF, MM, AR, and SR: critical proof reading. KP-G, FG-V, SA, AzK, Saima, AF, KZ, SF, MM, AR, and SR: data analysis. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research by KP-G, FG-V, and SA were supported by the King Abdullah University of Science and Technology (KAUST) through the baseline fund and the Award No. FCC/1/1976-25 and REI/1/4446-01 from the Office of Sponsored Research (OSR). For computer time, this research used the resources of the Supercomputing Laboratory at KAUST. We are grateful to our patients for their cooperation. This work was generated within the ERN ITHACA (European Reference Network for Intellectual Disability, Telehealth, Autism, and Congenital Anomalies). The Cell lines and DNA bank of Rett Syndrome, X-linked mental retardation, and other genetic diseases, member of the Telethon Network of Genetic Biobanks (project nos. GTB12001 and GFB18001), funded by Telethon Italy, and of the EuroBioBank network provided us with specimens.

ACKNOWLEDGMENTS

We thank all the individuals for their participation in this study.

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., et al. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249. doi: 10.1038/nmeth0410-248

Bilder, D. A., Pinborough-Zimmerman, J., Bakian, A. V., Miller, J. S., Dorius, J. T., et al. (2013). Prenatal and perinatal factors associated with intellectual disability. *Am. J. Intellect. Dev. Disabil.* 118, 156–176. doi: 10.1352/1944-7558-118.2.156

Borrego-Diaz, E., Kerff, F., Lee, S. H., Ferron, F., Li, Y., and Dominguez, R. (2006). Crystal structure of the actin-binding domain of alpha-actinin 1: evaluating two competing actin-binding models. *J. Struct. Biol.* 155, 230–238. doi: 10.1016/j.jsb.2006.01.013

Bourgeron, T. (2015). From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat. Rev. Neurosci.* 16, 551–563. doi: 10.1038/nrn3992

Boycott, K. M., Rath, A., Chong, J. X., Hartley, T., Alkuraya, F. S., Baynam, G., et al. (2017). International cooperation to enable the diagnosis of all rare genetic diseases. *Am. J. Hum. Genet.* 100, 695–705. doi: 10.1016/j.ajhg.2017.04.003

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., and Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* 7:e46688. doi: 10.1371/journal.pone.0046688

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

Dulac, C. (2010). Brain function and chromatin plasticity. *Nature* 465, 728–735. doi: 10.1038/nature09231

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Ridwan Amode, M., et al. (2021). Ensembl. *Nucleic Acids Res.* 49. D884–D891. doi: 10.1093/nar/gkaa942

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi: 10.1038/s41586-021-03819-2

Kennedy, S. P., Weed, S. A., Forget, B. G., and Morrow, J. S. (1994). A partial structural repeat forms the heterodimer self-association site of all beta-spectrins. *J. Biol. Chem.* 269, 11400–11408. doi: 10.1016/S0021-9258(19)78138-1

Khan, A., Miao, Z., Umair, M., Ullah, A., Alshabeeb, M. A., Bilal, M., et al. (2020). Two Cases of recessive intellectual disability caused by NDST1 and METTL23 variants. *Genes* 11:1021. doi: 10.3390/genes11091021

Khan, A., Wang, R., Han, S., Umair, M., Abbas, S., Khan, M. I., et al. (2019). Homozygous missense variant in the TTN gene causing autosomal recessive limb-girdle muscular dystrophy type 10. *BMC Med. Genet.* 20:166. doi: 10.1186/s12881-019-0895-7

Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., et al. (2015). Excess of rare, inherited truncating mutations in autism. *Nat. Genet.* 47, 582–588. doi: 10.1038/ng.3303

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, J., Chen, K., Zhu, R., and Zhang, M. (2020). Structural basis underlying strong interactions between ankyrins and spectrins. *J. Mol. Biol.* 432, 3838–3850. doi: 10.1016/j.jmb.2020.04.023

Mayer, B. J., Ren, R., Clark, K. L., and Baltimore, D. (1993). A putative modular domain present in diverse signaling proteins. *Cell* 73, 629–630. doi: 10.1016/0092-8674(93)90244-K

Monies, D., Abouelhoda, M., Assoum, M., Moghrabi, N., Rafiullah, R., Almontashiri, N., et al. (2019). Lessons learned from large-scale, first-tier clinical exome sequencing in a highly consanguineous population. *Am. J. Hum. Genet.* 104, 1182–1201. doi: 10.1016/j.ajhg.2019.04.011

Musante, L., and Ropers, H. H. (2014). Genetics of recessive cognitive disorders. *Trends Genet.* 30, 32–39. doi: 10.1016/j.tig.2013.09.008

Nicolas, G., Pedroni, S., Fournier, C., Gautero, H., Craescu, C., Dhermy, D., et al. (1998). Spectrin self-association site: characterization and study of beta-spectrin mutations associated with hereditary elliptocytosis. *Biochem. J.* 332 (Pt 1), 81–89. doi: 10.1042/bj3320081

Nøstvik, M., Kateta, S. M., Schönewolf-Greulich, B., Afenjar, A., Barth, M., Boschann, F., et al. (2021). Clinical and molecular delineation of PUS3-associated neurodevelopmental disorders. *Clin. Genet.* 100, 628–633. doi: 10.1111/cge.14051

Ortiz-Gonzalez, X., and Wierenga, K. (2020). "SPTBN4 disorder," in *GeneReviews® [Internet]*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, G. Mirzaa, and A. Amemiya (Seattle, WA: University of Washington), 1993–2021.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17, 405–423. doi: 10.1038/gim.2015.30

Ropers, H. H. (2010). Genetics of early onset cognitive impairment. *Annu. Rev. Genomics Hum. Genet.* 11, 161–187. doi: 10.1146/annurev-genom-082509-141640

Rosenfeld, J. A., Xiao, R., Bekheirnia, M. R., Kanani, F., Parker, M. J., Koenig, M. K., et al. (2021). Heterozygous variants in SPTBN1 cause intellectual disability and autism. *Am. J. Med. Genet. A* 185, 2037–2045. doi: 10.1002/ajmg.a.62201

Shao, D. D., Straussberg, R., Ahmed, H., Khan, A., Tian, S., Hill, R. S., et al. (2021). A recurrent, homozygous EMC10 frameshift variant is associated with a syndrome of developmental delay with variable seizures and dysmorphic features. *Genet. Med.* 23, 1158–1162. doi: 10.1038/s41436-021-01097-x

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., and Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 40, W452–W457. doi: 10.1093/nar/gks539

Stabach, P. R., and Morrow, J. S. (2000). Identification and characterization of beta V spectrin, a mammalian ortholog of Drosophila beta H spectrin. *J. Biol. Chem.* 275, 21385–21395. doi: 10.1074/jbc.C000159200

Stradal, T., Kranewitter, W., Winder, S. J., and Gimona, M. (1998). CH domains revisited. *FEBS Lett.* 431, 134–137. doi: 10.1016/S0014-5793(98)00751-0

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347:1260419. doi: 10.1126/science.1260419

Umair, M., Ballow, M., Asiri, A., Alyafee, Y., Al Tuwaijri, A., Alhamoudi, K. M., et al. (2020). EMC10 homozygous variant identified in a family with global developmental delay, mild intellectual disability, and speech delay. *Clin. Genet.* 98, 555–561. doi: 10.1111/cge.13842

Umair, M., Khan, A., Hayat, A., Abbas, S., Asiri, A., Younus, M., et al. (2019). Biallelic missense mutation in the ECEL1 underlies distal arthrogryposis type 5 (DA5D). *Front. Pediatr.* 7:343. doi: 10.3389/fped.2019.00343

Vissers, L. E., Gilissen, C., and Veltman, J. A. (2016). Genetic studies in intellectual disability and related disorders. *Nat. Rev. Genet.* 17, 9–18. doi: 10.1038/nrg3999

Wang, C., Geng, H., Liu, W., and Zhang, G. (2017). Prenatal, perinatal, and postnatal factors associated with autism: a meta-analysis. *Medicine* 96:e6696. doi: 10.1097/MD.0000000000006696

Wilfert, A. B., Turner, T. N., Murali, S. C., Hsieh, P., Sulovari, A., Wang, T., et al. (2021). Recent ultra-rare inherited variants implicate new autism candidate risk genes. *Nat. Genet.* 53, 1125–1134. doi: 10.1038/s41588-021-00899-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

2.4 Natural history of KBG syndrome in a large European cohort

In the context of a European collaborative study, we collected 49 patients affected by KBG syndrome (KBGS) from prenatal age to forties with the aim of characterising the KBGS-associated clinical features progression. A combined array-CGH and NGS approach investigated both genomic CNVs and SNVs, permitting us to assign a diagnosis when rare or unique clinical signs hamper the medical ascertainment.

Our results allow us to broaden the spectrum of KBGS phenotype progression and demonstrate that the combined use of ES and array-CGH supports the diagnosis of KBGS, particularly in paediatric patients with immature clinical findings.

OXFORD

Human Molecular Genetics, 2022, Vol. 00, 00, 1–12
https://doi.org/10.1093/hmg/ddac167
Advance access publication date: 21 July 2022
Original Article

Natural history of KBG syndrome in a large European cohort

Lorenzo Loberti[1,2,27,†], Lucia Pia Bruno[1,2,†], Stefania Granata[1,2,27], Gabriella Doddato[1,2], Sara Resciniti[1,2], Francesca Fava[1,2,27], Michele Carullo[1,2], Elisa Rahikkala[3], Guillaume Jouret[4], Leonie A. Menke[5], Damien Lederer[6], Pascal Vrielynck[7], Lukáš Ryba[8], Nicola Brunetti-Pierri[9], Amaia Lasa-Aranzasti[10], Anna Maria Cueto-González[10], Laura Trujillano[10], Irene Valenzuela[10], Eduardo F. Tizzano[10], Alessandro Mauro Spinelli[11], Irene Bruno[12], Aurora Currò[13], Franco Stanzial[13], Francesco Benedicenti[13], Diego Lopergolo[14], Filippo Maria Santorelli[14], Constantia Aristidou[15], George A. Tanteles[15], Isabelle Maystadt[6], Tinatin Tkemaladze[16], Tiia Reimand[17,18], Helen Lokke[17,18], Katrin Õunap[17,18], Maria K. Haanpää[19], Andrea Holubová[8], Veronika Zoubková[8], Martin Schwarz[8], Riina Žordania[17], Kai Muru[17,18], Laura Roht[17,18], Annika Tihveräinen[20], Rita Teek[17], Ulvi Thomson[21], Isis Atallah[22], Andrea Superti-Furga[22], Sabrina Buoni[23], Roberto Canitano[23], Valeria Scandurra[23], Annalisa Rossetti[24], Salvatore Grosso[24], Roberta Battini[25,26], Margherita Baldassarri[1,2], Maria Antonietta Mencarelli[27], Caterina Lo Rizzo[27], Mirella Bruttini[1,2,27], Francesca Mari[1,2,27], Francesca Ariani[1,2,27], Alessandra Renieri[1,2,27,*] and Anna Maria Pinto[27]

[1]Medical Genetics, University of Siena, Siena 53100, Italy
[2]Med Biotech Hub and Competence Centre, Department of Medical Biotechnologies, University of Siena, Siena 53100, Italy
[3]Department of Clinical Genetics, PEDEGO Research Unit, and Medical Research Center Oulu, University of Oulu and Oulu University Hospital, Oulu 90014, Finland
[4]National Center of Genetics (NCG), Laboratoire national de santé (LNS), L-3555 Dudelange, Luxembourg
[5]Amsterdam UMC location University of Amsterdam, Department of Pediatrics, Amsterdam 1100, The Netherlands
[6]Institut de Pathologie et de Génétique; Centre de Génétique Humaine, Gosselies 6041, Belgium
[7]William Lennox Neurological Hospital, Reference Center for Refractory Epilepsy UCLouvain, Ottignies 1340, Belgium
[8]Department of Biology and Medical Genetics, Charles University – 2nd Faculty of Medicine and University Hospital Motol, Prague 150 00, Czech Republic
[9]Department of Translational Medicine, University of Naples "Federico II", Naples 80125, Italy
[10]Area of Clinical and Molecular Genetics, Vall d'Hebron University Hospital, Barcellona 08035, Spain
[11]Regional Coordinating Center for Rare Diseases, Udine 33100, Italy
[12]Institute for Maternal and Child Health, Trieste 34100, Italy
[13]Genetic Counseling Service, Department of Pediatrics, Regional Hospital of Bolzano, Bolzano 39100, Italy
[14]IRCCS Stella Maris Foundation, Molecular Medicine for Neurodegenerative and Neuromuscular Disease Unit, Pisa 98125, Italy
[15]Department of Clinical Genetics and Genomics, The Cyprus Institute of Neurology & Genetics, Nicosia 1683, Cyprus
[16]Department of Molecular and Medical Genetics, Tbilisi State Medical University, Tbilisi 0162, Georgia
[17]Department of Clinical Genetics, Genetic and Personalized Medicine Clinic, Tartu University Hospital, Tartu 50406, Estonia
[18]Institute of Clinical Medicine, University of Tartu, Tartu 50406, Estonia
[19]Department of Genomics and Clinical Genetics, Turku University Hospital, Turku 20500, Finland
[20]Department of Child Neurology, Turku University Hospital, Turku 20500, Finland
[21]Centre for Neurological Diseases, West-Tallinn Central Hospital, Tallinn 10617, Estonia
[22]Division of Genetic Medicine, Lausanne University Hospital (CHUV) and University of Lausanne, 1011 Lausanne, Switzerland
[23]Division of Child and Adolescent Neuropsychiatry, University of Siena, Siena 53100, Italy
[24]Clinical Paediatrics, Department of Molecular Medicine and Development, University of Siena, Siena 53100, Italy
[25]IRCCS Stella Maris Foundation, Department of Developmental Neuroscience, Pisa 98125, Italy
[26]Department of Clinical and Experimental Medicine, University of Pisa, Pisa 56122, Italy
[27]Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena 53100, Italy
*To whom correspondence should be addressed at: Medical Genetics Unit, University of Siena, Policlinico Santa Maria alle Scotte, Viale Bracci, 2, 53100 Siena, Italy. Tel: 39 0577 233303; Fax: 39 0577 233325; Email: alessandra.renieri@unisi.it

†Co-first authors.

Abstract

KBG syndrome (KBGS) is characterized by distinctive facial gestalt, short stature and variable clinical findings. With ageing, some features become more recognizable, allowing a differential diagnosis. We aimed to better characterize natural history of KBGS. In the context of a European collaborative study, we collected the largest cohort of KBGS patients (49). A combined array- based Comparative Genomic Hybridization and next generation sequencing (NGS) approach investigated both genomic Copy Number Variants and SNVs. Intellectual disability (ID) (82%) ranged from mild to moderate with severe ID identified in two patients. Epilepsy was present in 26.5%. Short stature was consistent over time, while occipitofrontal circumference (median value: −0.88 SD at birth) normalized over years. Cerebral anomalies, were identified in 56% of patients and thus represented the second most relevant clinical feature reinforcing clinical suspicion in the paediatric age when short stature and vertebral/dental anomalies are vague. Macrodontia, oligodontia and

Received: June 1, 2022. **Revised:** June 22, 2022. **Accepted:** July 7, 2022

dental agenesis (53%) were almost as frequent as skeletal anomalies, such as brachydactyly, short fifth finger, fifth finger clinodactyly, pectus excavatum/carinatum, delayed bone age. In 28.5% of individuals, prenatal ultrasound anomalies were reported. Except for three splicing variants, leading to a premature termination, variants were almost all frameshift. Our results, broadening the spectrum of KBGS phenotype progression, provide useful tools to facilitate differential diagnosis and improve clinical management. We suggest to consider a wider range of dental anomalies before excluding diagnosis and to perform a careful odontoiatric/ear-nose-throat (ENT) evaluation in order to look for even submucosal palate cleft given the high percentage of palate abnormalities. NGS approaches, following evidence of antenatal ultrasound anomalies, should include *ANKRD11*.

Introduction

KBG syndrome (KBGS) (OMIM#148050) was firstly described in 1975 by Hermann *et al.* (1) who reported many family members with a distinct phenotype. KBGS is a pan-ethnic syndrome presenting with variable clinical expressivity even across single families and for which an incomplete penetrance has been previously hypothesized (2–5). The prevalence of KBGS is currently unknown and about 200 patients have been described up to date, with male patients generally being more severely affected (2,3,6).

Classical clinical features include macrodontia of the upper central incisors, distinctive craniofacial findings, short stature, skeletal anomalies and neurologic involvement encompassing global developmental delay, electroencephalographic anomalies with or without seizures and intellectual disability (ID) (2). Uncommon clinical features, reported in sporadic reports, include juvenile idiopathic arthritis, dysfunctional dysphonia, multiple dental agenesis, oral frenula, idiopathic precocious thelarche, motor tics, lipoma of the corpus callosum, pilomatrixoma and endothelial corneal polymorphic dystrophy (7).

Monoallelic variants in *ANKRD11* and deletions encompassing *ANKRD11* cause KBGS (8,9). *ANKRD11* is located on chromosome 16, has 13 exons and encodes ankyrin repeat domain 11 (10). Sequencing of *ANKRD11* allows the resolution of more than half of the individuals clinically diagnosed with KBGS (5,8,11–13). Most of the pathogenic variants are truncating, whereas missense variants are rare; plenty of them cluster in exon 10, the largest exon of ANKRD11 (14). Genomic rearrangements of the 16q24.3 chromosomal region, including *ANKRD11*, have also been detected in KBGS patients leading to variable phenotypes depending on the contribution of other genes involved in the rearrangement (9,15).

ANKRD11 is a chromatin regulator of the neural stem cell fate through histone acetylation (16). Mice with heterozygous deletion of *ANKRD11* (17) in neural crest cells present with mild midfacial hypoplasia (reduced midfacial width and a persistent open fontanelle), reproducing the human facial phenotype (16). Mice with homozygous mutations in neural crest cells are not vital at birth (18). *ANKRD11* expression levels reflect the development of the bony structures during craniofacial development, strongly regulating intramembranous ossification and palate development (18). Interestingly, *ANKRD11* is frequently affected by loss of heterozygosity in cancer, and it is supposed to enhance the tumour-suppressive function of TP53 oncogene (19). It has been speculated that ANKRD11 haploinsufficiency may lead to an increased cancer risk in KBGS patients but only one patient with a 16p24.3 deletion involving *ANKRD11* developed a paratesticular rhabdoid tumour (12).

Although broader availability of whole exome sequencing (WES) is increasing the diagnostic rate, a high percentage of KBGS cases remain undiagnosed due to the lack of specific diagnostic clues or mild clinical presentation, except for a recurrent facial gestalt (8,20). In children, younger than 6 years of age, in particular, a clinical diagnosis of KBGS may not be obvious since facial features are subtler and part of the physical characteristics become distinguishable only with age progression. Indeed, an age-dependent evolution of the facial features from Cornelia de Lange (CdLS) to KBGS, from infancy to adolescence, has been previously highlighted (3). Nevertheless, a clinical overlap with the Aarskog, CdLS, Kabuki, Filippi, Silver-Russell and Cohen syndromes has been noted (2,21–23). In milder affected individuals, the short stature and the triangular shape of the face in late childhood could pose a differential diagnosis with Noonan syndrome (24).

In the context of a European collaborative study, we employed a combined array-based Comparative Genomic Hybridization (array-CGH), gene panel analysis and exome sequencing approach on the largest cohort of KBGS patients currently available ($n=49$) to establish new genotype–phenotype correlations, define novel clinical signs and characterize the clinical evolution of KBGS with ageing.

Results

We analysed a cohort of 49 individuals affected by KBGS with an equal distribution of male and female patients (25/24) (Supplementary Material, Table S1). We divided the cohort into four age-groups (0–4, 5–10, 11–20, ≥21 years) differentiated for the gender (Supplementary Material, Fig. S1).

Molecular characterization

Detection of *ANKRD11* variants was obtained by a combined exome sequencing, gene panel analyses and array-CGH. In most patients, we detected alterations only in the *ANKRD11* gene. In two patients genetic defects in additional genes were detected. All patients carried constitutive variants except for patient 22, who was found to carry the variant in a mosaic state.

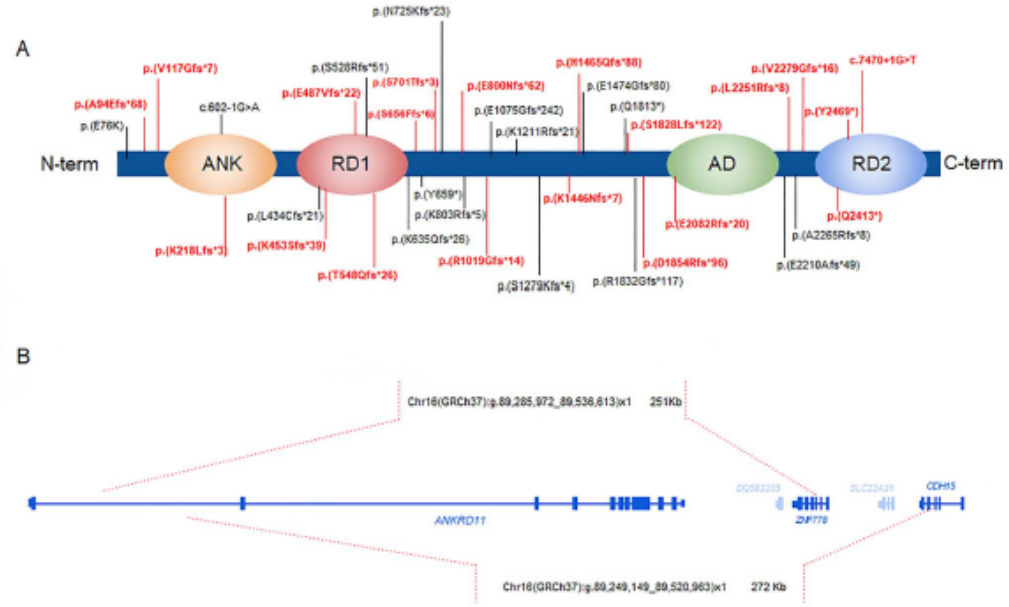


Figure 1. *ANKRD11* variants localization and genomic rearrangements (**A**) Genetic pathogenic variants identified in our study in the ANKRD11 protein, with a schematic representation of its domains (ANK: Ankyrin repeats (p.126–295); RD1 Repression domain 1 (p.318–611); AD: Activation domain (p.2076–2145); RD2: Repression domain 2 (p.2369–2663)). The novel mutations identified in the patients described in this study are indicated in red. (**B**) Genomic coordinates of the microdeletions (highlighted with red dashed lines) as identified by array-CGH in two KBGS patients.

Five patients inherited the alteration from an affected parent, whereas in 35 patients the variants were *de novo*. In nine patients segregation analysis of the variants was not available. Two twins shared the same pathogenic variant. We detected 36 different variants: one missense variant, indeed lying at the −1 position of the 5′ splice site and thus able to affect splicing, two splice site variants, 29 frameshift and four nonsense variants (Supplementary Material, Table S1 and Supplementary Material, Table S2). In 35 patients out of 49, variants were located in exon 10. Among the detected variants, 17 have been previously reported, whereas 19 are unpublished (Fig. 1A). We did not detect any specific genotype–phenotype correlation. Furthermore, array-CGH analysis revealed two microdeletions extending to the neighbouring Zinc Finger 778 gene (*ZNF778*), one segregating in an affected mother and her three affected children (patient 8, 9, 10), and another arising *de novo* in patient 6 (Fig. 1B).

Clinical features prevalence

The prevalence of the main clinical features is summarized in Figure 2. As expected, the most prevalent feature was mild ID. However, cerebral anomalies were found in 56% of patients (18/32) who underwent brain magnetic resonance imaging (MRI). Skeletal and dental anomalies were also frequent. Congenital heart defects were reported in 16.3% of cases.

Growth

At birth, all growth parameters tended to be lower than average; weight at birth had a median of −0.81 SD, length had a median of −0.65 SD and occipitofrontal circumference (OFC) had a median of −0.88 SD. With ageing, all auxological parameters, except for height, tended to normalize (Fig. 3). Weight reached a median of −0.645 SD between 11 and 21 years of age, and −0.48 SD above 20 years. OFC also increased to a median of +0.03 SD and +0.62 SD over time, with values in the range of normality above 21 years of age (*P*-value = 0.04 in the range ≥ 21 years versus 5–10 years). Persistent microcephaly was detected in 12% (6/49) of patients. Height did not improve with ageing with a deflection in the growth curve around the age of 10 and a median of −1.54 SD in the cohort above 21 years of age. The difference of the median height SD for patients ≥21 years versus the median length SD at birth was statistically significantly different (*P*-value = 0.0268). Delayed bone maturation was reported in five cases (patients 5, 17, 41, 48 and 49).

Dysmorphic features

In all patients, a typical facial gestalt was detected (Fig. 4). The main features include prominent forehead, brachycephaly, triangular face, synophrys, hypertelorism, long philtrum, thin upper lip, depressed and wide nasal bridge, anteverted nares and hypertrichosis.

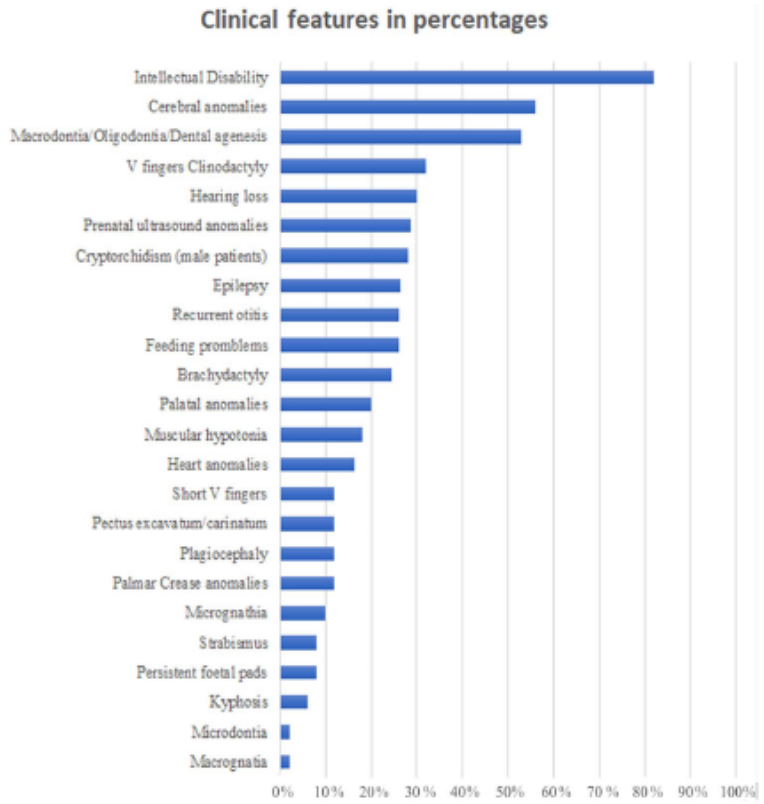


Figure 2. Clinical features distribution: The horizontal bars represent the different percentages (x axis) of each clinical feature (y axis) as observed in our KBGS patient cohort.

Facial dysmorphic features are summarized in Supplementary Material, Table S1 and illustrated in Figure 4.

Neurodevelopmental problems

In our cohort of patients, 82% (36/44) presented with ID which mostly ranged from mild to moderate (72.2% and 19.4%, respectively) in line with previous reports (25). Seven patients (patients 7, 8, 9, 13, 20, 26, 38 in Supplementary Material, Table S1) out of the 44 evaluable individuals (15,6%), did not display ID. Severe ID was identified only in two patients with concomitant genetic conditions (12,16), namely maternal 15q11.2 deletion and paternal inherited exostosis due to *EXT1* mutation, respectively. About 18% of patients presented hypotonia at birth or at the first evaluation. Epilepsy was present in 26.5% of patients (13/49) and mostly correlated with cerebral anomalies at the MRI. Indeed, out of 13 patients who developed seizures, eight presented brain anomalies, in three of them (12,25,26) no cerebral alteration was detected by brain imaging and in two of them no MRI was performed. Electroencephalography (EEG) often revealed mild non-specific background abnormalities with modest voltage asymmetry, slow background activity (patient 2, 8, 47) or an irregular brain activity during sleeping (patient 41). Generalized slow spikes and waves, generalized polyspikes and generalized paroxysmal fast activity during sleep were only registered in patient 12, a 34-year-old male, who displayed a refractory Lennox–Gastaut syndrome with severe ID and who indeed harboured a deletion of the 15q11.2 region including *TUBGCP5*, *CYFIP1*, *NIPA1*, *NIPA2*. In patient 45 slight pointed plurifocal anomalies in the left hemispheric area were already registered at 1 month of age. Behavioural problems were detected in 65% of individuals (32/49) with atrial septal defects (ASD) or attention deficit hyperactivity disorder (ADHD) reported in 30% of patients (15/49), tantrums and aggressive behaviour in 20% of patients (10/49), anxiety in 6% (3/49), emotional lability and shyness in 8.1% of individuals (4/48).

Cerebral anomalies

Cerebral abnormalities were identified in 56% of patients (18/32) who underwent brain MRI. Enlarged cisterna magna (including mega cisterna magna) was the most shared clinical finding among affected individuals being

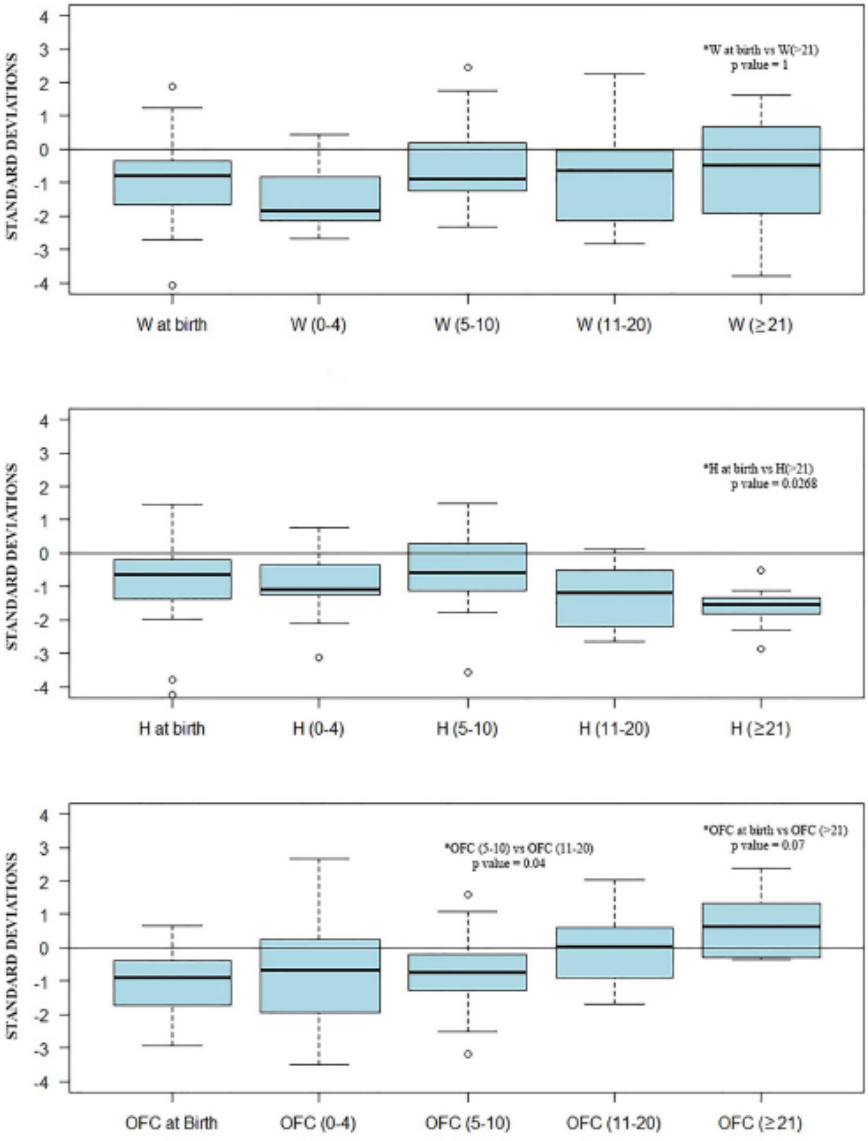


Figure 3. Auxological parameters distribution over time: Standard deviations for auxological parameters divided into age groups (0–4 years old, 5–10 years old, 11–20 years old and ≥21 years old). The statistically significant differences (P-value < 0.05) of the median standard deviations for the auxological parameters for the different demographic classes are indicated on the box plots.

detected by MRI in 23% of patients (6/32). Common clinical findings were also arachnoid cysts. Other brain anomalies include periventricular nodular heterotopia (PNH) (patients 10 and 13), short and thin corpus callosum (patients 26, 27 and 45), microcephaly (22% of patients, 11/49), cortical atrophy and dysplasia (patients 2 and 10), trigonocephaly (patient 17) (Fig. 6).

Dental and skeletal anomalies

A wide range of dental defects were observed in 53% (26/49) of cases: macrodontia, oligodontia, dental fusion, dental crowding and dental agenesis. Delayed dental eruption was observed in two patients. Patient 16 displayed macrodontia with a fusion of the maxillary lateral. Microdontia was observed in one patient. In 10%

Figure 4. Patient's clinical features: Clinical features of patients at different ages (**A**:18, **B**:15, **C**:42, **D**:13, **E**:48, **F–G**:9, **H–I**:14, **J–K**:44, **L**:40, **M–N**:33, **O–P**:36). Common clinical features are triangular face, long philtrum, depressed and wide nasal bridge, anteverted nares.

of patients, micrognathia was present, while only one patient presented macrognathia. Palatal anomalies were detected in 20% of cases (10/49) with 12% of patients (6/49) presenting with a highly arched palate. Cleft soft palate was described in one patient (patient 14) and velopharyngeal dysfunction in another patient (patient 19).

Skeletal anomalies including pectus excavatum, scoliosis, fifth finger clinodactyly, brachydactyly, vertebral malformations such as malformation of fifth and sixth thoracic vertebrae, were detected in 89% of cases (44/49). About 32% of patients (16/49) presented with fifth finger clinodactyly. Short fifth finger was also common (12%, 6/49), along with brachydactyly (24.4%, 12/49) and pectus excavatum/carinatum (12%, 6/49). Kyphotic posture was described in three patients (6%). Delayed bone maturation was reported in five cases (patients 5, 17, 41, 48 and 49). In patient 45 a reduced development of the growth nuclei of the bones and absence of the growth nucleus of the radius was observed. Large anterior fontanelle were detected in 10% of cases.

Hearing loss and recurrent otitis

Hearing loss was found in 30% (15/49) of patients in accordance with previous reports (5). Conductive hearing loss was present in nine of them, mixed hearing loss was found in two, one patient didn't pass otoacoustic emissions (OAE) and only in one case sensorineural hearing loss was reported (patient 2). In patient 29 conductive hearing loss in his left ear and sensorineural hearing loss in his right ear coexisted. Only three patients presented unilateral hearing loss (patient 6, 25 and 35).

Recurrent otitis were found in 26% (13/49) of patients. Of the 15 patients that presented hearing loss only half of them presented recurrent otitis (7/15, 46%).

Additional congenital anomalies

Anomalies of palmar creases were found in 12% of patients and they included decreased palmar creases in two cases (patients 2 and 6), deeper palmar crease in one case (patient 4) and a single palmar fold in three cases (patients 5, 40 and 45). Persistence of foetal pads was noted in 8% of cases (patient 4, 14, 16 and 28).

About 16% of patients (8/49) displayed heart defects including ASD, ventricular septal defects, atrioventricular defect and valve insufficiency. Bilateral or unilateral cryptorchidism was detected in 28% of patients (7/25). Vesicoureteral reflux was reported in patients 31, 32 and 48. About 8% of patients had strabismus (4/49). Hydrocele and double spleen were found, respectively, in patients 29 and 42. Anteriorly placed anus and anal atresia were detected in patient 9, two fistulas in the anal region emerged in patient 36 (Fig. 5). Lingual frenulum alterations were observed in four patients with one of them presenting longer than normal frenulum and the other three presenting shorter frenulum (two of them needed surgery). Feeding difficulties were present in 13/49 patients (26%) mostly consisting of low feeding activity, unable to breastfeed, discoordinated

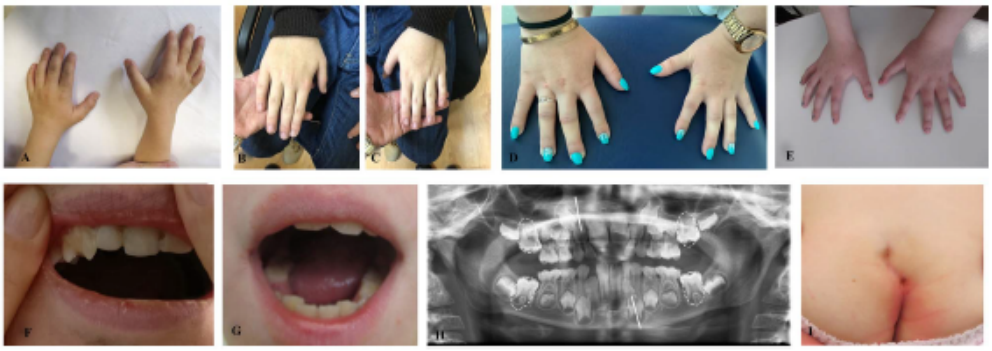


Figure 5. Other Clinical features: Brachydactyly patients 48 (**A**), 42 (**B–C**), 13 (**D**) and 44 (**E**), dental anomalies patients 16 (**F**) and 44 (**G**), mild eruption delay of permanent first molars and small size of the inferior lateral left incisor and of the superior lateral right incisor in patient 49 (**H**) and anal fistula patient 36 (**I**).

sucking-swallowing-breathing rhythm, difficulty in chewing or swallowing.

Antenatal ultrasound anomalies

In 28.5% of individuals (14/49), antenatal ultrasound anomalies were reported namely increased nuchal translucency, polyhydramnios and intrauterine growth restriction (IUGR).

Coexistence of two conditions

In patient 12, a maternal 15q11.2 microdeletion including *TUBGCP5*, *CYFIP1*, *NIPA1* and *NIPA2* genes was also detected. He presented with severe ID and major behavioural disorder. He also suffered from Lennox–Gastaut syndrome from age 5. At 33 years of age, language was limited to a few words.

In patient 16 a previously described pathogenic variant in *EXT1* gene was also identified. His skeletal phenotype was indeed more complex. He presented multiple exostoses in the context of a Madelung deformity, curved forearms and short upper limbs. He also presented with a severe ID. Considering that mutations of *EXT1* gene are not associated with ID, his more severe cognitive phenotype is probably a random association. In conclusion, the only two severe ID individuals collected in this study had a coexistence of KBGS with a second unrelated condition.

Discussion

Short stature, macrodontia and distinctive facial features are the most common clinical traits in KBGS (2). Even if none of the individual symptoms identified so far is a KBGS diagnostic requisite, the combination of them can lead to disease suspicion. In line with previous reports, our study reinforces the concept that general clinical features, which justify genetic investigation for KBGS, are developmental delay, recognizable facial gestalt (becoming more apparent with ageing) and skeletal/dental anomalies. In accordance with previous studies (5) hearing loss was present in about 30% of patients and feeding difficulties in about 26% (5). Height below average was present in 20/49 patients, while short stature (height < −2SD) was detected in 5/49 individuals. ID generally ranges from mild to moderate in line with previous reports (27) and seizures generally correlated with cerebral anomalies. The occurrence of cerebral anomalies suggests that performing a brain MRI might help reaching a diagnosis of KBGS. In contrast with previous reports (13), we did not find a correlation between ID severity and location of the *ANKRD11* variants because both mild and moderate ID were observed in patients with variants in the RD2 domain as well as in patients with variants lying in the region between RD1 and AD domains. We found only two individuals with severe ID that both presented with a concomitant genetic condition. Most of the adult patients with mild ID (26/49, 53% of cases) were able to attend high school with teaching aid. One of them is a kindergarten teacher and one patient is currently attending university. Behavioural issues were present in 65% of individuals with ASD and ADHD being the most prevalent finding followed by aggressive behaviour.

The finding of distinctive dysmorphisms is generally suggestive of a KBGS even if making a definitive diagnosis early in life is challenging because of milder features and the lack of some clinical signs such as dental anomalies and short stature (2). In order to provide useful criteria for clinical diagnosis, a presumptive scheme of KBGS natural history, showing the natural evolution of the major clinical findings over time, is provided in Figure 7. As recently underlined, atypical findings such as a wide anterior fontanelle accompanied by minor facial features, can be the only presenting features of KBGS in children (28). Enlarged fontanelles are mostly caused by a slow or incomplete closure of the skull bones (29). This finding, which we detected in about 10% of patients, is in line with the reports about *ANKRD11* role in mice in

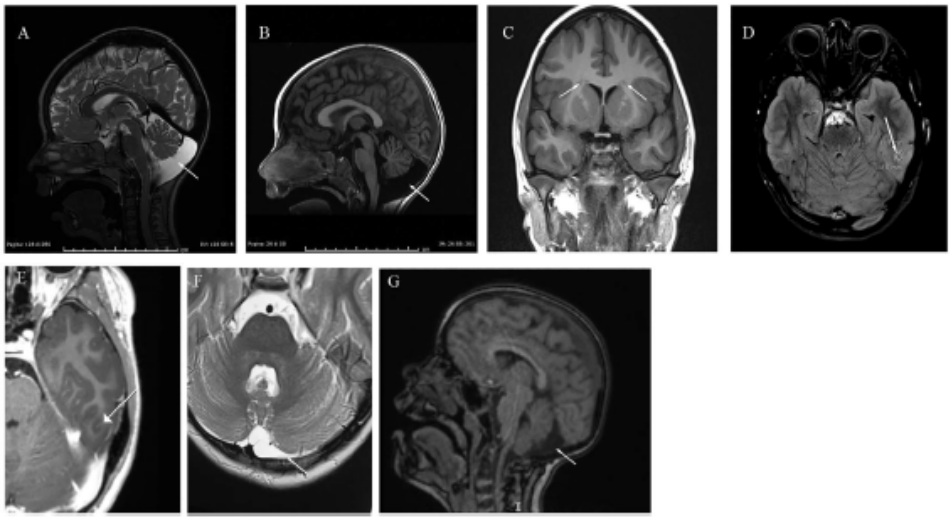


Figure 6. Brain MRI images. (**A**–**B**) Megacisterna Magna (patient 46); (**C**–**E**) Frontal heterotopy in the white matter near the lateral ventricles and in the temporo-basal region. T2 signal intensity suggestive of cortical dysplasia (patient 10); (**F**): Subarachnoid cyst (patient 16); (**G**) Retro-cerebellar cisterna expansion (patient 15).

which heterozygous deletion of *ANKRD11* in neural crest cells leads to reduced midfacial width and persistent open fontanelle (16).

Various brain abnormalities including cerebellar vermis hypoplasia (30), enlarged cisterna magna, Chiari I malformation, PNH, pineal cyst, dysgenesis of the corpus callosum, arachnoid cysts have been previously reported as single case reports. However, the frequency of brain malformations was not known because brain MRI has not been previously performed in large cohorts of affected individuals. Notably, being detected in more than 50% of the 32 patients who underwent brain MRI, our cohort suggests that cerebral anomalies are far more frequent in KBGS than previously estimated. We found an enlarged or a mega cisterna magna (greater than 10 mm) in 23% of patients. Furthermore, we report a series of unique clinical features such as trigonocephaly, small hypophysis and hypotrophy of the optic nerve, broadening the cerebral phenotypic spectrum of KBGS. Nevertheless, according to our data, we suggest considering a wider range of dental anomalies including microdontia before ruling out KBGS diagnosis. This clinical finding has been previously reported in another patient displaying maxillary canine microdontia (31). We also highlight the importance of performing a careful dental/ENT evaluation to look for even submucosal cleft palate given the high frequency of palate abnormalities.

In conclusion, broadening the diagnostic criteria can help the ascertainment of KBGS with an improvement of the clinical management and can facilitate differential diagnosis (25). In some individuals, KBGS symptoms overlap with other syndromes leading to a misdiagnosis. In patients 11 and 16, CdLS was suspected in the first years of life. A clinical diagnosis of Nicolaides–Baraitser syndrome was proposed for patient 14. Thus, consistent with previous reports, our study supports an age-dependent phenotypic evolution going from Coffin-Siris/Nicolaides Baraitser or CdLS to KBGS from infancy to adolescence. This also highlights the importance of clinical follow-ups to achieve the proper diagnosis (3). Differential diagnosis with RASopathies spectrum disorders is often raised as well, due to some common features, namely short stature and triangular shape of the face. Notably, an increased nuchal translucency, often observed in RASopathies, has been reported in the prenatal era in some of our KBGS patients, thus highlighting the importance of a wider targeted panel approach which should include *ANKRD11* in the context of prenatal diagnosis.

None of our patients developed cancer and thus we were not able to establish a correlation between *ANKRD11* loss of function variants and cancer predisposition, as postulated in previous studies (12). However, as the eldest patient in our cohort is 41 years old, we cannot exclude that KBGS patients have an increased cancer risk in adulthood. Notably, in patient 16 we were not able to establish the maternal origin for the pathogenic variant since the mother died in her forties due to myeloid acute leukaemia, a kind of tumour often related to TP53 pathway dysregulation. However, available iconographic documents strongly suggest that she also carried a pathogenic variant in *ANKRD11*. Thus, although at the moment no specific measures for cancer surveillance need to be recommended for KBGS patients, a systematic follow-up is necessary to collect additional useful information to clarify if there is a cancer risk in KBGS.

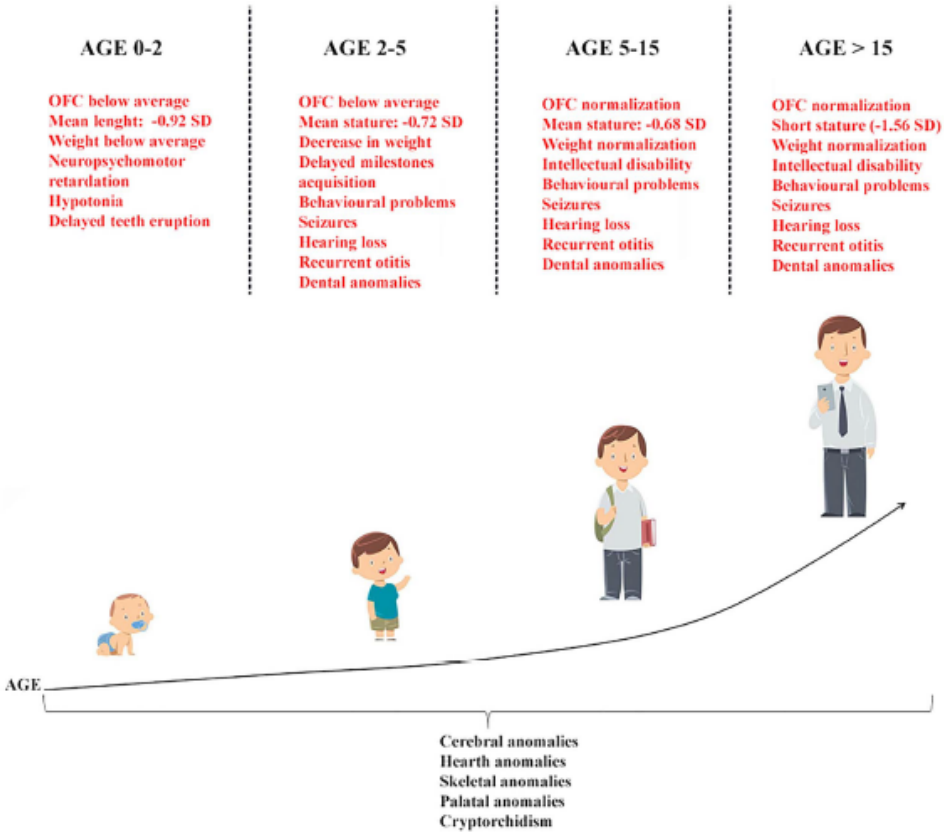


Figure 7. Natural history of KBGS phenotype: Graphic representation of the major clinical findings over time. Our cohort was divided into age groups. Clinical findings spanning all age groups are listed below the main figure. Clinical signs typical of each age group are indicated in red above the main figure.

According to previous studies, pathogenic *ANKRD11* variants were almost all exclusively frameshift, followed by rare reports of more rarely nonsense variants, confirming the haploinsufficiency as disease mechanism. Interestingly, in our cohort, the only apparent missense mutation lies at −1 position of the 5′ splice site that can affect mRNA processing (14). This finding emphasizes the importance of careful evaluation of *ANKRD11* missense variants as causative of KBGS and highlights that a splicing effect should be considered when clinical diagnosis is convincing. We cannot indeed exclude though that *ANKRD11* missense variants could be responsible for a distinct phenotype. In our cohort, we detected a 251 kb microdeletion involving the *ANKRD11* and *ZNF778* genes, segregating in a family with a phenotype partially distinct from KBGS. Recently, a microdeletion, causing haploinsufficiency of the *ANKRD11*-flanking genes, has been found in patients with a 16q24.3 microdeletion syndrome (32,33). Our patients presented KBGS-like features that overlapped with those of the 16q24.3 microdeletion syndrome (dyslalia, strabismus, mild ID and heterotopy), defining a 'KBGS plus phenotype'. We identified a mosaic *ANKRD11* variant in patient 22. A mosaic state has been established for a previously reported patient with a mild phenotype (26). Our patient with mosaic KBGS presented mild ID, ASD and minor facial dysmorphisms confirming that a milder phenotype is likely associated with *ANKRD11* mosaic variants.

In conclusion, our results demonstrate that the combined use of exome sequencing analysis and array-CGH can result in the diagnosis of KBGS, particularly in paediatric patients with immature clinical findings (34,35). Furthermore, our data underline the importance of considering additional signs, such as cerebral anomalies and suggest that anamnestic obstetric history, positive

for IUGR and/or increased nuchal translucency, supports KBGS diagnosis in the paediatric age.

Materials and Methods

Selection of patients and DNA samples' preparation

For each patient clinical evaluation and genetic counselling was performed in order to collect family history and better define facial and physical characteristics. In all patients audiological examination was performed to check for hearing loss as well as EEG to look for parossistic activities. ENT evaluation followed the suspicion of palatal anomalies and/or velopharyngeal dysfunction, brain MRI was performed in 32/49 patients (65%). All patients or their parents gave their written informed consent to the study that was carried out according to the Declaration of Helsinki.

Genomic DNA was extracted, for some samples, from EDTA peripheral blood samples using MagCore HF16 (Diatech Lab Line), according to the manufacturer's instructions. The DNA quantity was estimated using the NanoDropTM 2000/2000c Spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA).

For other samples DNA was extracted from peripheral blood using the standard salting out method or the QiaSymphony DSP DNA midi kit. The DNA quantity was estimated using the NanoDropTM 1000 Spectrophotometer (Thermo Fisher Scientific).

Whole-Genome array-CGH

High-resolution whole-genome (array-CGH) analysis was performed on genomic DNA of patients 6, 8, 9, 10 in Supplementary Material, Table S1, using the SurePrint G3 Human CGH Microarray 8 × 60 k (Agilent Technologies, Santa Clara, CA, USA), a dual-colour array containing 60-mer high-quality probes with 41 Kb genome-wide median probe spacing. Copy Number Variants (CNVs) were analysed and mapped using the Human Genome Assembly GRCh37/hg19.

Slides were scanned using an Agilent G2565CA Microarray Scanner (Agilent Technologies, Santa Clara, CA, USA) and processed using Feature Extraction software (v10.5.1.1). Results were analysed using Agilent CytoGenomics software (v5.1) with default settings. The results included imbalances with at least three consecutive probes with abnormal log2 ratios. The Database of Genome Variants (DGV- http://dgv.tcag.ca/dgv/app/home;), DECIPHER (DatabasE of Chromosomal Imbalances and Phenotypes using Ensembl Resources- https://www.deciphergenomics.org/), PubMed (https://pubmed.ncbi.nlm.nih.gov;), UCSC genome browser (https://genome.ucsc.edu;), Database of Human CNVs (http://gvarianti.homelinux.net/gvariantib37/index.php;), SFARI (Simon's Foundation Autism Research Initiative) Gene Database (https://gene.sfari.org;) and OMIM (Online Mendelian Inheritance in Man- https://www.omim.org/;) databases were used for the interpretation of the results. Each DNA sample was analysed twice through array-CGH, in order to confirm the result.

Exome sequencing and data analysis

For most patients, exome sequencing was performed using the Life Technologies Ion Proton sequencer (Life Technologies, Carlsbad, CA, USA) on genomic DNA samples of proband and parents, when available. This system enables >92% of bases covered 20X. Sample preparation and sequencing were performed with AmpliSeq™Exome strategy, following the manufacturer's protocol (Life Technologies). The library preparation was performed using the Ion AmpliSeq™Exome Kit (Life Technologies), which allows us to target ~33 Mb of coding exons plus 15 Mb of flanking regions for a total of ~58 Mb, in total more than 97% of the coding regions described by Consensus Coding Sequences (CCDS) annotation. In total 12 primer pools for highly specific enrichment of exons within the human genome were used. Taking advantage of a barcode system, three samples were loaded together in a single run and sequenced. Data analysis was performed with Torrent Suite™ Software v3.6.2 (Life Technologies). The provider generated at least 30 effective mean depths per sample. Using specific parameters, we were able to remove the adaptors' contamination and low-quality sequences, so the total amount of clean data was mapped to the UCSC/hg19 reference genome. Indel and variant calls were made using GATK version 2.7 (Broad Institute, Cambridge, MA, USA) (and its recommended parameters) and then the variants were also annotated against external datasets, including 1000 genomes and dbSNP (36).

For Finnish patients an NGS-based clinical trio exome was performed with Sophia Genetics custom clinical exome solution and Illumina sequencing. Variant annotation and filtering were performed with Sophia DDM™ and visualized with IGV (Broad Institute).

Prioritization of the variants was obtained excluding polymorphisms (minor allele frequency, MAF <0,01), synonymous variants, variants classified as benign or likely benign. Frameshift, stopgain and splice site variants were prioritized as pathogenic. The potential impact of variants on splicing was evaluated using Alamut® Visual software—version 2.11–0 (Interactive Biosoftware, France), which employs five different algorithms: SpliceSiteFinder-like, MaxEntScan, NNSPLICE, GeneSplicer and HumanSplicingFinder.

The following public databases were used for the interpretation of the variants: ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), LOVD (https://databases.lovd.nl/shared/genes), the Human Genome Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac). Each disease-candidate variant was confirmed by Sanger sequencing.

Additional genetic tests

Estonian patients (Supplementary Material, Table S1 patients 29–40) were investigated by TruSight One (TSO) and TruSight One Expanded (TSOE) panels (Illumina

Inc., San Diego, California). These panels cover ~4800 and ~6700 genes associated with known genetic disorders or clinical phenotypes. A detailed description of the method was published earlier (37). Detected variants were validated by Sanger sequencing.

In patients 17 and 18 in Supplementary Material, Table S1, a targeted panel approach including *ANKRD11* was employed to ascertain the disease. The targeted panel was realized via Next Generation Sequencing on the 454 Junior platform (Roche) with specific amplification. Analysis was carried out using the GS Amplicon Variant Analyzer (AVA) application version 2.9 (Roche). In alternative amplification and high-throughput sequencing of *ANKRD11* coding regions and exon-intron junctions was obtained by the NimbleGen SeqCapEZ Custom Enrichment Kit (Roche) and Illumina platform Nextseq550 (sensitivity >99%). Results interpretation focused on exonic regions with a read depth >20X.

Informed Consent Statement

The patients/participants or their parents provided their written informed consent to participate in this study.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We are grateful to our patients for their cooperation. This work is generated within the European Reference Network for Rare Malformation Syndromes, Intellectual and Other Neurodevelopmental Disorders (ERN ITHACA). 'Cell lines and DNA bank of Rett Syndrome, X-linked mental retardation and other genetic diseases', member of the Telethon Network of Genetic Biobanks (project no. GTB12001), funded by Telethon Italy, and of the EuroBioBank network, provided us with specimens. This work has also been partially supported by grant-RC Linea 1'Studio fenotipo-genotipo delle malattie genetiche rare ad espressione neuropsichiatrica in età evolutive' and the 5x 1000 voluntary contributions, Italian Ministry of Health (R.B.).

Conflict of Interest statement. The authors declare that they have no conflict of interest.

Funding

K.Õ., T.R., K.M. and L.R. were supported by Estonian Research Council grant PRG471.

Authors Contribution

Author Contributions: L.B. and L.L. made important contributions in the interpretation of the molecular results and drafted the manuscript. A.M.P. and A.R. made substantial contributions to the conception and design of the study and reviewed the manuscript. All co-authors performed clinical data acquisition, clinical evaluations and variants interpretation. All authors approved the final version of the manuscript for publication.

References

1. Herrmann, J., Pallister, P.D., Tiddy, W. and Opitz, J.M. (1975) The KBG syndrome-a syndrome of short stature, characteristic facies, mental retardation, macrodontia and skeletal anomalies. *Birth Defects Orig. Artic. Ser.*, **11**, 7–18.
2. Morel Swols, D., Foster, J. and Tekin, M. (2017) KBG syndrome. *Orphanet J. Rare Dis.*, **12**, 183.
3. Parenti, I., Mallozzi, M.B., Hüning, I., Gervasini, C., Kuechler, A., Agolini, E., Albrecht, B., Baquero-Montoya, C., Bohring, A., Bramswig, N.C. *et al.* (2021) ANKRD11 variants: KBG syndrome and beyond. *Clin. Genet.*, **100**, 187–200.
4. Ropers, H.H. and Wienker, T. (2015) Penetrance of pathogenic mutations in haploinsufficient genes for intellectual disability and related disorders. *Eur. J. Med. Genet.*, **58**, 715–718.
5. Low, K., Ashraf, T., Canham, N., Clayton-Smith, J., Deshpande, C., Donaldson, A., Fisher, R., Flinter, F., Foulds, N., Fryer, A. *et al.* (2016) Clinical and genetic aspects of KBG syndrome. *Am. J. Med. Genet. A*, **170**, 2835–2846.
6. Wojciechowska, K., Nurzyńska-Flak, J., Styka, B., Kacprzak, M. and Lejman, M. (2021) Case Report: Two Newly Diagnosed Patients With KBG Syndrome—Two Different Molecular Changes. *Front. Pediatr.*, **9**, 9.
7. Gnazzo, M., Lepri, F.R., Dentici, M.L., Capolino, R., Pisaneschi, E., Agolini, E., Rinelli, M., Alesi, V., Versacci, P., Genovese, S. *et al.* (2020) KBG syndrome: Common and uncommon clinical features based on 31 new patients. *Am. J. Med. Genet. A*, **182**, 1073–1083.
8. Sirmaci, A., Spiliopoulos, M., Brancati, F., Powell, E., Duman, D., Abrams, A., Bademci, G., Agolini, E., Guo, S., Konuk, B. *et al.* (2011) Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am. J. Hum. Genet.*, **89**, 289–294.
9. Sacharow, S., Li, D., Fan, Y.S. and Tekin, M. (2012) Familial 16q24.3 microdeletion involving ANKRD11 causes a KBG-like syndrome. *Am. J. Med. Genet. A*, **158A**, 547–552.
10. Zhang, A., Yeung, P.L., Li, C.-W., Tsai, S.-C., Dinh, G.K., Wu, X., Li, H. and Chen, J.D. (2004) Identification of a novel family of ankyrin repeats containing cofactors for p160 nuclear receptor coactivators. *J. Biol. Chem.*, **279**, 33799–33805.
11. Goldenberg, A., Riccardi, F., Tessier, A., Pfundt, R., Busa, T., Cacciagli, P., Capri, Y., Coutton, C., Delahaye-Duriez, A., Frebourg, T. *et al.* (2016) Clinical and molecular findings in 39 patients with KBG syndrome caused by deletion or mutation of ANKRD11. *Am. J. Med. Genet. A*, **170**, 2847–2859.
12. Behnert, A., Auber, B., Steinemann, D., Frühwald, M.C., Huisinga, C., Hussein, K., Kratz, C. and Ripperger, T. (2018) KBG syndrome patient due to 16q24.3 microdeletion presenting with a paratesticular rhabdoid tumor: Coincidence or cancer predisposition? *Am. J. Med. Genet. A*, **176**, 1449–1454.
13. Li, Q., Sun, C., Yang, L., Lu, W. and Luo, F. (2021) Comprehensive analysis of clinical spectrum and genotype associations in Chinese and literature reported KBG syndrome. *Transl. Pediatr.*, **10**, 834–842.
14. Ockeloen, C.W., Willemsen, M.H., de Munnik, S., van Bon, B.W.M., de Leeuw, N., Verrips, A., Kant, S.G., Jones, E.A., Brunner, H.G., van Loon, R.L.E. *et al.* (2015) Further delineation of the KBG syndrome

phenotype caused by ANKRD11 aberrations. *Eur. J. Hum. Genet. EJHG*, **23**, 1176–1185.
15. Palumbo, O., Palumbo, P., Di Muro, E., Cinque, L., Petracca, A., Carella, M. and Castori, M. (2020) A Private 16q24.2q24.3 Microduplication in a Boy with Intellectual Disability, Speech Delay and Mild Dysmorphic Features. *G. E. N.*, **11**, 707.
16. Roth, D.M., Baddam, P., Lin, H., Vidal-García, M., Aponte, J.D., De Souza, S.-T., Godziuk, D., Watson, A.E.S., Footz, T., Schachter, N.F. *et al.* (2021) The Chromatin Regulator Ankrd11 Controls Palate and Cranial Bone Development. *Front. Cell Dev. Biol.*, **9**, 645386.
17. Walz, K., Cohen, D., Neilsen, P.M., Foster, J., Brancati, F., Demir, K., Fisher, R., Moffat, M., Verbeek, N.E., Bjørgo, K. *et al.* (2015) Characterization of ANKRD11 mutations in humans and mice related to KBG syndrome. *Hum. Genet.*, **134**, 181–190.
18. Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H. *et al.* (2017) Correction: Corrigendum: High-throughput discovery of novel developmental phenotypes. *Nature*, **551**, 398–398.
19. Neilsen, P.M., Cheney, K.M., Li, C.-W., Chen, J.D., Cawrse, J.E., Schulz, R.B., Powell, J.A., Kumar, R. and Callen, D.F. (2008) Identification of ANKRD11 as a p53 coactivator. *J. Cell Sci.*, **121**, 3541–3552.
20. Crippa, M., Rusconi, D., Castronovo, C., Bestetti, I., Russo, S., Cereda, A., Selicorni, A., Larizza, L. and Finelli, P. (2015) Familial intragenic duplication of ANKRD11 underlying three patients of KBG syndrome. *Mol. Cytogenet.*, **8**, 20.
21. Cucco, F., Sarogni, P., Rossato, S., Alpa, M., Patimo, A., Latorre, A., Magnani, C., Puisac, B., Ramos, F.J., Pié, J. *et al.* (2020) Pathogenic variants in EP300 and ANKRD11 in patients with phenotypes overlapping Cornelia de Lange syndrome. *Am. J. Med. Genet. A*, **182**, 1690–1696.
22. Scarano, E., Tassone, M., Graziano, C., Gibertoni, D., Tamburrino, F., Perri, A., Gnazzo, M., Severi, G., Lepri, F. and Mazzanti, L. (2019) Novel Mutations and Unreported Clinical Features in KBG Syndrome. *Mol. Syndromol.*, **10**, 130–138.
23. Ahmed, A., Mufeed, A., Ramachamparambathu, A.K. and Hasoon, U. (2016) Identifying Aarskog Syndrome. *J. Clin. Diagn. Res. JCDR*, **10**, ZD09–ZD11.
24. Sayed, I.S.M., Abdel-Hamid, M.S. and Abdel-Salam, G.M.H. (2020) KBG syndrome in two patients from Egypt. *Am. J. Med. Genet. A*, **182**, 1309–1312.
25. Lo-Castro, A., Brancati, F., Digilio, M.C., Garaci, F.G., Bollero, P., Alfieri, P. and Curatolo, P. (2013) Neurobehavioral phenotype observed in KBG syndrome caused by ANKRD11 mutations. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **162**, 17–23.
26. Khalifa, M., Stein, J., Grau, L., Nelson, V., Meck, J., Aradhya, S. and Duby, J. (2013) Partial deletion of ANKRD11 results in the KBG phenotype distinct from the 16q24.3 microdeletion syndrome. *Am. J. Med. Genet. A*, **161**, 835–840.
27. Alfieri, P., Caciolo, C., Lazzaro, G., Menghini, D., Cumbo, F., Dentici, M.L., Digilio, M.C., Gnazzo, M., Demaria, F., Pironi, V. *et al.* (2021) Cognitive and Adaptive Characterization of Children and Adolescents with KBG Syndrome: An Explorative Study. *J. Clin. Med.*, **10**, 1523.
28. Kutkowska-Kaźmierczak, A., Boczar, M., Kalka, E., Castañeda, J., Klapecki, J., Pietrzyk, A., Barczyk, A., Malinowska, O., Landowska, A., Gambin, T. *et al.* (2021) Wide Fontanels, Delayed Speech Development and Hoarse Voice as Useful Signs in the Diagnosis of KBG Syndrome: A Clinical Description of 23 Cases with Pathogenic Variants Involving the ANKRD11 Gene or Submicroscopic Chromosomal Rearrangements of 16q24.3. *G. E. N.*, **12**, 1257.
29. Kiesler, J. and Ricer, R. (2003) The Abnormal Fontanel. *Am. Fam. Physician*, **67**, 2547–2552.
30. Zollino, M., Battaglia, A., D'Avanzo, M.G., Della Bruna, M.M., Marini, R., Scarano, G., Cappa, M. and Neri, G. (1994) Six additional cases of the KBG syndrome: clinical reports and outline of the diagnostic criteria. *Am. J. Med. Genet.*, **52**, 302–307.
31. Parloir, C., Fryns, J.P., Deroover, J., Lebas, E., Goffaux, P. and van den Berghe, H. (1977) Short stature, craniofacial dysmorphism and dento-skeletal abnormalities in a large kindred. A variant of K.B.G. syndrome or a new mental retardation syndrome. *Clin. Genet.*, **12**, 263–266.
32. Willemsen, M.H., Fernandez, B.A., Bacino, C.A., Gerkes, E., de Brouwer, A.P., Pfundt, R., Sikkema-Raddatz, B., Scherer, S.W., Marshall, C.R., Potocki, L. *et al.* (2010) Identification of ANKRD11 and ZNF778 as candidate genes for autism and variable cognitive impairment in the novel 16q24.3 microdeletion syndrome. *Eur. J. Hum. Genet.*, **18**, 429–435.
33. Novara, F., Rinaldi, B., Sisodiya, S.M., Coppola, A., Giglio, S., Stanzial, F., Benedicenti, F., Donaldson, A., Andrieux, J., Stapleton, R. *et al.* (2017) Haploinsufficiency for ANKRD11-flanking genes makes the difference between KBG and 16q24.3 microdeletion syndromes: 12 new cases. *Eur. J. Hum. Genet.*, **25**, 694–701.
34. Clark, M.M., Stark, Z., Farnaes, L., Tan, T.Y., White, S.M., Dimmock, D. and Kingsmore, S.F. (2018) Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom. Med.*, **3**, 16.
35. Niguidula, N., Alamillo, C., Shahmirzadi Mowlavi, L., Powis, Z., Cohen, J.S. and Farwell Hagman, K.D. (2018) Clinical whole-exome sequencing results impact medical management. *Mol. Genet. Genomic Med.*, **6**, 1068–1078.
36. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. *et al.* (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, **20**, 1297–1303.
37. Pajusalu, S., Kahre, T., Roomere, H., Murumets, Ü., Roht, L., Simenson, K., Reimand, T. and Õunap, K. (2018) Large gene panel sequencing in clinical diagnostics-results from 501 consecutive cases. *Clin. Genet.*, **93**, 78–83.

2.5 Additional analysis and collaborations

In the context of the exome analysis, we collected 101 positive cases out of a selection of 330 ID/ASD patients, 49 of those suffering from KBGS.

From a molecular point of view, we detected a statistically significant difference between male and female patients for non-disruptive variants and disruptive variants in the ID/ASD genes and in new candidate ID/ASD genes, with the first being enriched in the males and the latter found more abundant in the females (ratio variants non-disruptive/variants disruptive) (fig.1). From a clinical point of view, most of the female patients (6/8) showing missense PVs in syndromic genes suffered from less symptoms and displayed more non-syndromic conditions with regards to the male patients with the same kind of variant.

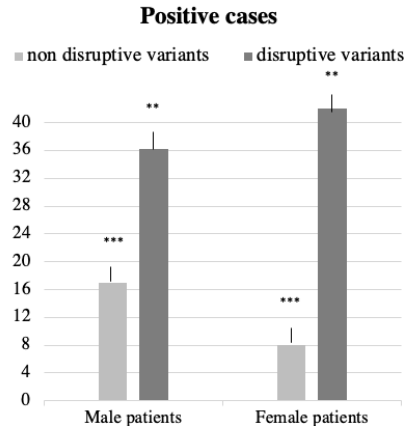


Figure 1: Number of disruptive and of non-disruptive variants in the positive cases.

We collaborated to identify a novel ID gene: *CAPRIN1*. *CAPRIN1* encodes a ubiquitous protein implicated in the transport and translation of neuronal mRNAs critical for synaptic plasticity, and of mRNAs encoding proteins required for cell proliferation and migration in multiple cell types.

The subsequent clinical features were reported for 12 cases with LoF variants in *CAPRIN1:* language delay (100%), ID (83%), ADHD (82%), ASD (67%), respiratory problems (50%), limb/skeletal anomalies (50%), developmental delay (42%) feeding difficulties (33%), seizures (33%) and ophthalmologic problems (33%). In patient-derived lymphoblasts and fibroblasts, a monoallelic expression of the wild- type allele, and a reduction of the transcript and protein consistent with a half dose emerged. *CAPRIN1*+/- human iPSCs (hiPSCs) were differentiated into neuronal progenitor cells and cortical neurons. *CAPRIN1* loss caused reduced neuronal processes, overall disruption of the neuronal organisation, an increased neuronal degeneration, and an alteration of mRNA translation. *CAPRIN1*+/- neurons also

showed an impaired calcium signalling and increased oxidative stress, both possibly impairing neuronal network development and functionality. Measurements of activity in *CAPRIN1*+/- neurons via micro-electrode arrays (MEA) indicated lower spike rates and bursts, with an overall reduced activity[2].

By WES we detected the novel acceptor splice site variant c.26-2A>G (falling in intron 2 of *CNOT3*, in correspondence to the donor splicing site of exon 3, in a family with syndromic ID. The proband showed ID, language delay, structural cerebral anomalies, physical dysmorphisms and cardiac defects whereas his mother manifested milder clinical traits and shared peculiar facial traits with his son (broad forehead, frontal drafts, long palpebral fissures, prominent columella, anteverted nares, short philtrum, thin lips, and micrognathia). According to the ACMG and to the bioinformatic splicing prediction tools the variant was predicted to be pathogenic leading to an alteration on the canonical splice site at 3'end. Sanger sequencing demonstrated the retention of intron 2, and showed from the beginning of intron retention, an amplification of exons 1-2, reconnecting to exon 3 without altering the reading frame. We reported the case of a girl affected by a composite developmental disorder, predominantly characterised by ID, language disorder and dyspraxia. Her urinary glycosaminoglycans (GAGs) dosages were above the normal values, raising the suspect of lysosomal storage disorder. WES sustained her diagnosis of Mucopolysaccharidosis IIIB.

We finally documented the finding by WES of two independent genetic syndromes due to *SMARCC2* and *RAF1* defects in two children and their mother suffering from NDDs. The first daughter carried a novel PV (p.(Ala1095Glnfs*)) in the *SMARCC2* gene, recently associated with Coffin-Siris syndrome 8. We reported down slanting palpebral fissures, short philtrum, downturned mouth, pointed chin, prominent and simplified ears, and arteriovenous malformation as novel characteristics of the CSS 8[3]. Her mother, instead, held the PV (p.(Thr661Met)) in *RAF1*. Finally, the second daughter showed both the PVs in *SMARCC2* and in *RAF1*.

2.6 Site-directed mutagenesis on *NLGN4X*

In this section we illustrate a site-specific mutagenesis approach that we performed with the aim of investigating *NLGN4X*-missense variants acting according to a likely GoF mechanism. We collected 4 unreported missense variants from an international collaboration (G174E, R232W, R437Q, R338K) and 5 VUS (variant of unknown significance) obtained from the

scientific literature (L406I, V371I, V522M, R753S, R583Q). All the variants fell in the extracellular portion of the Neuroligin 4 X-linked protein apart for the mutation R753S.

Patient 1 -Mutation R338K

The mutation (p.(Arg338Lys)) was shared by many members of a family showing overall non syndromic ASD and ID clinical features.

Patient 2 -Mutation R232W

The patient 2, a 22-year-old young man, inherited the novel PV c.694C>T, (p.(Arg232Trp)) from his mother. The following series of problems was reported: short stature, loss of ambulatory ability, interdigital webbing with abnormal hand creases, left preauricular pit, possibly familial, severe ID with ASD, intractable epilepsy, sensorineural bilateral hearing loss, arachnoid cyst, dysmorphic craniofacial features, overweight, paediatric, BMI corresponding to the 85.0-94.9 percentile for age, gastrostomy tube dependence, exotropia.

Patient 3 -Mutation R437Q

The patient 3 was a 2 years and 6 months old child born from eutocic birth to term of normal spontaneous pregnancy. Regular weaning was reported at 6 months of age.

He acquired the sitting station at 6 months, and he started to walk autonomously at 15 months. His lallation began in time, but he never pronounced entire words. From the age of 15 months, he showed poor interest throughout the others with absent speech, stereotypic movements, and food selectiveness. He was diagnosed with ASD, absent language, and showed no physical and facial peculiarities. ES identified the variant c.1310G>A (p.(Arg437Gln)) rs1377267584 (AF in gnomAD 0.0010%). He inherited the variant from her unaffected mother.

Patient 4-Mutation G174E

The patient 4 was a 5-year-old boy displaying the subsequent symptoms: global severe developmental delay, seizures. He did not receive a formal diagnosis of ASD. He showed microcephaly with lissencephaly and diffuse brain abnormality in old MRI, spastic quadriparesis with more recent increase in muscle spasticity, bilateral profound sensorineural hearing loss, bilateral optic atrophy with macular scars and vertical nystagmus, feeding disorder and G-tube placement, respiratory muscle weakness with poor secretion clearance and aspiration pneumonias, obstructive sleep apnea, recurrent bronchitis, and asthma, short stature, joint contractures at elbows and knees, scoliosis, dysmorphic features including ptosis and large low set ears. He underwent CMA with normal outcome. The trio ES highlighted the

hemizygous variant of uncertain significance c.521G>A, (p.(Gly174Glu)), inherited from his unaffected mother.

*Patient 5 -Mutation R649Lfs*2*

The proband was a 12-year-old male patient with past medical history of ASD, ID, and intermittent explosive disorder. On objective examination, he was overall non dysmorphic. Prior genetic testing revealed a frameshift mutation in *NLGN4X*, c.1946_1947delAC (p.(His649Leufs*2)) that he inherited from his mother which likely explained his neurodevelopmental presentation. His behaviour was difficult to control.

To sum up, bilateral profound sensorineural hearing loss, short stature and gastrostomy-tube dependence appeared as shared clinical features not reported before for the *NLGN4X*-phenotype.

First, to get an idea of the effects of our variants on the encoded protein, we analysed them with HOPE and Phyre2 tools, which predicted damage effects in the proteins to different extent. A structural damage was predicted by Phyre2 for R338K, caused by the replacement of a buried hydrophobic residue with a hydrophilic one.

The variant G174E triggered disallowed phi/psi alert, leading to phi/psi angles in the outlier region for the mutant residues. Finally, the substitution of arginine with a glutamine in position 437 replaced a buried charged residue with an uncharged one. The substitution also disrupted all side-chain/ side-chain H-bonds formed by a buried residue. Phyre2 did not report structural damage for the variants R232W, V371I, R583Q, L406I, V522M but an evaluation of the detrimental effects was obtained by HOPE. The residue changes of the variants R232W and V371I were located on the surface of the protein, and mutation of these residues could disturb interactions with other molecules or other parts of the protein. The missense variant R583Q lay in a domain that is important for binding of other molecules. The mutated residues were in contact with residues in another domain, and these contacts could be disturbed. L406I, V522M were in a region that was annotated to form an α-helix. The mutations converted the wild-type residues into new ones disrupting the α-helices.

Secondly, site-directed mutagenesis was performed, followed by molecular imaging on Neuro2A or murine neuroblastoma cell lines. Mutant's cell morphology was compared to the one of the *NLGN4X* wild type cells and to cells with plasmids containing GFP only. The cells transfected with the *NLGN4X* plasmid developed processes in >70% of the cases whereas the

7% of the cells containing only GFP manifested processes. We were unable to reproduce the mutation G174E, that resulted not vital in bacteria.

We stained the endoplasmic reticulum (ER) with an anti-RTN4 antibody to highlight a possible retention of NL4 in it. The RTN4 (Reticulon 4) gene belongs to the family of reticulon encoding genes. Reticulons are associated with the endoplasmic reticulum and are involved in neuroendocrine secretion or in membrane trafficking in neuroendocrine cells. Reticulons are required to induce the formation and stabilisation of ER tubules[4]. Neither the GFP signal of wild type cells nor of the mutant cells overlapped to the one of the anti-RTN4, with exception of the cells carrying the mutation R232W (fig.2 A-H).

For the variant R232W a phenotypic regression to an undifferentiated stage was observed with cells showing neuritis only in 15% of the cases and an overall reduction of the area's distribution of the neuroligin 4 X-linked. Morphologically, the Neuro2A mutant cells revealed an overall decrease in the neuritis length and a general decrease of membrane's area distribution of the neuroligin 4 X-linked compared to the wt cells, suggesting different pathogenic mechanisms (fig. 3). Due to the various Neuro 2A cell morphology, no differentiation between dendrites and axons was made.

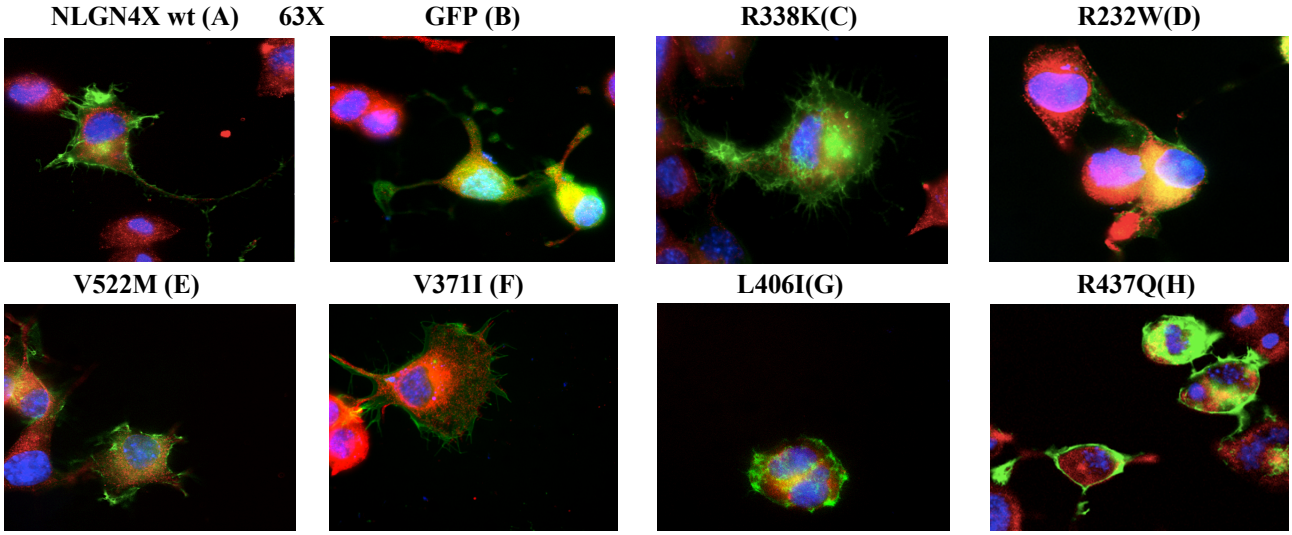


Figure 2 (A-H): Images at electron microscopy of Neuro2A wt and mutant cells containing GFP-conjugated plasmids, stained with the anti-RTN4 antibody. The staining did not reveal an overlap between the green signal by GFP and the red signal (produced by anti-RTN4) in the transfected cells, except for the R232W mutation.

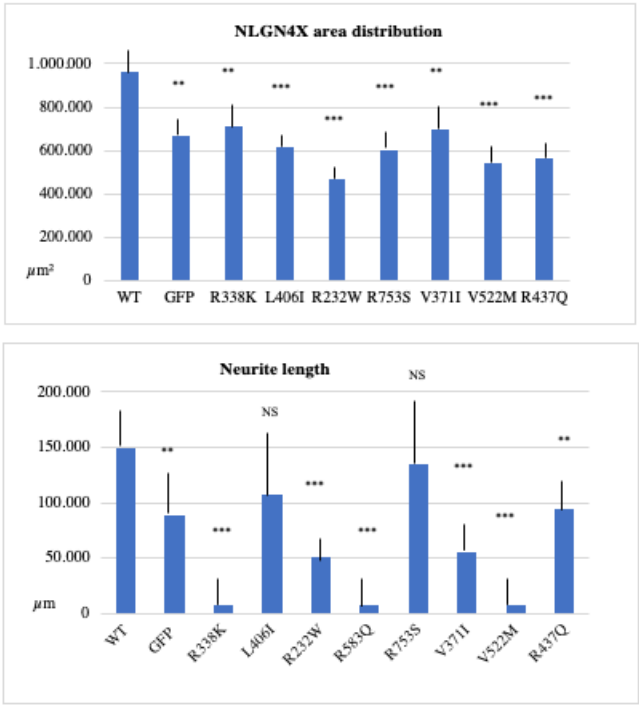
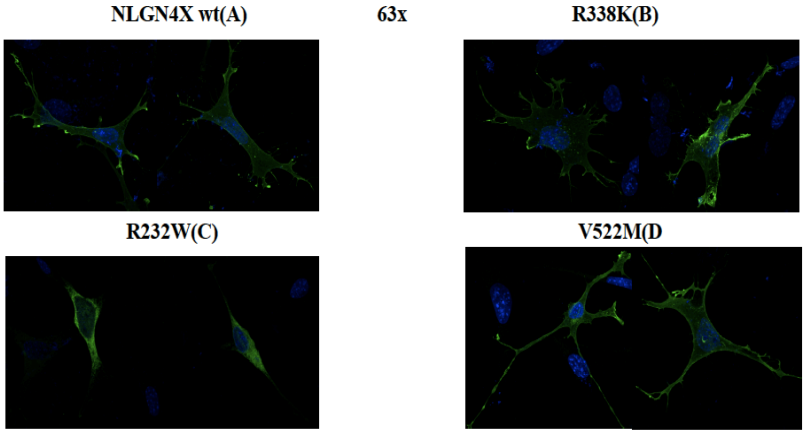


Figure 3: Graphs of the differences between Neuro2A wt and mutant cells in the area distribution and the neurite length (the point stands for a comma). 30 cells were evaluated for each plasmid at time. The experiments and the analysis were repeated three times. An overall decrease in the neuritis' length and in the distribution of the neuroligin 4 X-linked protein was detected.

To deepen our knowledge regarding the effect of the variants on the actin organisation on growth cone, we proceeded with a confocal microscopy examination on GN11 mouse neuronal cell line. For cells containing the plasmid WT, GFP, L406I and R232W, actin was visualised (fig.5). These photos confirmed a reduction of the membrane distribution and of the number of cell processes in the cells carrying the mutation R232W(fig.4-5).

We hypothesised this could be explained as a retention of the neuroligin 4 X-linked in the endoplasmic reticulum. Furthermore, a decrease in the number of the actin filaments was apparent in the mutant cells.



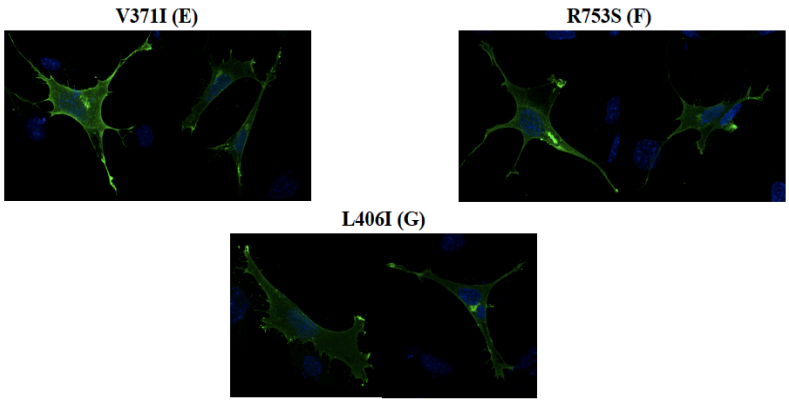


Figure 4 (A-G): Images at confocal microscopy of the GN11 cells transfected with wt and mutant plasmids. An overall decrease in the length of the processes and in the area distribution of the neuroligin 4 X-linked was observed. A simplified morphology was apparent for cells carrying the R232W mutation.

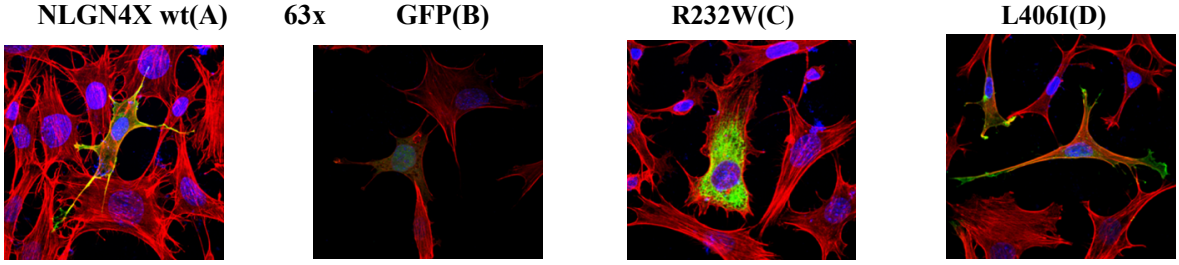


Figure 5 (A-D): Images at confocal microscopy of the GN11 cells transfected with wt and mutant plasmids with actin marked in red. A decrease in the number of the actin filaments was noted in the mutant cells, and a reduced area of distribution for the neuroligin 4 X-linked protein was visualised for the mutation R232W.

We decided to get insight into the differences in protein expression in the mutant cells with respect to the wt ones by performing a Western blot. Neuro2A cells were transfected with *NLGN4X-GFP* (for the WT, R437Q, R338K, R232W, R583Q, R753S e V371I) and with a plasmid containing RFP protein to compare the efficiency of transfection in the mutant cells. Bands signals were normalised choosing Glyceraldehyde 3 Phosphate Dehydrogenase (GAPDH) as housekeeping gene. Analysis of the optical density revealed an increased expression of the neuroligin 4 X-linked protein for all the mutant cells (fig.6-7).

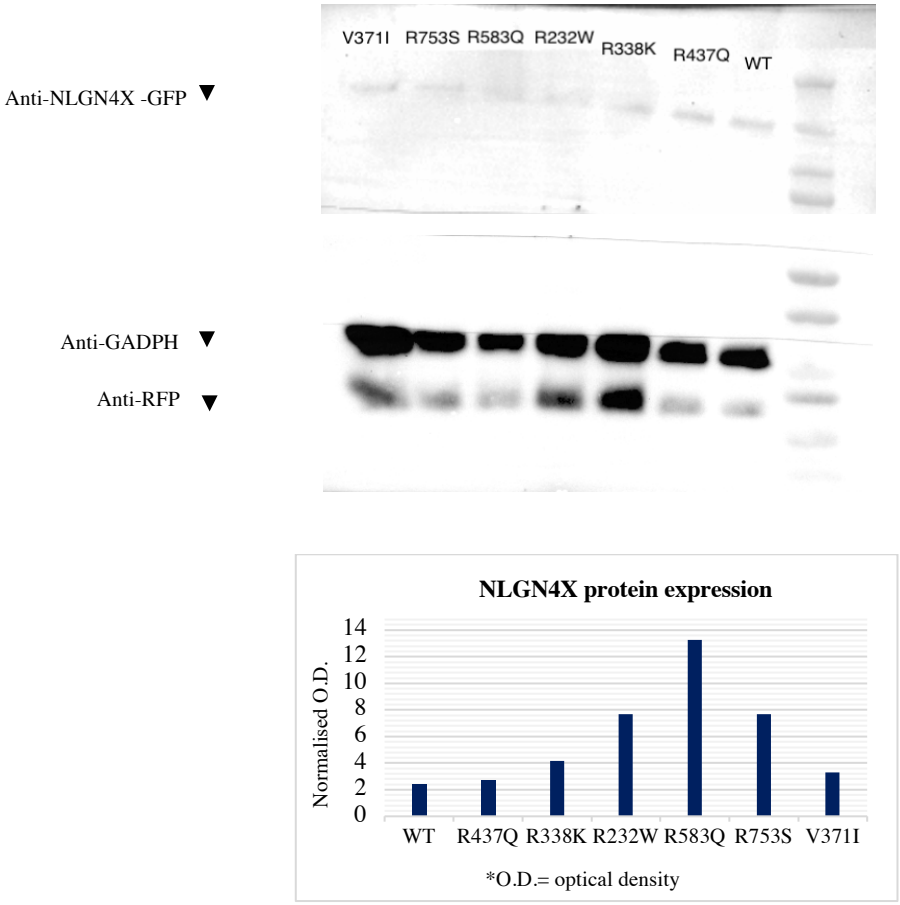


Figure 6-7: Western blot on Neuro2A cell line reporting NLGN4X-GFP, GAPDH and RFP signals going from above to below. Primary antibodies used were Turbo-GFP, anti-RFP and anti-GAPDH. Below, differences in the normalised optical density measured in the mutant cells and in the wt cells, showing an increase in the protein expression in the mutant cells.

3.Materials and Methods

3.1 Whole Exome Sequencing

3.1.1. Selection of Patients and DNA Samples Preparation

Genetic counselling was carried out to evaluate each patient's personal and familial history. Parents provided and signed a written informed consent at the Medical Genetics department of the University of Siena for exome sequencing analysis, clinical data usage, and the use of DNA samples from the tested individuals for both research and diagnosis purposes. We analysed several patients affected by ID and ID/ASD recruited from January 2019 until the end of August 2022.

Genomic DNA from the probands and parents was isolated from EDTA peripheral blood samples using MagCore HF16 (Diatech Lab Line, Jesi, Ancona, Italy) according to the manufacturer's instructions.

3.1.2. Libraries 'preparation and analysis

Sample preparation was performed following the "Illumina DNA prep with enrichment" manufacturer's protocol. This protocol involves the use of a bead-based transposome complex to perform tagmentation, a process that fragments the genomic DNA and then tags it with adapter sequences in one step. After saturation with input DNA, the bead-based transposome complex fragments a set number of DNA molecules. This fragmentation provides flexibility to use a wide DNA input range to generate normalised libraries with a consistent tight fragment size distribution. Then, a limited-cycle PCR adds adapter sequences to the ends of a DNA fragment. A subsequent target enrichment workflow is applied. Following pooling, the double-stranded DNA libraries are denatured and biotinylated. Illumina Exome Panel v1.2 (CEX) probes are hybridised to the denatured library fragments. Then Streptavidin Magnetic Beads (SMB) capture the targeted library fragments within the regions of interest. The indexed libraries are eluted from the beads and further amplified before sequencing. Whole-exome sequencing analysis was performed on the Illumina NovaSeq6000 System (Illumina San Diego, CA, USA) according to the NovaSeq6000 System Guide. Reads were mapped against the hg19 reference genome using the Burrow–Wheeler aligner (BWA) [5].Variant

calling was obtained using an in-house pipeline which takes advantage of the GATK Best Practices workflow [6].

3.1.3. Filtering and Variant Prioritisation

All variants were screened according to frequency, location, mutation category, literature, and mutation database data (ClinVar database, LOVD database, HGMD database). Polymorphisms (minor allele frequency, MAF< 0.01) were excluded, and synonymous variants were assumed to be benign or likely benign. Missense variants were predicted to be damaging by CADD-Phred prediction tools for functional effect prediction. Frameshift, stop-gain, and splice site variants were prioritised as pathogenic. A prediction of damage for the unreported missense variants in the new candidate genes came from HOPE and Phyre2 tools. The potential impact of the variants on splicing was evaluated using Alamut Visual software—version 2.11-0 (Interactive Biosoftware, Rouen, France), which employs five different algorithms: SpliceSiteFinder-like, MaxEntScan, NNSPLICE, GeneSplicer, and HumanSplicingFinder. The following public databases were used for the interpretation of the variants: ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), LOVD (https://databases.lovd.nl/shared/genes), Human Genome Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/index.php3).

3.1.4 Variant's identification

The identification of the ID/ASD causative variants proceeded in many steps:

1. Assessment of pathogenicity: nonsense and frameshift variants were assumed to be pathogenic based on their intrinsic nature, whereas a damage prediction was made for synonymous, splice site, and nonsynonymous variants by bioinformatic tools and by splicing analysis; 2. Definition of candidate ID/ASD genes: the genes with variants predicted to be pathogenic were analysed regarding their brain-expression patterns, the presence of animal models, evolutionary conservation; 3. Determination of novel ID/ASD gene: genes with mutations occurred in patients with symptoms overlapping to those previously described for other persons with PVs/LPVs in the same gene (fig.8).

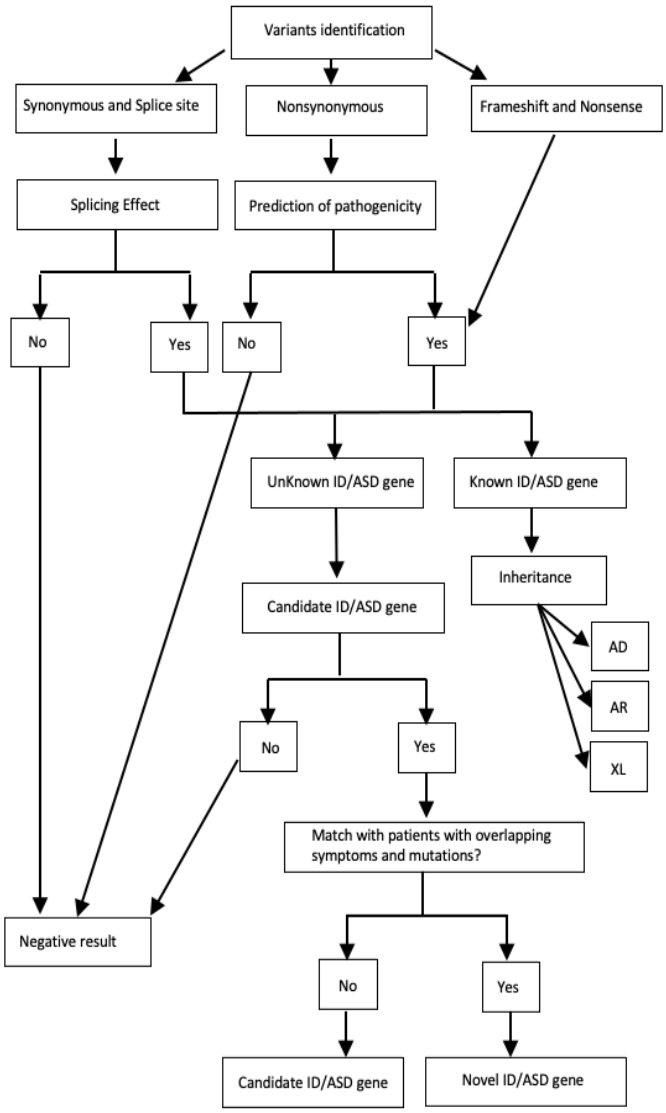


Figure 8: Workflow of the variants analysis for the ID/ASD patients.

3.1.5. Sanger Sequencing

Variants identified by WES were confirmed by Sanger sequencing. PCRs were performed using HS-TaqDNA Polymerase reagents (10x Buffer; 50mM $MgCl_2$, 5U/µl HS- Taq DNA Polymerase) by Experteam (Venezia, Italy), in addition with 25 mM dNTPs, milliQ H_2O, specific primers specially designed and 100 ng of cDNA, for a final volume of 50 µL. Amplification was performed with the following PCR program: 95°C for 5 min; 35 cycles at 95°C for 30sec, at the specific annealing temperature for 30sec, at 72°C for 30sec; 72°C for 7 min. 3 µL of amplified PCR products were mixed with the loading dye, containing Bromophenol Blue, and separated by electrophoresis on a 1.2% agarose gel. Bands were displayed in a UV transilluminator. PCR products were purified using AMPure XP magnetic beads (Beckman Coulter, Inc. Brea, CA, USA). DNA samples were sequenced using the PE

Big Dye Terminator Cycle Sequencing Kit on an ABI Prism 3130 analyzer (Applied Biosystems). The data were analysed using the Sequencher version 5.4.6 software.

3.2. Splicing Analysis

3.2.1 RNA Extraction and Reverse Transcriptase-PCR (RT-PCR)

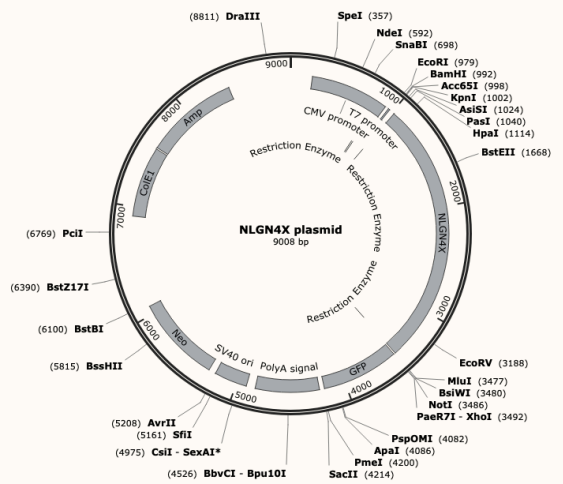
Total RNA was isolated from PAXgene blood RNA tubes (PreAnalytiX®, Qiagen, Hilden, Germany) (http:/www.qiagen.com), with PAXgene Blood RNA Kit (IVD) (PreAnalytiX®), following the manufacturer's instructions, and quantified by NanoDropTM 2000/2000c Spectrophotometer (ThermoFisher Scientific). For each sample, 1 µg of RNA was reverse transcribed into cDNA using a dedicated Qiagen kit (QuantiTect®Reverse Transcription Kit, Qiagen) according to the manufacturer's instructions. cDNA was stored at -20°C until use.

3.2.2 Transcript Characterization

To verify the effect of the considered splicing variant on *CNOT3*, specific PCR primers were designed. Two negative control cDNA samples were included. PCRs were performed using the same reagents used for Sanger Sequencing, plus 100 ng of cDNA, for a final volume of 50 µL. Amplification was performed with the following PCR program: 95°C for 5 min; 40 cycles at 95°C for 1 min, at the specific annealing temperature for 1 min, at 72°C for 1 min, 72°C for 7 min. Then, 10 µL of amplified PCR products were mixed with the loading dye, separated on a 3% agarose gel and displayed in a UV transilluminator. The different amplification products were purified using a dedicated Qiagen kit (MinElute® Gel Extraction Kit, Qiagen); the negative control's unfractionated PCR products were purified using AMPure XP magnetic beads (Beckman Coulter, Inc. Brea, CA, USA). cDNA PCR products were sequenced using PE Big dye terminator cycle sequencing kit on ABI Prism 3130 Genetic Analyzer (Applied Biosystem), as for Sanger Sequencing.

3.4. Site-directed mutagenesis approach

The site-directed mutagenesis was realised using the Agilent QuikChange II XL Site-Directed Mutagenesis Kit. The following *NLGN4X* plasmid and pairs of primers were engineered for eleven variants (fig.9):



Variant	Forward	Reverse	Tm
R232W	GGGCTCCTGGATCAGATTCAAGCACTGTGGTGGATTGA GGAGAATGTGGG	CCCACATTCTCCTCAATCCACCACAGTGCTTGAATCTGA TCCAGGAGCCC	66,00
G174E	TCATGGTCTATATCCATGGGGAATCTTACATGGAGGGC ACCGG	CCGGTGCCCTCCATGTAAGATTCCCCATGGATATAGACC ATGA	85,62
R437Q	GGAAAACCCGGAGACGCGGCAGAAAACCCTGGTGGCT CTCTTT	AAAGAGAGCCACCAGGGTTTTCTGCCGCGTCTCCGGGTT TTCC	85,21
R583Q	GAGTGAGAGATCACTACCAGGCAACGAAAGTGGCTTTC TGG	CCAGAAAGCCACTTTCGTTGCCTGGTAGTGATCTCTCAC TC	85,97
V371I	CTACGACATCATGCTGGGCATCAACCAAGGGGAAGGCC TGAAGTT	AACTTCAGGCCTTCCCCTTGGTTGATGCCCAGCATGATG TCGTAG	84,78
L406I	GTGTCCAACTTCGTGGACAACATTTACGGCTACCCTGAA GGGAA	TTCCCTTCAGGGTAGCCGTAAATGTTGTCCACGAAGTTG GACAC	85,21
V522M	CGACGTCATGCTCAGCGCCATGGTCATGACCTACTGGA CGAACT	AGTTCGTCCAGTAGGTCATGACCATGGCGCTGAGCATG ACGTCG	85,21
R753S	GCAGGCACACGACACACTGAGTCTCACCTGCCCGCCAG ACTACAC	GTGTAGTCTGGCGGGCAGGTGAGACTCAGTGTGTCGTGT GCCTG	87,24

Figure 9: *NLGN4X* plasmid and primers for mutagenesis.

Cell Transformation was obtained adopting the One Shot TOP10 Chemically Competent E.coli. After having performed mini preparations of the plasmids, the mutations were checked via Sanger sequencing with the following primers:

Primers sequences

Forward

5'- CGTCGCTCCTCTTCCTCAACAT-3'
5'-CACCACAAATGATATCGCTCA-3'
5'- CCACCGAGCTCTTCAGTTGT-3'
5'-ACACCACGGACATGGTAGAA-3'
5'-AAGTACACTCGGATATTGGCAG-3'
5'-TTCAAGATCAAAATGAAGACTGC-3'
5'-GAACACCGTTACCCAATGAGA-3'

Reverse

3'- AAAGTGTGCAAAGGCTGCAT-5'
3'-GTTGTGCAAATGAGGAACGA-5'

3'- ACACCGTCCTCGTTATCCAC-5'
3'- GAAGACATAGGGGACCTCATCAC-5'
3'- TCCCACATTCTCCTCAATCC-5'
3'- GTAAACCAGATGGGCAGCAT-5'

3.5 Immunofluorescence staining and Duolink studies on cells

Cells are seeded 24 h before labelling in 12-well plates containing coverslips, 1 10^5 mouse neuroblasts Neuro 2A per well. Cells are washed with cold PBS and fixed with 4% formaldehyde for 20 min at room temperature. The cells are then incubated with NH4Cl (50mM) solution for 10 min, followed by permeabilization with 0.3% PBS-Triton solution for 7 min, still at room temperature. Saturation is performed with 1% PBS-BSA, 0.1% Tween 20, 15 min at room temperature. The primary antibody (or two if co-labeling) is incubated at 4°C overnight, diluted in a blocking buffer. The next day, after 3 washes with cold PBS, the 1/1000th secondary antibody is added for 45 min. After 3 final washes, the mount is made with a DAPI-containing medium and analysed by microscopy. For more advanced colocalization studies, the Duolink® In Situ Detection Reagents Orange kit is used, in combination with PLA Probe Anti-Rabbit PLUS and Anti-Mouse MINUS (Sigma).
Microscopic studies are performed at the GAD laboratory with a Zeiss microscope in epifluorescence, or at the Dimacell imaging platform of INRA in Dijon, in confocal. Image analysis, quantification, and evaluation of colocalization (using JACoP) is done with ImageJ software.

3.6 Western blot procedure

Cells were collected by trypsinization and lysed in 1X RIPA buffer (Cell Signalling) supplemented with PMSF (1 µM) and a protease inhibitor cocktail (Sigma). After 15 min on ice and centrifugation at 13000g at 4 °C for 15 min, the total lysate was recovered. Proteins were dosed using the BCA kit (Sigma Aldrich).
For the Western-blot, 40 µg of protein from cell lysates were run on 10% SDS-PAGE gel and migration was performed in 25 mM Tris buffer, 192 mM Glycine, 0.10% SDS at 80 V for 30 min and then 120 V for 3 h. The proteins were transferred for 1 hour at 120 V in Tris-Boric buffer (Tris 49 mM, Boric acid 48.5 mM) onto a millipore PVDF membrane (Immobilon®-P) previously activated with methanol. The membrane was saturated with a solution of TBS, 0.05% Tween 20 and 5% BSA for 30 min and incubated with primary antibodies overnight

at 4 °C. After 3 washes, HRP-conjugated secondary antibody was added for 1 hour. The labelled proteins were detected using ECL ClarityTM substrate (Bio-Rad Laboratories, Inc) and bands on western blots were visualised using ChemiDoc™ Imaging System (Bio-Rad). Molecular weights were determined using Image Lab software (Bio-Rad).

4. Discussion

Investigating the genetic causes of ID/ASD conditions is a long and complex process, both for the lack of a comprehensive first-tier genetic test for GDD and for the biological phenomena related to non-linear genotype-phenotype correlations in ID/ASD patients. Thus, the diagnostic outcome for many cases remains often controversial, requiring further tests.

Our results support the application of the WES both for diagnostic and for research purposes. WES has revealed its utility for the resolution of cases showing clinical symptoms common to more than one medical condition, such as KBG syndrome (KBGS), Cornelia de Lange, or Noonan syndrome, permitting to obtain an early diagnosis, particularly in paediatric patients with immature clinical features. Our research reinforces the concept that KBGS general clinical features are developmental delay, a recognisable facial gestalt, skeletal/dental anomalies, epilepsy, short stature, cerebral anomalies. Our findings support a phenotypic evolution from Coffin-Siris/ Nicolaides Baraitser and CdLS to KBGS from the childhood to the adulthood for KBGS patients and suggest a broadening of the diagnostic criteria for KBGS as atypical clinical findings as a wide anterior fontanelle can be the only clinical sign[7]. In this context a combined use of WES and array-CGH has optimised the diagnostic yield.

Many atypical presentations emerged for the *DDX3X, KCNQ3*, *KIF1A, SHANK3, POGZ,* genes. We documented the case of a patient mutated in *DDX3X* manifesting Rett-like spectrum features instead of the most common signs: hypotonia, movement disorder, behavioural problems, corpus callosum hypoplasia and epilepsy. Also, we detected a patient displaying a mutation in *KCNQ3* not associated with epilepsy but ID, ASD, stereotypies, bladder anomalies, craniofacial dysmorphisms. *KIF1A* is implicated in the NESCAV syndrome, for which GDD, spasticity, ID, speech delay, and behavioural abnormalities are described as typical symptoms. We reported a patient suffering from spastic paraparesis, behaviour disorder, enlargement of the interfolial spaces of the cerebellar hemispheres, hypertonia. We found the mutation (p.(Pro305Leu)) in two patients with clinical signs comprising in one case language delay, cerebellar and vermis atrophy, psychomotor delay, brain abnormalities, hypertrichosis, bilateral clinodactyly, and facial dysmorphisms, whereas in the other ataxia, spastic paraparesis, angioma, nystagmus, and seizures were observed. Finally, we found mutations in *SHANK3* in patients affected with “non-syndromic” ID, not presenting with the classical signs of the Phelan-McDermid syndrome [8]. For these patients,

WES did not detect any additional molecular cause possibly concurring to their phenotypes, reinforcing the associations with the reported PVs.

WES is therefore effective to make diagnosis for diseases characterised by variable clinical expressivity. We found the PV (p.(Met394Valfs*9)) in the *POGZ* gene in three family members. In the female proband craniofacial dysmorphisms, hyperactivity, blepharophimosis, kidney abnormalities and language delay, emerged. She shared ID, brachydactyly, and nail hypoplasia with her mother, the latter also manifested microcephaly. Her brother exhibited ID, hypotonia, obesity and had some craniofacial dysmorphisms.

Additionally, the discovery of variants in disease gene, not suspected in first instance to be causative of the probands presentations, has expanded the gene-associated clinical spectrum for *CNOT3* and *SMARCC2.*

The female protective effect could in part give explanation of the intrafamilial phenotypic variability observed in familiar members with a novel splice site mutation in *CNOT3* segregating from mother to son, with the latter more severely affected. Here, *CNOT3* is firstly related to anorectal dysplasia. Rectal atresia is an anorectal malformation (ARM) due to defects in the development of the rectum in early foetal life. As *CNOT3* takes part in the mesendoderm differentiation and the rectum is generated from the differentiation of the endoderm in the primordial gastrointestinal tract, this suggests a connection between *CNOT3* alterations and rectal atresia. We are not able to demonstrate if the exons 1-2 duplication induced by the variant c.26-2A>G, together with a retention of intron 2, represents an aberrant isoform, or a sort of Tandem Exonic Duplication (TED). TED is an evolutive strategy for substitution alternative splicing leading to several isoforms: the major ones retaining their function, while the minor isoform being selected according to their effect[9].This mechanism could explain the exons duplication as a sort of rescue of protein integrity, with one of the duplicated exons performing the original function, and the other retaining intron two potentially assuming pathogenetic significance.

Also, for what concerns the molecular characterization of the variants in our cohort, we detected an enrichment of disruptive SNVs in the female patients whereas more missense SNVs were detected in male patients. This is consistent with the female protective effect model according to which females would require an increased genetic burden to reach the threshold for ID/ASD diagnosis. Female patients in many cases may be not included for the

presence of mild symptoms due to missense variants capable of induce more severe effects in the corresponding male patients.

A coexistence of two genetic syndromes has been detected in a family with the mother and the old daughter holding respectively two mutations in *RAF1* and *SMARCC2,* the latter associated with novel clinical findings associated with the Coffin Siris 8 syndrome. Instead, the younger sister showed both the genetic defects, proving the WES utility in unmasking the genetic causes of complex phenotypes.

In medical genetics, a misdiagnosis can be attributed not exclusively to the patient's clinical heterogeneity but, as for mucopolysaccharidosis, also to the slow progress of the disease or to the high rate of false negative results in the urinary screening test for mucopolysaccharides. For these diseases, a positive molecular result can contradict false negative biochemical exams, it is needed to make the diagnosis, and to eventually pick out the patients for gene-editing therapies[10].

Finally, the use of WES as first tier analysis avoids non targeted therapies guaranteeing economic advantages, but above all, health benefits.

Also, it promises to shed light on that part of the exome still unknown. We identified seven ID/ASD candidate genes:

- Alterations in *CACNA2D1*, which encodes the alpha-2/delta subunit of skeletal muscle and brain voltage-dependent calcium, have been previously detected in patients with epilepsy and ID. Point mutations in mice in the *CACNA2D1* gene cause an altered synaptic transmission at the level of the central nervous system;
- *GPR14* (orphan G protein-coupled receptor 14 for Urotensin II), is broadly expressed in the brain and spinal cord;
- We report four families with variants in *SPTBN5* (Spectrin Beta, Non-Erythrocytic 5), being the last gene of the spectrin family to be related to NDDs. ID, DD, and aggression symptoms overlap with those described for the other spectrin genes. Low rates of seizures are attributable to variants in *SPTBN5* as compared to *SPTBN1*, *SPTBN4*, and individuals with *SPTBN1* and *SPTBN4* variants show cerebellar or cerebral atrophy that were absent in our families [11];
- The Myelin Basic Protein (*MBP*) is the principal constituent of myelin, needed for the constitution of the mature form of its multilamellar membrane structure. Mutations in the *MBP* gene in mice provoke a decreased myelination of the central nervous system,

tremors, and to a premature death. A gradient of transcription for *MBP* has been observed in the developing human brain, as it reflects the process of myelination. Our *MBP* mutated patient presented hand stereotypes, hyperactivity, sleep disorder, and speech delay;

- *PCDHA1* was found mutated in a patient with ADHD, language and psychomotor delay, simplified auricles, and abnormal cerebral structures. *PCDHA1* is part of the *PCDHA* gene cluster which participates in the axonal projection, learning, and memory;
- *PCDH15* PVs cause Usher syndrome, even though psychiatric symptoms and comorbidities with mental illnesses co-occur. Remarkably, SNVs in *PCDH15* in a heterozygous fashion were detected in a patient with ASD. We found a *PCDH15* frameshift variant in a patient with ID, epilepsy, and language, and psychomotor delay and we hypothesised that it might explain his phenotype;
- *PDPR*, encoding for the pyruvate dehydrogenase phosphatase regulatory subunit, is expressed at different levels in the brain, especially in the corpus callosum and in the cerebellum. A PV in homozygosity was previously found in a patient with GDD, Joubert-like symptoms, and MRI findings. We sustain that a *PDPR* variant might contribute to the language delay and behaviour disorder of our patient.

Reactome browser https://reactome.org/PathwayBrowser/ and Pathways Common https://apps.pathwaycommons.org/search?q=PCDHA1&type=Pathway accessed the 17/10/2022 revealed the involvement of the new candidate genes in different pathways, already linked to neurodevelopmental disorders: *CACNA2D1* is implied in the presynaptic depolarization and calcium-channel opening, *MBP* plays a role in the EGR2 and SOX10-mediated initiation of Schwann cell myelination, *GPR14* (or *UTS2R*) acts in the G-alpha (g) signalling events and in the class A/1 Rhodopsin-like receptors pathway, *PCDH15* contributes in the sensory processing of sound by inner /outer hair cells of the cochlea, *PDPR* contributes in the citric acid (TCA) cycle and respiratory electron transport, *SPTBN5* is present in the COPI-mediated anterograde transport, in the NCAM signalling for neurite out-growth, in the RAF-MAPK kinase cascade, and interacts with L1 CAM, *PCDHA1* gene participates to the WNT and to the Cadherin signal pathways.

Finally, we found VUS in ID/ASD known genes and in many novel genes, both alone and in combination. A large piece of literature proves the association between de novo SNVs and

ASD/ID, particularly for disruptive variants. The discovery of an enrichment of multiple de novo variants in various genes found via large-scale WES supports the involvement of an oligogenic model in patients with ASD. Our findings are consistent with the theory of a multi-hit model in the pathogenesis of ASD and ID[12].

We also contributed to identify the novel ID/ASD gene *CAPRIN1*, whose functional investigation has highlighted that *CAPRIN1* half-dose causes morphologic, growth and functional alterations in neuronal cells in vitro, supporting haploinsufficiency as a pathogenic mechanism.

Experimental evidence is required to characterise novel genes, but also to pinpoint novel mechanisms of action for known ID/ASD genes. Currently, missense variants in the *NLGN4X* gene are known to act through heterogeneous mechanisms, causing either the retention of the altered protein in the endoplasmic reticulum (LoF mechanism), or the increase of the number of the synapses (GoF mechanism), based on the gene role in different pathways (inhibitory and excitatory). We illustrated a site-specific mutagenesis approach on the Neuro2A murine cell line and on GN11 mouse neuronal cell line on 8 variants obtained both from our collaboration and from literature being reported as VUS. We detected morphological alterations as a decrease of the neurite's length and in the membrane distribution of the Neuroligin 4 X-linked protein in the mutant cells with respect to the wt cells, suggesting different pathogenic mechanisms. A regression to an undifferentiated phenotype was apparent for the cells with the mutation R232W, being possibly related to a retention of the protein in the endoplasmic reticulum as suggested by the anti-RTN4 staining.

A decrease in the number of actin filaments in the mutant cells was also observed, possibly affecting the neuritogenesis, existing a positive correlation between the actin filament number and the growth cone area, which ultimately has an impact on neuritogenesis [13]. Finally, western blot results suggest a likely GoF mechanism at the base of the altered biological role of the NLG4 mutant proteins. Interestingly, the new clinical features detected in our cohort may raise the possibility to enlarge the *NLGN4X* clinical spectrum. Nevertheless, an in-depth analysis is required to exclude other PVs potentially responsible for the phenotype of the patients.

5. Limitations of the study

The main limitation of this study relies on the size of the cohort, not allowing a proper statistical analysis of the contribution of the mentioned variants to the ID/ASD symptoms in patients. The second drawback came from the possible lack of molecular diagnosis for probands with a partial or no clinical diagnosis, that could have hampered the genetic investigation, as patients presenting mild symptoms may have been excluded. On the contrary, the analysis could have been misguided by clinical bias, with exclusion of the variants contributing to the phenotypes.

Furthermore, for most of the cases we were not able to detect mosaic events, as we did not evaluate the impact of mitochondrial mutations and of alterations in the methylation status in the probands. The last weakness consists in the interpretation of the WES results, limited to the status of the current scientific knowledge. A periodic re-evaluation of the identified VUS is thus recommended, potentially providing a revision of the analysis and additional diagnosis over time.

Additional evidence is needed to clarify how the morphological changes in the mutant cells are responsible for the patients' pathologies and if the *NLGN4X* variants must be considered causative of the novel clinical findings that we detected in our cohort. A statistical analysis of the decrease in the number of the actin filaments in the mutant must be addressed. At last, the overexpression of NLG4 alone is not sufficient to demonstrate a GoF mechanism and electrophysiological evidence should be provided to determine the effect of the mutant proteins on the synaptic functionality.

6. Concluding remarks

The reported data in the present thesis suggest that both the research and diagnostic fields could benefit from the WES application. The enlargement of the clinical spectrum of the ID/ASD genes and the discovery of new ID/ASD candidate genes appear the main achievements. Nevertheless, this thesis stresses the need of experimental evidence for characterising novel ID/ASD genes, to assign novel mechanisms of action and establish genotype-phenotype correlations.

7.Bibliography

[1] A. Thapar, M. Cooper, and M. Rutter, "Neurodevelopmental disorders," *The Lancet Psychiatry*, vol. 4, no. 4. 2017. doi: 10.1016/S2215-0366(16)30376-5.

[2] D. J. Morris-Rosendahl and M. A. Crocq, "Neurodevelopmental disorders-the history and future of a diagnostic concept," *Dialogues Clin Neurosci*, vol. 22, no. 1, 2020, doi: 10.31887/DCNS.2020.22.1/macrocq.

[3] M. Servetti *et al.*, "Neurodevelopmental Disorders in Patients With Complex Phenotypes and Potential Complex Genetic Basis Involving Non-Coding Genes, and Double CNVs," *Front Genet*, vol. 12, 2021, doi: 10.3389/fgene.2021.732002.

[4] A. K. Srivastava and C. E. Schwartz, "Intellectual disability and autism spectrum disorders: Causal genes and molecular mechanisms," *Neuroscience and Biobehavioral Reviews*, vol. 46, no. P2. Elsevier Ltd, pp. 161–174, Oct. 01, 2014. doi: 10.1016/j.neubiorev.2014.02.015.

[5] S. S. Jeste and D. H. Geschwind, "Disentangling the heterogeneity of autism spectrum disorder through genetic findings," *Nature Reviews Neurology*, vol. 10, no. 2. 2014. doi: 10.1038/nrneurol.2013.278.

[6] L. Kaufman, M. Ayub, and J. B. Vincent, "The genetic basis of non-syndromic intellectual disability: A review," *Journal of Neurodevelopmental Disorders*, vol. 2, no. 4. 2010. doi: 10.1007/s11689-010-9055-2.

[7] S. Bölte, S. Girdler, and P. B. Marschik, "The contribution of environmental exposure to the etiology of autism spectrum disorder," *Cellular and Molecular Life Sciences*, vol. 76, no. 7. 2019. doi: 10.1007/s00018-018-2988-4.

[8] A. Reichenberg *et al.*, "Discontinuity in the genetic and environmental causes of the intellectual disability spectrum," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 4. 2016. doi: 10.1073/pnas.1508093112.

[9] A. M. Alazami *et al.*, "Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families," *Cell Rep*, vol. 10, no. 2, pp. 148–161, Jan. 2015, doi: 10.1016/j.celrep.2014.12.015.

[10] N. N. Parikshak *et al.*, "XIntegrative functional genomic analyses implicate specific molecular pathways and circuits in autism," *Cell*, vol. 155, no. 5, 2013, doi: 10.1016/j.cell.2013.10.031.

[11] S. Srivastava *et al.*, "Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders and the NDD Exome Scoping Review Work Group", doi: 10.1038/s41436.

[12] N. Vadgama *et al.*, "De novo single-nucleotide and copy number variation in discordant monozygotic twins reveals disease-related genes," *European Journal of Human Genetics*, vol. 27, no. 7, 2019, doi: 10.1038/s41431-019-0376.

[13] V. Kalscheuer, S. Schweiger, N. Laura, L. Merckx, and H. van Esch, "Human Brain Models of Intellectual Disability: Experimental Advances and Novelties," 2022, doi: 10.3390/ijms23126476.

[14] T. W. Yu, M. Coulter, M. Chahrour, and C. A. Walsh, "Autism Spectrum Disorders," *Genomic and Personalized Medicine*, vol. 2, pp. 1067–1074, 2013, doi: 10.1016/B978-0-12-382227-7.00090-2.

[15] "Correction and Republication: Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012," *MMWR Morb Mortal Wkly Rep*, vol. 67, no. 45, 2018, doi: 10.15585/mmwr.mm6745a7.

[16] C. Farmer, A. Thurm, and P. Grant, "Pharmacotherapy for the core symptoms in autistic disorder: Current status of the research," *Drugs*, vol. 73, no. 4, 2013, doi: 10.1007/s40265-013-0021-7.

[17] K. Malpass, "Unlocking the secrets of autism through whole-exome sequencing," *Nat Rev Neurol*, vol. 8, no. 6, 2012, doi: 10.1038/nrneurol.2012.82.

[18] B. J. O ' Roak *et al.*, "Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations," *Nature*, vol. 485, no. 7397, 2012, doi: 10.1038/nature10989.

[19] S. J. Sanders *et al.*, "De novo mutations revealed by whole-exome sequencing are strongly associated with autism," *Nature*, vol. 485, no. 7397, 2012, doi: 10.1038/nature10945.

[20] T. Sierra-Arregui, J. Llorente, P. Giménez Minguez, J. Tønnesen, and O. Peñagarikano, "Neurobiological Mechanisms of Autism Spectrum Disorder and Epilepsy, Insights from Animal Models," *Neuroscience*, vol. 445. 2020. doi: 10.1016/j.neuroscience.2020.02.043.

[21] A. Thurm, C. Farmer, E. Salzman, C. Lord, and S. Bishop, "State of the field: Differentiating intellectual disability from autism spectrum disorder," *Frontiers in Psychiatry*, vol. 10. 2019. doi: 10.3389/fpsyt.2019.00526.

[22] I. Ivanović, "Psychiatric Comorbidities in Children With ASD: Autism Centre Experience," *Front Psychiatry*, vol. 12, 2021, doi: 10.3389/fpsyt.2021.673169.

[23] F. M. C. Besag, "Epilepsy in patients with autism: Links, risks and treatment challenges," *Neuropsychiatric Disease and Treatment*, vol. 14. 2018. doi: 10.2147/NDT.S120509.
[24] A. Polyak, R. M. Kubina, and S. Girirajan, "Comorbidity of intellectual disability confounds ascertainment of autism: Implications for genetic diagnosis," *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, vol. 168, no. 7, 2015, doi: 10.1002/ajmg.b.32338.
[25] I. Parenti, L. G. Rabaneda, H. Schoen, and G. Novarino, "Neurodevelopmental Disorders: From Genetics to Functional Pathways," *Trends in Neurosciences*, vol. 43, no. 8. 2020. doi: 10.1016/j.tins.2020.05.004.
[26] J. M. Savatt and S. M. Myers, "Genetic Testing in Neurodevelopmental Disorders," *Frontiers in Pediatrics*, vol. 9. 2021. doi: 10.3389/fped.2021.526779.
[27] A. Arteche-López *et al.*, "Towards a change in the diagnostic algorithm of autism spectrum disorders: Evidence supporting whole exome sequencing as a first-tier test," *Genes (Basel)*, vol. 12, no. 4, 2021, doi: 10.3390/genes12040560.
[28] M. I. Álvarez-Mora *et al.*, "Diagnostic yield of next-generation sequencing in 87 families with neurodevelopmental disorders," *Orphanet J Rare Dis*, vol. 17, no. 1, 2022, doi: 10.1186/s13023-022-02213-z.
[29] Y. Guo *et al.*, "Exome sequencing generates high quality data in non-target regions," *BMC Genomics*, vol. 13, no. 1, 2012, doi: 10.1186/1471-2164-13-194.
[30] K. Y. Sánchez-Luquez, M. X. Carpena, S. M. Karam, and L. Tovo-Rodrigues, "The contribution of whole-exome sequencing to intellectual disability diagnosis and knowledge of underlying molecular mechanisms: A systematic review and meta-analysis," *Mutation Research/Reviews in Mutation Research*, vol. 790, p. 108428, Jul. 2022, doi: 10.1016/J.MRREV.2022.108428.
[31] B. Rabbani, M. Tekin, and N. Mahdieh, "The promise of whole-exome sequencing in medical genetics," *Journal of Human Genetics*, vol. 59, no. 1. 2014. doi: 10.1038/jhg.2013.114.
[32] T. J. Dixon-Salazar *et al.*, "Exome Sequencing Can Improve Diagnosis and Alter Patient Management," *Sci Transl Med*, vol. 4, no. 138, 2012, doi: 10.1126/scitranslmed.3003544.
[33] A. C. Lionel *et al.*, "Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test," *Genetics in Medicine*, vol. 20, no. 4, 2018, doi: 10.1038/gim.2017.119.
[34] E. Lalonde, S. Rentas, F. Lin, M. C. Dulik, C. M. Skraban, and N. B. Spinner, "Genomic Diagnosis for Pediatric Disorders: Revolution and Evolution," *Frontiers in Pediatrics*, vol. 8. 2020. doi: 10.3389/fped.2020.00373.
[35] N. Jalkh *et al.*, "The added value of WES reanalysis in the field of genetic diagnosis: Lessons learned from 200 exomes in the Lebanese population," *BMC Med Genomics*, vol. 12, no. 1, 2019, doi: 10.1186/s12920-019-0474-y.
[36] J. R. Stojanovic *et al.*, "Diagnostic and Clinical Utility of Clinical Exome Sequencing in Children With Moderate and Severe Global Developmental Delay / Intellectual Disability," *J Child Neurol*, vol. 35, no. 2, 2020, doi: 10.1177/0883073819879835.
[37] M. M. Clark *et al.*, "Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases," *NPJ Genom Med*, vol. 3, no. 1, 2018, doi: 10.1038/s41525-018-0053-8.
[38] M. Shakiba and M. Keramatipour, "Effect of whole exome sequencing in diagnosis of inborn errors of metabolism and neurogenetic disorders," *Iranian Journal of Child Neurology*, vol. 12, no. 1. 2018. doi: 10.22037/ijcn.v12i1.19069.
[39] B. Zablotsky *et al.*, "Prevalence and trends of developmental disabilities among children in the United States: 2009–2017," *Pediatrics*, vol. 144, no. 4, 2019, doi: 10.1542/peds.2019-0811.
[40] L. E. L. M. Vissers *et al.*, "A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology," *Genetics in Medicine*, vol. 19, no. 9, 2017, doi: 10.1038/gim.2017.1.
[41] J. Xiang *et al.*, "Genetic Analysis of Children With Unexplained Developmental Delay and/or Intellectual Disability by Whole-Exome Sequencing," *Front Genet*, vol. 12, 2021, doi: 10.3389/fgene.2021.738561.
[42] Y. Zhai, Z. Zhang, P. Shi, D. M. Martin, and X. Kong, "Incorporation of exome-based CNV analysis makes trio-WES a more powerful tool for clinical diagnosis in neurodevelopmental disorders: A retrospective study," *Hum Mutat*, vol. 42, no. 8, 2021, doi: 10.1002/humu.24222.
[43] M. Yang *et al.*, "Genetic diagnoses in pediatric patients with epilepsy and comorbid intellectual disability," *Epilepsy Res*, vol. 170, 2021, doi: 10.1016/j.eplepsyres.2021.106552.
[44] L. Loberti *et al.*, "Natural history of KBG syndrome in a large European cohort," *Hum Mol Genet*, Jul. 2022, doi: 10.1093/hmg/ddac167.
[45] Y. Zhang *et al.*, "Genetic evidence of gender difference in autism spectrum disorder supports the female-protective effect," *Transl Psychiatry*, vol. 10, no. 1, 2020, doi: 10.1038/s41398-020-0699-8.

[46] B. Dan, "Sex differences in neurodevelopmental disorders," *Developmental Medicine and Child Neurology*, vol. 63, no. 5. 2021. doi: 10.1111/dmcn.14853.
[47] S. Jacquemont *et al.*, "A higher mutational burden in females supports a 'female protective model' in neurodevelopmental disorders," *Am J Hum Genet*, vol. 94, no. 3, 2014, doi: 10.1016/j.ajhg.2014.02.001.
[48] S. Baron-Cohen, "The extreme male brain theory of autism," *Trends in Cognitive Sciences*, vol. 6, no. 6. 2002. doi: 10.1016/S1364-6613(02)01904-6.
[49] N. Coscini, K. Williams, D. Chew, K. C. Pang, M. A. O'Connell, and T. May, "Association between early androgens and autistic traits: A systematic review and meta-analysis," *Research in Autism Spectrum Disorders*, vol. 85. 2021. doi: 10.1016/j.rasd.2021.101789.
[50] B. M. Nugent, C. M. O'Donnell, C. N. Epperson, and T. L. Bale, "Placental H3K27me3 establishes female resilience to prenatal insults," *Nat Commun*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-018-04992-1.
[51] J. Bachman, "Site-directed mutagenesis," in *Methods in Enzymology*, vol. 529, 2013. doi: 10.1016/B978-0-12-418687-3.00019-7.
[52] T. C. Südhof, "Neuroligins and neurexins link synaptic function to cognitive disease," *Nature*, vol. 455, no. 7215. 2008. doi: 10.1038/nature07456.
[53] T. Nagase, K. I. Ishikawa, R. Kikuno, M. Hirosawa, N. Nomura, and O. Ohara, "Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro," *DNA Research*, vol. 6, no. 5, 1999, doi: 10.1093/dnares/6.5.337.
[54] A. Rissone *et al.*, "Characterization of the neuroligin gene family expression and evolution in zebrafish.," *Dev Dyn*, vol. 239, no. 2, 2010, doi: 10.1002/dvdy.22196.
[55] S. G. Marro *et al.*, "Neuroligin-4 Regulates Excitatory Synaptic Transmission in Human Neurons," *Neuron*, vol. 103, no. 4, 2019, doi: 10.1016/j.neuron.2019.05.043.
[56] T. Yumoto *et al.*, "Autism-associated variants of neuroligin 4X impair synaptogenic activity by various molecular mechanisms," *Mol Autism*, vol. 11, no. 1, 2020, doi: 10.1186/s13229-020-00373-y.
[57] N. Kopp, I. Amarillo, J. Martinez-Agosto, and F. Quintero-Rivera, "Pathogenic paternally inherited NLGN4X deletion in a female with autism spectrum disorder: Clinical, cytogenetic, and molecular characterization," *Am J Med Genet A*, vol. 185, no. 3, 2021, doi: 10.1002/ajmg.a.62025.
[58] S. K. Singh and C. Eroglu, "Neuroligins provide molecular links between syndromic and nonsyndromic autism," *Science Signaling*, vol. 6, no. 283. 2013. doi: 10.1126/scisignal.2004102.
[59] M. Sarkar *et al.*, "CNOT3 interacts with the Aurora B and MAPK/ERK kinases to promote survival of differentiating mesendodermal progenitor cells," *Mol Biol Cell*, vol. 32, no. 22, 2021, doi: 10.1091/mb21-02-0089.
[60] R. Martin *et al.*, "De novo variants in CNOT3 cause a variable neurodevelopmental disorder," *European Journal of Human Genetics*, vol. 27, no. 11, 2019, doi: 10.1038/s41431-019-0413-6.
[61] X. Zheng *et al.*, "CNOT3-Dependent mRNA Deadenylation Safeguards the Pluripotent State," *Stem Cell Reports*, vol. 7, no. 5, 2016, doi: 10.1016/j.stemcr.2016.09.007.
[62] L. Martinez Gomez, F. Pozo, T. A. Walsh, F. Abascal, and M. L. Tress, "The clinical importance of tandem exon duplication-derived substitutions," *Nucleic Acids Res*, vol. 49, no. 14, 2021, doi: 10.1093/nar/gkab623.
[63] J. H. M. Schuurs-Hoeijmakers *et al.*, "Clinical delineation of the PACS1-related syndrome-Report on 19 patients," *Am J Med Genet A*, vol. 170, no. 3, 2016, doi: 10.1002/ajmg.a.37476.
[64] I. Parenti, L. G. Rabaneda, H. Schoen, and G. Novarino, "Neurodevelopmental Disorders: From Genetics to Functional Pathways," *Trends in Neurosciences*, vol. 43, no. 8. 2020. doi: 10.1016/j.tins.2020.05.004.
[65] K. Machol *et al.*, "Expanding the Spectrum of BAF-Related Disorders: De Novo Variants in SMARCC2 Cause a Syndrome with Intellectual Disability and Developmental Delay," *Am J Hum Genet*, vol. 104, no. 1, 2019, doi: 10.1016/j.ajhg.2018.11.007.
[66] T. Y. Tan *et al.*, "A head-to-head evaluation of the diagnostic efficacy and costs of trio versus singleton exome sequencing analysis," *European Journal of Human Genetics*, 2019, doi: 10.1038/s41431-019-0471.
[67] L. Pavinato *et al.*, "15 Alessandra Renieri 13,14,15 Danai Veltra," vol. 13, p. 25, doi: 10.1093/brain/awac278/6650380.
[68] Y. Yamamoto, A. Yoshida, N. Miyazaki, K. Iwasaki, and T. Sakisaka, "Arl6IP1 has the ability to shape the mammalian ER membrane in a reticulon-like fashion," *Biochemical Journal*, vol. 458, no. 1, 2014, doi: 10.1042/BJ20131186.
[69] H. Li and R. Durbin, "Fast and accurate long-read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 26, no. 5, 2010, doi: 10.1093/bioinformatics/btp698.
[70] R. Poplin *et al.*, "Scaling accurate genetic variant discovery to tens of thousands of samples," *bioRxiv*. 2017. doi: 10.1101/201178.
[71] L. Loberti *et al.*, "Natural history of KBG syndrome in a large European cohort," *Hum Mol Genet*, Jul. 2022, doi: 10.1093/hmg/ddac167.

[72] F. Valentino *et al.*, “Exome sequencing in 200 intellectual disability/autistic patients: New candidates and atypical presentations,” *Brain Sci*, vol. 11, no. 7, 2021, doi: 10.3390/brainsci11070936.
[73] “Identification of a Novel Pathogenic Variant in the NAGLU Gene in a Child with Neurodevelopmental Delay _ Enhanced Reader”.
[74] M. Engelhardt *et al.*, “SPTBN5, Encoding the βV-Spectrin Protein, Leads to a Syndrome of Intellectual Disability, Developmental Delay, and Seizures,” *Frontiers in Molecular Neuroscience | www.frontiersin.org*, vol. 1, p. 877258, 2022, doi: 10.3389/fnmol.2022.877258.
[75] L. P. Bruno *et al.*, “New Candidates for Autism/Intellectual Disability Identified by Whole-Exome Sequencing,” *Int J Mol Sci*, vol. 22, no. 24, p. 13439, Dec. 2021, doi: 10.3390/ijms222413439.
[76] N. J. F. Gatford *et al.*, “Neuroligin-3 and neuroligin-4X form nanoscopic clusters and regulate growth cone organization and size,” *Hum Mol Genet*, vol. 31, no. 5, 2022, doi: 10.1093/hmg/ddab277.