

RESEARCH

Open Access



# Diagnostic yield of next-generation sequencing in 87 families with neurodevelopmental disorders

María Isabel Álvarez-Mora<sup>1,2</sup>, Aurora Sánchez<sup>1,2</sup>, Laia Rodríguez-Revenga<sup>1,2</sup>, Jordi Corominas<sup>3</sup>, Raquel Rabionet<sup>4</sup>, Susana Puig<sup>2,5</sup> and Irene Madrigal<sup>1,2\*</sup> 

## Abstract

**Background:** Neurodevelopmental disorders (NDDs) are a group of heterogeneous conditions, which include mainly intellectual disability, developmental delay (DD) and autism spectrum disorder (ASD), among others. These diseases are highly heterogeneous and both genetic and environmental factors play an important role in many of them. The introduction of next generation sequencing (NGS) has led to the detection of genetic variants in several genetic diseases. The main aim of this report is to discuss the impact and advantages of the implementation of NGS in the diagnosis of NDDs. Herein, we report diagnostic yields of applying whole exome sequencing in 87 families affected by NDDs and additional data of whole genome sequencing (WGS) from 12 of these families.

**Results:** The use of NGS technologies allowed identifying the causative gene alteration in approximately 36% (31/87) of the families. Among them, de novo mutation represented the most common cause of genetic alteration found in 48% (15/31) of the patients with diagnostic mutations. The majority of variants were located in known neurodevelopmental disorders genes. Nevertheless, some of the diagnoses were made after the use of GeneMatcher tools which allow the identification of additional patients carrying mutations in *THOC2*, *SETD1B* and *CHD9* genes. Finally the use of WGS only allowed the identification of disease causing variants in 8% (1/12) of the patients in which previous WES failed to identify a genetic aetiology.

**Conclusion:** NGS is more powerful in identifying causative pathogenic variant than conventional algorithms based on chromosomal microarray as first-tier test. Our results reinforce the implementation of NGS as a first-test in genetic diagnosis of NDDs.

**Keywords:** Neurodevelopmental disorders, Whole exome sequencing, Whole genome sequencing, Genetic diagnostic, Genetic diagnostic yield

## Background

Neurodevelopmental disorders (NDDs) are a group of heterogeneous conditions that result from an abnormal brain development that may give rise to impaired

cognition, communication, adaptive behaviour, and psychomotor skills [1]. NDDs include several disorders such as autism spectrum disorder (ASD), social communication disorders, intellectual disability (ID), developmental delay (DD) or attention deficit hyperactivity disorder (ADHD), among others. NDDs constitute a serious health problem in our society, affecting >3% of children worldwide [2]. Reaching a diagnosis is a challenge, given the considerable clinical and genetic heterogeneity associated with these rare

\*Correspondence: imadbajo@clinic.cat

<sup>1</sup> Biochemistry and Molecular Genetics Department, Hospital Clinic of Barcelona, IDIBAPS (Institut de Investigacions Biomèdiques August Pi I Sunyer), Barcelona, Spain

Full list of author information is available at the end of the article



diseases [3]. Large-scale studies have identified substantial phenotypic heterogeneity and overlap between conditions, making specific genetic NDDs difficult to identify clinically. Currently, there are no biomarkers to diagnose NDDs or to differentiate between them. Rather, these disorders are categorized based on clinical presentation which is problematic since many symptoms are not unique to single NDDs, and several NDDs have clusters of symptoms in common. Similar phenotypes may be associated with different genetic mechanism. Moreover, different variants in the same gene may lead to different disease phenotypes that increase clinical heterogeneity [4, 5]. Exhaustive consultations and traditional genetic investigations are costly and often fail to arrive at a final diagnosis when no recognizable syndrome is suspected.

Next-generation sequencing (NGS) technologies have improved diagnostics of rare genetic diseases. Specifically, whole exome sequencing (WES) has led to enormous progress in deciphering monogenic forms of NDDs [2, 6]. If considering the genetic aetiology of NDDs, each of these disorders could be considered as a rare disease with overlapping clinical phenotypes. Therefore, the simultaneous analysis of several potentially mutated genes improves the diagnostic process, making it faster and more efficient. Nevertheless and despite the revolutionary advances of these approaches, in 40–60% of the families with NDDs the genetic defect remains unknown.

Current recommendations of genetic studies for diagnosis NDDs are based in the guidelines reported by Miller and collaborators reported more than 10 years ago in which chromosomal microarray (CMA) is considered the first-tier clinical diagnosis test for individuals with DD or congenital anomalies [7]. The diagnostic yield of CMA is established in approximately 15–20%. There is accumulating evidence that suggest that the diagnostic yield of whole exome sequencing in NDDs is markedly greater (30–50%) and therefore WES should be placing the first-tier clinical diagnostic test rather than CMA [8]. In the past months, it has been published the new recommendations of the American College of Medical Genetics and Genomics (ACMG) that strongly recommends applying NGS as a first- (or second-) line test in patients with congenital anomalies, DD or ID [9]. Herein, we report diagnostic yields of applying WES in 87 families with NDDs and additional data of whole genome sequencing (WGS) from 12 patients in which previous WES failed to identify a genetic aetiology. Our results support the implementation of WES as a first-tier test for NDDs.

## Material and methods

### Patients

A total of 87 families with one or several members affected by affected by NDDs were referred to the Biochemistry and Molecular Genetics department for genetic testing. This cohort reflects a heterogeneous collection of clinical presentations, making it representative of a typical genetic clinic. Before enrolment, all patients underwent an extensive diagnostic workup, including clinical evaluation and genomic profiling (fragile X syndrome, copy number variations (CNVs) in subtelomeric regions by MLPA and CMA (4 × 44 K, Agilent Technologies). The institutional review board approved the collection and use of these samples for research purposes (Ethics Committee of Hospital Clinic of Barcelona 2011/6625). Written informed consent was obtained from all subjects prior to their participation.

### Cohort structure

We used four different family series in order to test the impact of the testing strategy of an extended family analysis (more than one generation affected) *versus* a trio-based analysis or solo-based analysis (sporadic case). As shown in Table 1, 87 families were selected for WES: (i) 8 familial cases with several affected generations and many affected members. WES was performed in quartets consisting in two patients and two healthy relatives, (ii) 15 sibling couple. WES was performed in the two affected siblings with or without sequencing their parents, (iii) 26 trios sequencing the affected patient and both parents, and (iv) 38 sporadic cases in which WES was only performed to the index case. Moreover, WGS was performed in 12 patients out of the 87 families in which no pathogenic variant was identified by WES.

### Sequencing and bioinformatics pipeline

Analysis of cases that underwent NGS has been previously described [10–15]. Briefly, allelic variants with MAF > 0.03 in any of the databases used (GnomAD,

**Table 1** Summary of detected variants in each family cohort

NGS strategy	Number of studies	Solved cases	% diagnosis
<b>WES</b>	83	29	35%
Large families	8	6	75%
Sibling couple	15	3	20%
Trio	26	8	31%
Sporadic cases	34	12	35%
<b>WGS</b>	12*	1	8%
<b>Total</b>	83	30	36%

\* Cases with no pathogenic variants were identified by WES

ExAC and 1000Genomes) and minimal coverage lower than 20% were discarded. Putative candidate variants were prioritized according to the predicted impact on coding sequence, their presence in ClinVar [16] or the Human Gene Mutation Database (HGMD) [17], zygosity and genetic mode of inheritance, clinical features and the function of the gene. The ACMG standards for interpretation of sequence variants were used to classify all reported variants [18]. Segregation analysis was performed in candidate variants by Sanger sequencing and CMA when necessary.

## Results

Our patient cohort consisted in 87 families with one or more individuals diagnosed with NDDs in which Fragile X syndrome and pathogenic CNVs were previously discarded. According to the high clinical heterogeneity of NDDs, approximately 75% (67/87) of our patients referred heterogeneous co-occurring conditions, where DD/ID was present in 90% of the patients (79/87). Other clinical co-occurring conditions included ADHD, speech disorders, epilepsy, micro/macrocephaly, and the presence of dysmorphic features/ malformations. Clinical features and genetic results of the diagnosed patients are shown in Table 2.

The use of NGS technologies allowed identifying the causative genetic alteration in 36% of the cases, which is a significant increase over conventional CMA testing. The initial cohort included 97 cases, with a diagnostic value of 8.2% and 32% for CMA and NGS, respectively. Using NGS we achieved genetic diagnosis in 31 families identifying all patterns of inheritance (Table 3). Variants of unknown significance (VUS) have not been included in this report.

In Fig. 1 are shown clinical features of some of the patients with definitive diagnosis. WES identified causative alterations in 30 families out of 87. Diagnostic yields ranged from 20 to 75% depending on the cohort structure (Table 1). Statistically analysis identified significant differences between the group of large families, with several generations affected, and the rest of the groups ( $p < 0.001$ ). While the diagnostic yield in the large families was 75%, the other achieved similar rates of pathogenic variant detection (sibling couple 20%; 31% trio-based; and 34% in sporadic cases). Among the 57 families that remained undiagnosed, WGS was applied to 12 patients and causative variants were only found in one (8%) of the families corresponding to a couple of affected siblings (Tables 1, 2).

### Autosomal dominant inheritance

The most common mode of inheritance identified in our cohort was AD which was found in approximately half of

the cases (Table 3). Among them, 14 out of the 15 cases were found to be caused by de novo mutations and the remaining case (patient 1) was found to carry a loss-of-function (LOF) variant in *DYNC1H1* gene inherited from his affected mother (Tables 2, 3).

Most of the de novo pathogenic variants were found to be novel and detected in different genes previously associated with NDDs (*ADNP*, *ANKRD11*, *CNOT3*, *DHDDS*, *NSD1*, *PHIP*, *PPP2R5D*, *SETD1B*, *TUBB5* and *WAC*). We also found three patients with previously reported variants, two of them with de novo variants in *NIPBL* gene and one with a de novo variant in the *KCNT2* gene (Table 2). Patient 20 was found to carry the missense variant c.385T>C (p.Ser129Pro) in *NIPBL* present in gnomAD database with an extremely low allelic frequency (rs141976717); patient 21 carried the splicing variant c.5329-15A>G in *NIPBL* reported in the literature to result the skipping of exon 28 without altering the reading frame [19]; and patient 29 carried the missense variant c.569G>A (p.Arg190His) in the *KCNT2* gene previously reported in the literature [20, 22].

On the other hand, we identified a de novo missense variant in the *CHD9* gene (patient 10), which is currently not associated with ID but some patients with autistic features have been already reported in HGMD database.

### X-linked inheritance

Our results have identified 10 families with pathogenic variants in 9 genes located in the X-chromosome (*ACSL4*, *CLCN4*, *DDX3X*, *GRIA3*, *IQSEC2*, *MID1*, *OCRL*, *SMC1A* and *THOC2*), representing 32% of the solved cases. We identified two families (patient 8 and patient 16) with pathogenic variants in the *THOC2* recently associated with the Mental retardation, X-linked 12/35 syndrome (OMIM # 300,957).

All pathogenic XL variants, except by one, were detected in affected male patients and were associated with XL-recessive pattern of inheritance (Tables 2, 3). Segregation analysis demonstrated that pathogenic variants were inherited from asymptomatic mothers. However, segregation analysis was not available in patient 26 who was found to carry a hemizygous frameshift variant in *MID1* gene (Table 2).

On the other hand, one female (patient 25) was found to carry a de novo missense pathogenic variant in *DDX3X* gene, associated with and ID disorder with XL-dominant pattern of inheritance.

### Autosomal recessive inheritance

In our cohort, 16% (5/31) of the families were found to carry genetic alterations associated with an AR pattern of inheritance (Table 3). WES identified three families with compound heterozygous variants in AR associated

**Table 2** Clinical findings and summary of pathogenic variants identified by WES or WGS

Case	Sex	Clinical findings	Gene	Transcript	Variant	Mode of inheritance and origin	OMIM	References
<b>Group (i)</b>		<b>Familial cases several generations</b>						
Patient_1	M	ID	<i>DYNC1H1</i>	NM_001376.5	c.4462dupC; p.Arg1488Profs*5	AD (mat)	ID, autosomal dominant 13 (#614,563)	
Patient_2	M	ID, microcephaly, dysmorphism	<i>VPS13B</i>	NM_017890.4	c.5998_5999delCT; p.Leu2000Alafs*2 (mat) c.10475_10476delAA; p.Lys3492Argfs*19 (pat)	AR	Cohen syndrome (#216,550)	[13]
Patient_3	M	ID	<i>IQSEC2</i>	NG_021296.1	c.3116-3_3116-2delCA	X-linked (mat)	ID, X-linked 1/78 (#309,530)	[14]
Patient_4	M	ID	<i>SMC1A</i>	NM_006306.2	c.1405C>T; p.Arg469Cys	X-linked (mat)	Cornelia de Lange syndrome 2 (#300,590)	[13]
Patient_5	M	ID	<i>OCRL</i>	NM_000276.3	c.1567G>A; p.Asp523Asn	X-linked (mat)	Dent disease (#300,555)	[13]
Patient_6	F	ID	<i>UBE3A</i>	NM_130839.4	c.2009delA; p.Asp671Metfs*3	Imprinting (mat)	Angelman syndrome (#105,830)	
<b>Group (ii)</b>		<b>Sibling couple</b>						
Patient_7	M	ID, DD, GR	<i>PYCR1</i>	NM_001282280.1	c.797+2delTGGG	AR	Cutis laxa type IIB (#612,940)	
Patient_8	M	Severe ID, ASD	<i>THOC2</i>	NM_001081550.1	c.3323C>T; p.Ser1108Leu	X-linked (germline mosaicism)	ID, X-linked 12/35 (#300,957)	[11]
Patient_9	M	ID, SD	<i>CLCN4</i>	NM_001830.4	c.758 G>A; p.Arg253Gln	X-linked (mat)	Raynaud-claes syndrome (#300,114)	
<b>Group (iii)</b>		<b>Trios</b>						
Patient_10	M	ASD, no speech development	<i>CHD9</i>	NM_001308319	c.3772A>C; p.Thr1258Pro	AD (de novo)		
Patient_11	F	Severe ID, DD and PDD, myoclonic seizures	<i>DHDDS</i>	NM_024887.3	c.632G>A; p.Arg211Gln	AD (de novo)	DD and seizures with or without movement abnormalities (#617,836)	
Patient_12	M	ID, DD	<i>PHIP</i>	NM_017934.7	c.5317 C>T; p.Arg1773Ter	AD (de novo)	DD, ID obesity, dysmorphic features (#617,991)	
Patient_13	M	Moderate ID, epilepsy, neonatal hypotonia, obesity, SD, BHD	<i>SETD1B</i>	NM_001353345.1	c.3772 C>T; p.Arg1258Ter	AD (de novo)	ID with seizures and language delay (#619,000)	[12]
Patient_14	F	Microcephaly	<i>TUBB5</i>	NM_001293213.2	c.1201 G>A; p.Glu401Lys	AD (de novo)	Cortical dysplasia, complex, with other brain malformations 6 (#615,771)	[15]
Patient_15	M	ID	<i>ACSL4</i>	NM_022977.2	c.1030T>C; p.Ser344Pro	X-linked (mat)	ID, X-linked 63 (#300,387)	
Patient_16	M	ID, severe SD, infantile hypotonia	<i>THOC2</i>	NM_001081550.1	c.3361A>G; p.Arg1121Gly	X-linked (mat)	ID, X-linked 12/35 (#300,957)	[11]
Patient_17	M	Moderate ID, macrocephaly	<i>GRIA3</i>	NM_000828.4	c.1892G>A; p.Arg631His	X-linked (mat)	ID, X-linked 94 (#300,699)	

**Table 2** (continued)

Case	Sex	Clinical findings	Gene	Transcript	Variant	Mode of inheritance and origin	OMIM	References
<b>Group (iv)</b>		<b>Sporadic cases</b>						
Patient_18	F	BHD, LD, bilateral hearing loss, cardiac malformation	<i>ANKRD11</i>	NM_013275.5	c.1940_1941delinsT; p.Arg647Leufs*6	AD (de novo)	KBG syndrome (#148,050)	
Patient_19	M	ID, BHD	<i>ADNP</i>	NM_015339.4	c.1792C>T; p.Gln598Ter	AD (de novo)	Helsmoortel-van der Aa syndrome (#615,873)	
Patient_20	M	GR, ID, compatible with Cornelia de Lange	<i>NIPBL</i>	NM_133433.4	c.385T>C; p.Ser129Pro	AD (de novo)	Cornelia de Lange syndrome 1 (#122,470)	
Patient_21	F	ID, GR, facial dysmorphism, compatible with Cornelia de Lange	<i>NIPBL</i>	NM_133433.4	c.5329-15A>G	AD (de novo)	Cornelia de Lange syndrome 1 (#122,470)	
Patient_22	F	ID, DD	<i>NSD1</i>	NM_022455.4	c.2276C>G; p.Ser759Ter	AD (de novo)	Sotos syndrome 1 (#117,550)	
Patient_23	F	PDD, macrocephaly, hypotonia	<i>CSPP1</i>	NM_024790.6	c.363_367delTAAAT; p.Leu123Rfs*19 (pat) c.2243_2244delAA; p.Glu750Gfs*30 (mat)	AR	Joubert syndrome 21 (#615,636)	
Patient_24	M	PDD	<i>EMC1</i>	NM_015047.2	c.797T>G; p.Leu266Ter (mat) c.285 T>C; p.Phe953Ser (pat)	AR	Cerebellar atrophy, visual impairment, and PDD (#616,875)	
Patient_25	F	ID, DD, microcephaly, hypotonia	<i>DDX3X</i>	NM_001356.3	c.1415A>G; p.His472Arg	X-linked (de novo)	ID, X-linked 102 (#300,958)	
Patient_26	M	Phenotype compatible with Opitz syndrome	<i>MID1</i>	NM_0000381.3	c.602_605del; p.Val201GlyfsTer11	X-linked*	Opitz GBBB syndrome, type I (#300,000)	
Patient_27	F	Dysmorphic features and LD	<i>PPP2R5D</i>	NM_006245.3	c.592G>A; p.Glu198Lys	AD (de novo)	Mental retardation, autosomal dominant 35 (#616,355)	
Patient_28	M	ASD, ID	<i>WAC</i>	NM_016628.5	c.1280_1281delCTinsGAG	AD (de novo)	Desanto-Shinawi syndrome (#616,708)	
Patient_29	M	ID and dysmorphic features	<i>KCNT2</i>	NM_198503.5	c.569G>A; p.Arg190His	AD (de novo)	Developmental and epileptic encephalopathy 57 (#617,771)	[20, 21]
Patient_30	F	ID, absent speech, dysmorphic features	<i>CNOT3</i>	NM_014516.3	c.169C>T (p.Arg57Trp)	AD (de novo)	Intellectual developmental disorder with speech delay, autism, and dysmorphic facies	
<b>Group (v)</b>		<b>Whole genome sequencing</b>						
Patient_31	M	Severe ID, absent speech, BHD	<i>TRAPPC9</i>	NM_031466.7	c.1037G>A; p.Gly346Glu (mat) arr8q24(141382973_141473138) × 1 (pat)	AR	ID, autosomal recessive 13 (#613,192)	[10]

ID: intellectual disability; DD: developmental delay; GR: growth retardation; ASD: autistic spectrum disorder; SD: speech delay; BHD: behavioural disorder; PDD: psychomotor development delay; LD: learning disabilities; \*mother not available

**Table 3** Diagnostic yield by WES and WGS in the analyzed cohorts

Total positive diagnosis	Cases (n = 31)
Autosomal dominant	15 (48.4%)
De novo	14
Inherited	1
X-linked	10 (32.3%)
Dominant	1
Recessive	9
Autosomal recessive	5 (16.1%)
Homozygous	1
Compound heterozygous	4
Imprinting	1 (3.2%)

genes, namely, *VPS13B* (patient 2), *CSPP* (patient 23) and *EMC1* (patient 24) and one consanguineous family with a homozygous variant in the *PYCR1* (patient 7) gene (Table 2). In addition, WGS allowed the identification of an structural variant (SV) and a missense variant in compound heterozygosis in the *TRAPPC9* (patient 31) gene (Table 2).

#### Imprinting inheritance

Finally, mutations in imprinted genes represented 3% of our diagnostic cohort (Table 3). Only one case (patient 6) was found to carry a LOF variant in the imprinted *UBE3A* gene inherited from her asymptomatic mother.

#### Discussion

The past decade has seen a rapid development of advancements in genetics and genomics, allowing an unprecedented identification of mutations that are involved in complex neurodevelopmental conditions. In the last years, WES has emerged as a comprehensive and cost-effective approach for discovering pathogenic variants in rare monogenic diseases.

In this report we present pathogenic variants detected by WES in a cohort of 87 families with NDDs. The resulting diagnostic yield in the sequenced cohort was 34%. After considering testing strategy based on family series, our analysis determined that extended families approach had a positive impact on diagnostic yield. However, the most common type of family that arrives to a medical genetics consultation for NDDs counselling is based on an isolated cases or a couple of siblings within a family, representing in our cohort more than 70% (64/87) of the patients. On the other hand, no differences were found in the diagnosis rate between the other cohort structures (siblings, trios and sporadic cases). Nevertheless, trio-based analysis reduced the turnaround time of genetic diagnosis (data not shown) and despite not having WES

data from parents, segregation analysis were in most cases essential in order to definitively classify variants as disease causing.

De novo variation has been proposed as a major pathomechanism in NDDs representing approximately 20–30% of ASD and ID cases [22, 23]. Our results identified monogenic aetiology due to de novo mutation in 48.4% (15/31) of the patients with genetic diagnosis, 14 of them in autosomal genes and 1 in an X-linked gene, representing the main source of genetic alteration in our cohort. Besides de novo variants, XL recessive disorders were found in 29% of the diagnosed patients (9/31), highlighting the role of the X-chromosome in underlying genetic basis of NDDs in affected males. To date, there are more than 140 genes on the X-chromosome related to cognition associated with ID [24]) and thus, variants in the X chromosome greatly contribute to ID in males.

Regardless the pattern of inheritance, most of the identified genes are implied in syndromic neurodevelopmental disorders, such as Cornelia de Lange syndrome (CdLS), Dent disease or Cohen syndrome, which are considered clinically recognizable syndromes. Nevertheless, several of the identified variants are rare and in some cases the clinical manifestations do not correspond to the classic syndromic forms or clinical features are not really specific. For example, patient 29 was referred to clinical consultation for presenting ID and dysmorphic features. Results from NGS identified a de novo previously reported variant in the *KCNT2* gene, usually associated with Developmental and epileptic encephalopathy type 57 (OMIM # 617771). This syndrome is characterized by epileptic encephalopathy, global DD with hypotonia, variably impaired intellectual development and poor or absent language. The *KCNT2*-associated developmental and epileptic encephalopathy comprises West syndrome, Lennox-Gastaut syndrome as well as epilepsy of infancy with migrating focal seizures. Currently, there are 4 patients with de novo mutations in the Arg190 residue described in the literature [20, 21]). Our patient was 3 years old and presented with development and speech delay, bruxism, attention deficit, repetitive behaviors and physical contact avoidance. He presented some facial dysmorphisms such as elongated face with a broad-based nose, a short filter and little marked slight prognathism. He had blonde hair color, unlike her parents and siblings who had brown hair color. Curiously there were no epilepsy attacks. In consonance with our results, the patients recently reported by Jackson and collaborators (2021) present a recognizable neurodevelopmental disorder without epilepsy [21]). All patients described up to date with variants affecting p.Arg190 of the *KCNT2* gene present ID, neonatal hypotonia, hirsutism, thick hair, prominent eyebrows, long eyelashes, and diastema [21].



**Fig. 1** Clinical pictures of 9 of the patients with genetic diagnosis identified in this study. Clinical manifestations are summarised in Table 2

NGS is also unveiling a large number of VUS, which effectively represent mutations where the pathogenicity and the function of the gene involved is unclear, hampering their endorsement to a clinical phenotype. The

process of variant prioritization is performed based on several factors, including their possible presence in databases such as gnomAD, ClinVar and HGMD; the relevance to the reported clinical features, the impact of

the variant or the associated mode of inheritance, among others. In order to determine whether a VUS may be associated with a patient's phenotype, additional studies are needed including segregation analysis or functional studies. There are several possibilities of functional tests that can be performed to answer questions about the function of genes and the functional consequences of genetic variants. RNA studies are the easiest and fastest ways to determine if a variant of unknown significance could be pathogenic. Nevertheless it is only indicated in variants that produce changes in mRNA expression levels, for example, variants that cause alternative splicing events or that cause loss of function. Protein expression studies have already proven their value to demonstrate pathogenicity of genetic variants. Another strategy is to test the genetic variants in model systems (for example cell culture models or animal models) [reviewed in 25].

Another approach is the reinterpretation of VUS in a given period later from initial analysis, which has been proven to increase diagnostic efficacy [26]. In our cohort, we identified an ultra-rare missense variant in the *NIPBL* gene (c.385T>C; p.Ser129Pro; rs141976717) in case 20 clinically diagnosed with CdLS. This patient has global cognitive and growth retardation, microcephaly, and dysmorphism. Physical examination revealed obvious clinical signs of CdLS such as long philtrum, long eyelashes, high-arched palate and anteverted nares, hypospadias (Fig. 2). Other clinical manifestations were the absence of language, mild hypertonia of the lower and upper extremities, and syndactyly of the 2nd and 3rd toes. Initially, this patient underwent to an external laboratory to perform a CdLS gene panel focused on the five current genes associated to this syndrome. A missense c.385T>C variant in a non-conserved position was reported. In silico tools indicated benign computational verdict based

on 10 benign versus 2 pathogenic predictions and it was described in gnomAD database with an ultra low allelic frequency (0.004%). On the basis of these observations, this variant was initially discarded to be responsible of the patient's phenotype. WES was later performed in this patient and no additional genetic alterations were identified associated to NDDs. After reanalysis this variant, segregation analysis was performed revealing that the variant c.385T>C was originated de novo in the patient. Although we cannot rule out the presence of an intronic variant in any of the CdL genes, taking into account that the majority of CdLS patients have de novo heterozygous missense variant in *NIPBL*, this variant was reclassified as likely pathogenic.

Despite the significant improvement in diagnostics of rare diseases using WES, a significant proportion of cases with a likely genetic aetiology remain undiagnosed. A step further WES, is WGS which allows the identification of sequence variants in non-coding regions making up the vast majority of human DNA [27]). However, there is evidence that the increased rate of WGS compared to WES is due to its efficiency for detecting potential disease-causing mutations within single nucleotide variants (SNV) in regions not properly covered by WES or SV not detected by this technology [28]). In our cohort, WGS was applied to 12 out of the 57 patients which remain undiagnosed, identifying only causative variants in 1 of the patients (Table 1). This patient (patient 31) was found to carry a missense variant and an intragenic deletion of 90 Kb, not detected neither by CMA nor WES [10].

The application of NGS technologies and the development of several web-based tools such as GeneMatcher (<http://www.genematcher.org>) to share phenotype and genotype data to broaden the exchange between scientific and medical teams have allowed the discovery of new



**Fig. 2** Clinical picture of patient 20 carrier of a de novo variant in *NIPBL* gene (NM\_133433.4: c.385T>C; p.Ser129Pro)

genes responsible for the disease. In our cohort, we identified two families diagnosed with pathogenic variants in *THOC2* gene (Table 2), who were previously reported as a series of cases describing the phenotype associated with THOC-related syndrome [11]. In addition, GeneMatcher tool allowed the delineation of a genome-wide DNA methylation signature for SETD1B-related syndrome, in which the patient herein reported (patient 13) was also included [12]. Interestingly, one of the patients of our cohort was found to carry a de novo likely pathogenic variant in the *CHD9* gene, which is a strong genetic candidate for NDDs that not yet been associated with ID. The CHD (Chromodomain Helicase DNA-binding protein) gene family consists of nine genes (*CHD1–CHD9*), and pathogenic variants in all CHD genes except *CHD9* are implicated in NDDs [29]. Recently, other groups have suggested a possible implication of *CHD9* variants in ASD patients [30–32]. Given the role of this gene as a transcriptional regulator and chromatin decompressor and the involvement of the rest of the CHD family genes in several NDDs, we consider that *CHD9* is a candidate gene for ASD.

## Conclusions

Currently, there are more than 2000 genes associated with NDDs making the implementation of NGS in the diagnostic flowchart of NDDs essential. Our results support the use of NGS for genetic diagnosis of NDDs as a first-tier test since it has a clear higher diagnostic rate compared to CMA (33% vs 10%). However, it should be taken into consideration that pathogenic CNVs were discarded by CMA in all patients presented in this report before performing NGS studies which might be a limitation of the study. On the other hand, new bioinformatic pipelines allow the detection not only of SNV but also reliable CNVs through WES data which further reinforces this change in diagnostic algorithm. The genetic architecture of NDDs is complex and interdisciplinary approaches combining genetics, functional genomics and biological models will be essential.

## Abbreviations

ACMG: American College of Medical Genetics and Genomics; ADHD: Attention deficit hyperactivity disorder; ASD: Autism spectrum disorder; CdLS: Cornelia de Lange syndrome; CGH: Comparative genomic hybridization; CHD: Chromodomain Helicase DNA-binding; CMA: Chromosomal microarray; CNV: Copy number variant; DD: Developmental delay; HGMD: Human Gene Mutation Database; GR: Growth retardation; ID: Intellectual disability; LOF: Loss-of-function; MLPA: Multiplex Ligation-dependent Probe Amplification; NDDs: Neurodevelopmental disorders; NGS: Next-generation sequencing; OMIM: Online Mendelian Inheritance in Man; SNV: Single nucleotide variant; SV: Structural variant; VUS: Variant of unknown significance; WES: Whole exome sequencing; WGS: Whole genome sequencing.

## Acknowledgements

We want to thank the patients and families for their collaboration on this study. We also want to thank the Dr. Christian Gilissen's group from the Department of Human Genetics, division Genomic Diagnostic, Radboud University Medical Center, Nijmegen, The Netherlands for their support on Bioinformatic analysis.

## Authors' contributions

MIAM: analyzed the NGS data, conducted the literature search and analyses and draft, and revised the manuscript. AS: clinical genetist who examined the patients and revised the manuscript. LRR: performed previous genetic studies from the patients (conventional karyotyping and Fragile X syndrome), participated in the original design of the study and revised the manuscript. JC: performed bioinformatic analysis from part of the NGS data herein reported and revised the manuscript. RR: performed part of the trio and sibling whole exome sequencing data. SP: performed segregation analysis and revised the manuscript. IM: supervised conceptualization of the study and the collection of data, had full access to all of the study data and assisted in the interpretation of the data, and draft and revised the manuscript before submission. All authors read and approved the final manuscript.

## Funding

This work was supported by the Instituto de Salud Carlos III (COHORTES/ Programa de Enfermedades Raras No Diagnosticadas, ENOD), the PERIS-Departament de Salut, Fondo Europeo de Desarrollo Regional (FEDER) "una manera de hacer Europa", CERCA Programme/Generalitat de Catalunya and Agència de Gestió d'Ajuts Universitaris i de Recerca from the Autonomous Catalan Government (2017 SGR1134), Fundación Mútua Madrileña (FundMM\_2019) and Fundación Alicia Koplowitz (AKOPLOWITZ18\_001). The 'CIBER de Enfermedades Raras' is an initiative of the Instituto de Salud Carlos III.

## Availability of data and materials

Please contact corresponding author for data requests.

## Declarations

### Ethics approval and consent to participate

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Boards of the Hospital Clinic of Barcelona (2011/6625; approved 14 April 2011). Informed consent was obtained from all subjects involved in the study.

### Consent for publication

Consent for publication was obtained from all subjects involved in the study.

### Competing interests

All authors declare that they have no competing interests.

### Author details

<sup>1</sup>Biochemistry and Molecular Genetics Department, Hospital Clinic of Barcelona, IDIBAPS (Institut de Investigacions Biomèdiques August Pi I Sunyer), Barcelona, Spain. <sup>2</sup>CIBER de Rare Diseases (CIBERER), Barcelona, Spain.

<sup>3</sup>Department of Human Genetics, Radboud UMC, Nijmegen, The Netherlands.

<sup>4</sup>Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain. <sup>5</sup>Department of Dermatology, Hospital Clínic of Barcelona and Melanoma Unit, IDIBAPS, Barcelona, Spain.

Received: 9 September 2021 Accepted: 6 February 2022

Published online: 19 February 2022

## References

- Moreno-De-Luca A, Myers SM, Challman TD, Moreno-De-Luca D, Evans DW, Ledbetter DH. Developmental brain dysfunction: revival and expansion of old concepts based on new genetic evidence. *Lancet Neurol.* 2013;12(4):406–14. [https://doi.org/10.1016/S1474-4422\(13\)70011-5](https://doi.org/10.1016/S1474-4422(13)70011-5).
- Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemssen MH, et al. Genome sequencing identifies major causes of severe

- intellectual disability. *Nature*. 2014;511(7509):344–7. <https://doi.org/10.1038/nature13394>.
3. Stevenson RE, Procopio-Allen AM, Schroer RJ, Collins JS. Genetic syndromes among individuals with mental retardation. *Am J Med Genet A*. 2003;123A(1):29–32.
  4. Deng Y, Pan W. Significance testing for allelic heterogeneity. *Genetics*. 2018;210:25–32. <https://doi.org/10.1002/ajmg.a.20492>.
  5. Posey JE, O'Donnell-Luria AH, Chong JX, Harel T, Jhangiani SN, Coban Akdemir ZH, et al. Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet Med*. 2019;21:798–812. <https://doi.org/10.1038/s41436-018-0408-7>.
  6. Trujillano D, Bertoli-Avella AM, Kumar Kandaswamy K, Weiss ME, Köster J, Marais A, et al. Clinical exome sequencing: results from 2819 samples reflecting 1000 families. *Eur J Hum Genet*. 2017;25:176–82. <https://doi.org/10.1038/ejhg.2016.146>.
  7. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, 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(5):749–64. <https://doi.org/10.1016/j.ajhg.2010.04.006>.
  8. Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, 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(11):2413–21. <https://doi.org/10.1038/s41436-019-0554-6>.
  9. Manickam K, McClain MR, Demmer LA, Biswas S, Kearney HM, Malinowski J, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(11):2029–37. <https://doi.org/10.1038/s41436-021-01242-6>.
  10. Álvarez-Mora MI, Corominas J, Gilissen C, Sanchez A, Madrigal I, Rodríguez-Revenga L. Novel Compound Heterozygous Mutation in *TRAPPC9* Gene: the relevance of whole genome sequencing. *Genes*. 2021;12:557–64. <https://doi.org/10.3390/genes12040557>.
  11. Kumar R, Gardner A, Homan CC, Douglas E, Mefford H, Wiczkorek D, et al. Severe neurocognitive and growth disorders due to variation in *THOC2*, an essential component of nuclear mRNA export machinery. *Hum Mutat*. 2018;39(8):1126–38. <https://doi.org/10.1002/humu.23557>.
  12. Krzyzewska IM, Maas SM, Henneman P, Lip KVD, Venema A, Baranano K, et al. A genome-wide DNA methylation signature for SETD1B-related syndrome. *Clin Epigenet*. 2019;11(1):156. <https://doi.org/10.1186/s13148-019-0749-3>.
  13. Madrigal I, Álvarez-Mora MI, Karlberg O, Rodríguez-Revenga L, Elurbe DM, Rabionet R, et al. Efficient application of next-generation sequencing for the diagnosis of rare genetic syndromes. *J Clin Pathol*. 2014;67(12):1099–103. <https://doi.org/10.1136/jclinpath-2014-202537>.
  14. Madrigal I, Álvarez-Mora MI, Rosell J, Rodríguez-Revenga L, Karlberg O, Sauer S, et al. A novel splicing mutation in the *IQSEC2* gene that modulates the phenotype severity in a family with intellectual disability. *Eur J Hum Genet*. 2016;24(8):1117–23. <https://doi.org/10.1038/ejhg.2015.267>.
  15. Madrigal I, Rabionet R, Álvarez-Mora MI, Sanchez A, Rodríguez-Revenga L, Estivill X, et al. Spectrum of clinical heterogeneity of  $\beta$ -tubulin TUBB5 gene mutations. *Gene*. 2019;695:12–7. <https://doi.org/10.1016/j.gene.2019.02.002>.
  16. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitpiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44:D862–8. <https://doi.org/10.1093/nar/gkv1222>.
  17. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, et al. Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat*. 2003;21(6):577–81. <https://doi.org/10.1002/humu.10212>.
  18. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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*. 2015;17:405–24. <https://doi.org/10.1038/gim.2015.30>.
  19. Teresa-Rodrigo ME, Eckhold J, Puisac B, Pozojevic J, Parenti I, Baquero-Montoya C, et al. Identification and functional characterization of two intronic NIPBL mutations in two patients with Cornelia De Lange syndrome. *Biomed Res Int*. 2016;2016:8742939. <https://doi.org/10.1155/2016/8742939>.
  20. Ambrosino P, Soldovieri MV, Bast T, Turnpenny PD, Uhrig S, Biskup S, et al. De novo gain-of-function variants in *KCNT2* as a novel cause of developmental and epileptic encephalopathy. *Ann Neurol*. 2018;83(6):1198–204. <https://doi.org/10.1002/ana.25248>.
  21. Jackson A, Banka S, Stewart H, Genomics England Research Consortium, Robinson H, Lovell S, Clayton-Smith J. Recurrent *KCNT2* missense variants affecting p.Arg190 result in a recognizable phenotype. *Am J Med Genet A*. 2021;185(10):3083–91. <https://doi.org/10.1002/ajmg.a.62370>.
  22. Wang W, Corominas R, Lin GN. De novo mutations from whole exome sequencing in neurodevelopmental and psychiatric disorders: from discovery to application. *Front Genet*. 2019;10:258. <https://doi.org/10.3389/fgene.2019.00258>.
  23. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542(7642):433–8. <https://doi.org/10.1038/nature21062>.
  24. Tejada MI, Ibarluzea N. Non-syndromic X linked intellectual disability: current knowledge in light of the recent advances in molecular and functional studies. *Clin Genet*. 2020;97(5):677–87. <https://doi.org/10.1111/cge.13698>.
  25. Rodenburg RJ. The functional genomics laboratory: functional validation of genetic variants. *J Inher Metab Dis*. 2018;41(3):297–307. <https://doi.org/10.1007/s10545-018-0146-7>.
  26. Vears DF, Niemiec E, Howard HC, Borry P. Analysis of VUS reporting, variant reinterpretation and recontact policies in clinical genomic sequencing consent forms. *Eur J Hum Genet*. 2018;26(12):1743–51. <https://doi.org/10.1038/s41431-018-0239-7>.
  27. Turro E, Astle WJ, Megy K, Gräf S, Greene D, Shamardina O, et al. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature*. 2020;583(7814):96–102. <https://doi.org/10.1038/s41586-020-2434-2>.
  28. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci USA*. 2015;112(17):5473–8. <https://doi.org/10.1073/pnas.1418631112>.
  29. Marfella CG, Imbalzano AN. The Chd family of chromatin remodelers. *Mutat Res*. 2007;618(1–2):30–4. <https://doi.org/10.1016/j.mrfmmm.2006.07.012>.
  30. Wu J, Yu P, Jin X, Xu X, Li J, Li Z, et al. Genomic landscapes of Chinese sporadic autism spectrum disorders revealed by whole-genome sequencing. *J Genet Genom*. 2018;45(10):527–38. <https://doi.org/10.1016/j.jgg.2018.09.002>.
  31. Li J, Wang L, Guo H, Shi L, Zhang K, Tang M, et al. Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. *Mol Psychiatry*. 2017;22(9):1282–90. <https://doi.org/10.1038/mp.2017.140>.
  32. Lim ET, Uddin M, De Rubeis S, Chan Y, Kamumbu AS, Zhang X, et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat Neurosci*. 2017;20(9):1217–24. <https://doi.org/10.1038/nn.4598>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

