Pediatric Neurology 99 (2019) 47–54

Contents lists available at ScienceDirect

Pediatric Neurology

journal homepage: www.elsevier.com/locate/pnu

Original Article

Proband-Only Clinical Exome Sequencing for Neurodevelopmental Disabilities

Se Hee Kim, MD [a, †], Borahm Kim, MD [b, †], Joon Soo Lee, MD, PhD [a], Heung Dong Kim, MD, PhD [a], Jong Rak Choi, MD, PhD [b], Seung-Tae Lee, MD, PhD [b, **], Hoon-Chul Kang, MD, PhD [a, *]

[a] Division of Pediatric Neurology, Department of Pediatrics, Severance Children's Hospital, Yonsei University College of Medicine, Epilepsy Research Institute, Seoul, Korea
[b] Department of Laboratory Medicine, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

ARTICLE INFO

Article history:
Received 19 December 2018
Accepted 20 February 2019
Available online 27 February 2019

Keywords:
Intellectual disability
Neurodiagnostics
Children

ABSTRACT

Background: Whole exome sequencing on family trios gives the highest diagnostic yield, but high cost limits its application. Here, we performed proband-only clinical exome sequencing in a population of patients with neurodevelopmental disabilities and tested the diagnostic yield.
Methods: This observational, retrospective study included 108 unrelated patients with neurodevelopmental disabilities who underwent clinical exome sequencing at the outpatient clinics of the Severance Children's Hospital, Seoul, South Korea, between March 2017 and May 2018. Clinical exome sequencing targeted 4503 disease-causing genes.
Results: The overall diagnostic rate was 38.0% (41 of 108) when proband-only clinical exome sequencing was performed without additional parental testing. Four sequence variants were reclassified as likely pathogenic after parental testing, representing an additional 3.7% of the diagnostic yield. The final diagnostic rate was 41.7% (45 of 108). Of 45 patients with genetic abnormalities, a total of 38 sequence variations were detected in 33 (30.6%) patients with five homozygous cases, and 13 chromosomal copy number variants were detected in 12 (11.1%) patients. Novel variants of known causal genes for neurodevelopmental disabilities were detected in 18 (16.7%) patients. These were variants that could be reclassified as likely pathogenic if the *de novo* nature of the mutation was confirmed after testing of parental samples.
Conclusions: Proband-only clinical exome sequencing is a practical diagnostic tool that may be implemented in the clinical setting for patients with neurodevelopmental disabilities. A cost-effective approach to neurodevelopmental disabilities would be a proband-only clinical exome sequencing followed by parental testing of selective candidate variants.

Financial disclosure: The authors have no financial relationships relevant to this article to disclose.

Funding source: This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, Republic of Korea (grant number: HI18C0586).

Potential conflicts of interest: The authors have no conflicts of interest relevant to this article to disclose.

Contributor's statement: Se Hee Kim and Borahm Kim conceptualized and designed the study, drafted the initial manuscript, and reviewed and revised the manuscript. Joon Soo Lee, Heung Dong Kim, and Jong Rak Choi designed the data collection instruments, collected data, carried out the initial analyses, and reviewed and revised the manuscript. Seung-Tae Lee and Hoon-Chul Kang conceptualized and designed the study, coordinated and supervised data collection, and critically reviewed the manuscript for important intellectual content. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

* Communications should be addressed to: Kang; Division of Pediatric Neurology; Department of Pediatrics; Severance Children's Hospital; Yonsei University College of Medicine; Epilepsy Research Institute; 50-1 Yonsei-ro, Seodaemun-Gu; Seoul 03722, South Korea.

** Communications should be addressed to: Lee; Department of Laboratory Medicine; Severance Hospital; Yonsei University College of Medicine; 50-1 Yonsei-ro, Seodaemun-gu; Seoul 03722, South Korea.

E-mail addresses: lee.st@yuhs.ac (S.-T. Lee), hipo0207@yuhs.ac (H.-C. Kang).

† Equal contributors.

https://doi.org/10.1016/j.pediatrneurol.2019.02.017
0887-8994/

Introduction

Etiologic diagnosis of neurodevelopmental disabilities (NDD) and intellectual disabilities (ID) is often challenging because of its phenotypical and etiologic heterogeneity. A molecular diagnosis is often required to predict prognosis. Conventional genetic tests, such as karyotype or chromosomal microarray (CMA) analysis, can only diagnose 3.5% to 13.0% or ~10% cases, respectively,[1–5] whereas metabolic tests can diagnose additional 0.6% to 1.3% cases.[6–8]

Recently, several multicenter studies found that whole exome sequencing could diagnose 16% to 57% of the remaining undiagnosed NDD or ID cases.[9–14] Notably, an additional 27% patients in whom a conclusive cause could not be identified with CMA could be diagnosed with whole exome sequencing, indicating that whole exome sequencing is a highly cost-effective diagnostic tool for ID.[15] Indeed, NDD or ID and related disorders were the most common referral reason for whole exome sequencing in a recent study performed with 2000 patients.[12]

However, many questions remain to be answered to use exome sequencing for the etiologic diagnosis of NDD or ID in the clinical setting.[16] For example, the diagnostic yield of proband-only clinical exome sequencing for NDD or ID has not been determined. Populations that would have a high diagnostic yield remain to be defined.

Here, we sought to identify the underlying genetic defects in a population of patients with NDD using proband-only clinical exome sequencing. In addition, we attempted to address issues that may arise using exome sequencing for NDDs in the clinical setting.

Methods

Patients and clinical information

From March 2017 to May 2018, a total of 108 unrelated patients with NDD or ID underwent proband-only clinical exome sequencing at the Neurodevelopmental Clinic of Severance Children's Hospital.

All patients met the following inclusion criteria: (1) intelligence quotient (IQ) or developmental quotient less than 70, (2) visited neurology clinic due to developmental issues, (3) parents wanted to identify etiology, (4) no causative abnormalities detected with previous genetic or metabolic tests, and (5) absence of magnetic resonance imaging (MRI) findings suggestive of a specific disorder. NDD was defined as a delay of motor, language, or social development, exceeding the upper limit of normal, as described in the Denver Developmental Screening Test or other equivalent tests, such as the Bayley Scales (Bayley-III). ID was diagnosed according to the following criteria provided by the American Association on Intellectual and Developmental Disabilities: (1) significant limitation in intellectual function, (2) significant limitation in adaptive behavior, and (3) origin before age 18 years.[17]

Previous genetic tests in each patient varied but usually included chromosomal analysis, Sanger sequencing for specific genes, and multiplex ligation-dependent probe amplification, using the SALSA MLPA P245 Microdeletion Syndrome kit (MRC Holland, Amsterdam, The Netherlands). These were mainly performed before referral to Severance Children's Hospital. Brain MRI was performed according to standardized neuro-MRI protocols, including coronal and axial views, on T2-weighted and fluid-attenuated inversion recovery at 1.5 to 3.0 T. Assays for metabolic disorders included plasma amino acid analysis, urine organic acid analysis, acylcarnitine profile (total and free carnitine levels), blood gas analysis, and measurement of homocysteine, serum lactate, pyruvate, and ammonia levels.

Information regarding developmental status, seizure history, admission to the neonatal intensive care unit, and head circumference was also collected. This study was approved by the Institutional Review Board of Severance Hospital (IRB 4-2018-0711).

Clinical exome sequencing

The xGen Inherited Diseases Panel (Integrated DNA Technologies, Coralville, IA, USA), which includes 4503 candidate genes, was used for exome sequencing. This panel includes genes associated with various neurodevelopmental disorders, including autism spectrum, epilepsy, and seizure disorders, as well as X-linked ID.

Genomic DNA was extracted using the QIAamp DNA Blood Mini Kit (QIAGEN, Hilden, Germany). Approximately 1.5 μg of genomic DNA was fragmented into 150- to 250-bp segments using the Bioruptor Pico Sonication System (Diagenode, Denville, NJ, USA). The resulting DNA was then end-repaired and ligated to Illumina adapters (Illumina, San Diego, CA, USA), and sequence indexes were added. Small fragments of ~100 bp and unligated adapters were removed using AMPure XP purification (Agencourt Bioscience, Beverly, MA, USA). Sequencing libraries were hybridized with probes, unbound DNA was removed, and capture probes were digested. Enriched DNA was then sequenced on a NextSeq 550 instrument (Illumina) with 2 × 151-bp reads.

Next-generation sequencing data analysis

Reads were aligned to human genomic reference sequences (GRCh37) using the Burrows-Wheeler alignment tool (0.7.12).[18] HaplotypeCaller in the Genome Analysis Toolkit (GATK) package (3.8-0) was used to identify single nucleotide variants (SNVs) and small indels, and Pindel (0.2.0) was used to detect large indels, All mutations were annotated using ANNOVAR and VEP (87) software.[19,20] Variants were further examined by visual inspection

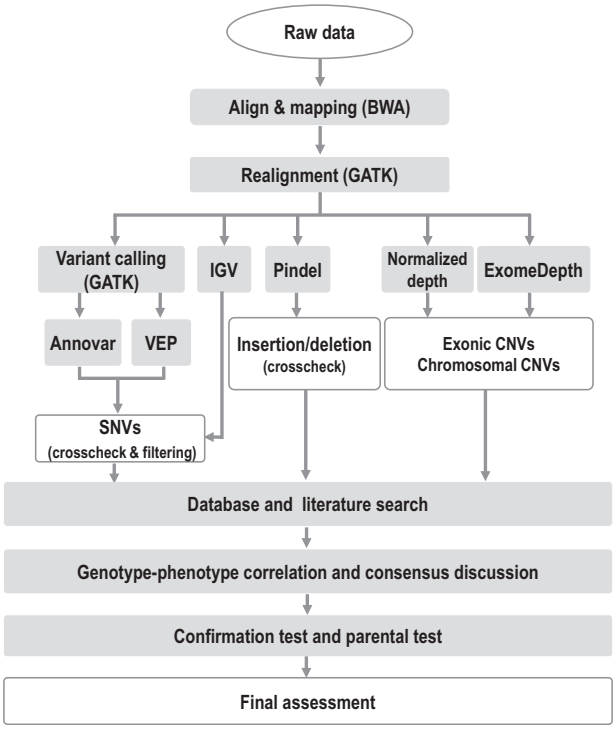

FIGURE 1. Flowchart detailing the workflow for clinical exome sequencing and downstream analysis.

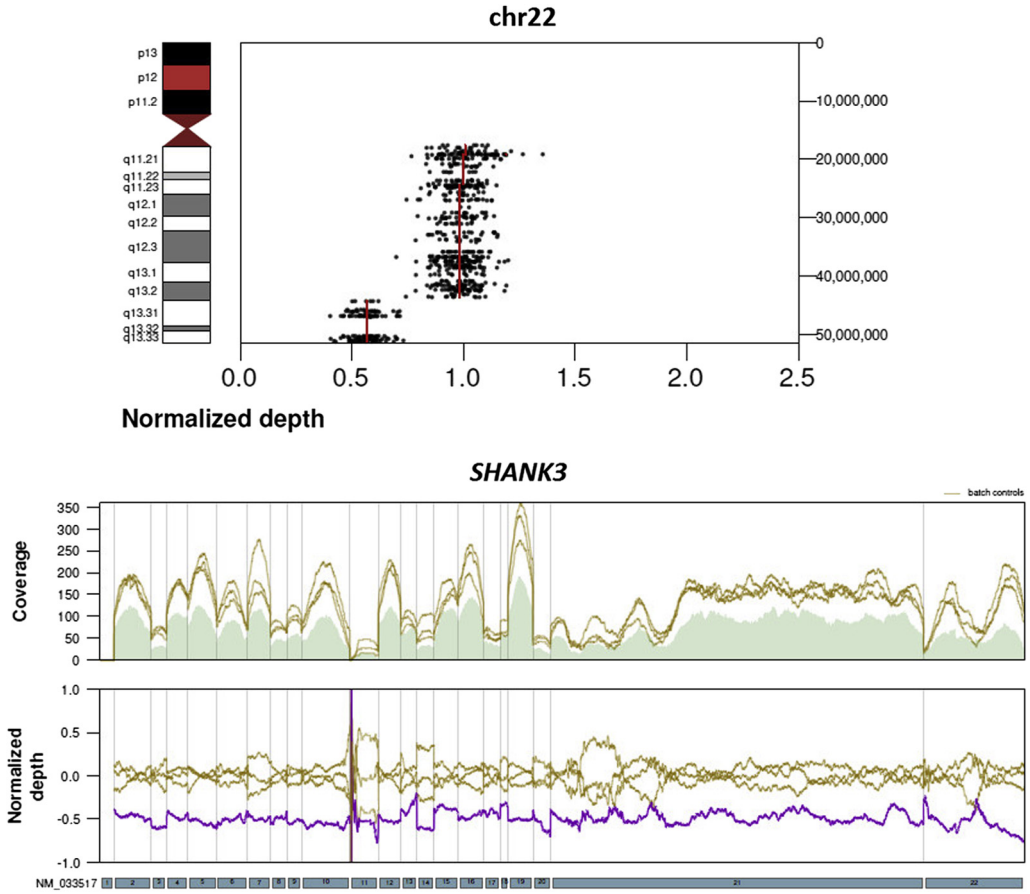

FIGURE 2. Example of a chromosomal copy number variation. Microdeletion of the 22q13 region was identified by exome sequencing. The color version of this figure is available in the online edition.

using the Integrative Genomic Viewer.[21] Variants confirmed to be true positives were manually scored and curated by searching the literature and databases such as ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). ExomeDepth (1.1.10) in the R package was used to detect chromosomal and exon-level copy number variants (CNVs) in target regions, followed by visualization using a base-level read depth normalization algorithm implemented in the DxSeq Analyzer (Dxome, Seoul, Korea).[22]

The following databases were used for variant annotation: Online Mendelian Inheritance in Man, Human Gene Mutation Database, ClinVar, dbSNP, 1000 Genome, Exome Aggregation Consortium, Exome Sequencing Project, and the Korean Reference Genome Database. The pathogenicity of missense variants was predicted using five *in silico* prediction algorithms, including Sorting Tolerant from Intolerant, Polymorphism Phenotyping v2, MutationTaster, MutationAssessor, and Functional Analysis through Hidden Markov Models, implemented in dbNSFP, version 3.0a. Effects on splicing were predicted using SPIDEX, version 1.0, and dbscSNV, version 1.1.

Identified variants were described according to nomenclature recommendations of the Human Genome Variation Society (http://www.hgvs.org/mutnomen). Interpretation of variants followed the five-tier classification system recommended by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, using a step-by-step approach. Most nonsense, frameshift, and splice site variants were classified as pathogenic or likely pathogenic if the same variants had never been reported in the normal population and the mechanism of pathogenesis was loss of function. Missense variants were usually classified as variants of unknown significance (VOUS) in the absence of previous reports. Pathogenicity was determined based on multiple factors, including previous evidence, recurrence, predicted functional effects, and the absence of the variant among healthy controls.

CNV information was obtained by comparison of read depths among samples in the same batch, using a base-level normalization algorithm, and subjected to visual verification (Fig 1). For chromosomal CNVs, variations were classified as pathogenic or likely pathogenic if they had one of the following characteristics: (1) large deletions or duplications involving more than 3 Mb, (2) located in regions associated with a well-known syndromic disorder, such as Phelan-McDermid syndrome, which involves a 22q13 deletion (Fig 2); or (3) microdeletions or duplications encompassing well-characterized genes with a defined pathogenicity, such as the nuclear receptor binding SET domain protein 1 (*NSD1*) gene found in the 5q35 deletion.

Phenotype review, consensus discussion, parental tests, and secondary confirmation tests

The last step in our process involved determination of the clinical significance of variants as assessed by clinical geneticists and the patients' attending physicians. Clinicians performed an in-depth review of each patient phenotype and provided an opinion from their point of view. Each case was further reviewed and discussed in a monthly consensus meeting attended by all laboratory

TABLE 1
Characteristics of Patients With and Without Pathogenic Mutations

Characteristics	Group With Pathogenic Variants (n = 45)	Group Without Pathogenic Variants (n = 63)	*P* value
Age, yr, median (interquartile range)	4.0 (1.9, 6.6)	4.6 (2.6, 7.6)	0.422
Male sex, n (%)	35 (77.8)	34 (54.0)	0.011
IQ level, n (%)			
50-70	10 (22.2)	15 (24.2)	0.894
36-49	9 (20.0)	10 (16.1)	
20-35	11 (24.4)	13 (21.0)	
<20	15 (33.3)	24 (38.7)	
Onset of developmental delay, n (%)			
0-6 mo	19 (42.2)	24 (38.1)	0.837
6-12 mo	10 (22.2)	10 (15.9)	
13-24 mo	13 (28.9)	19 (30.2)	
25-36 mo	2 (4.4)	5 (7.9)	
37 mo	1 (2.2)	3 (4.7)	
Not specified		2 (3.2)	
MRI abnormality, n (%)			
No	27 (62.8)	27 (50.0)	0.208
Yes	16 (37.2)	27 (50.0)	
Prenatal problems, n (%)			
No	31 (68.9)	53 (84.1)	0.060
Yes	14 (31.1)	10 (15.8)	
Preterm infant, n (%)			
No	38 (84.4)	56 (88.9)	0.498
Yes	7 (15.6)	7 (11.1)	
Other anomaly, n (%)			
No	24 (53.3)	49 (77.8)	0.008
Yes	21 (46.7)	14 (22.2)	
Abnormal head circumference, n (%)			
Normocephaly	14 (34.2)	37 (64.9)	0.011
Microcephaly	19 (46.4)	14 (24.6)	
Macrocephaly	8 (19.6)	6 (10.5)	
Presence of seizures, n (%)			
No	21 (46.7)	31 (49.2)	0.795
Yes	24 (53.3)	32 (50.8)	

Abbreviations:
IQ = Intelligence quotient
MRI = Magnetic resonance imaging

personnel, bioinformaticians, geneticists, and clinicians. When pathogenic or likely pathogenic variants were consistent with patient phenotype, a final validation using other confirmatory assays and a parental study was planned. VOUS, especially missense variants, were prioritized according to population frequency, American College of Medical Genetics and Genomics score, and the patient's clinical phenotype. A follow-up parental study was scheduled to investigate whether the variant is a *de novo* mutation or an inherited one.

For identified variants requiring parental study or confirmation with a secondary method, PCR and Sanger sequencing were performed using a 3730 DNA Analyzer with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Chromosomal CNVs were confirmed with CMA using the CytoScan 750K array (Affymetrix, Santa Clara, CA, USA).

Statistical analysis

To compare the molecular diagnosis rate among groups, the Mann-Whitney U test, chi-square test, and Kruskal-Wallis test were utilized, followed by logistic regression analysis, with single or multiple variables. Variables included in multiple logistic levels were chosen by adapting stepwise as a variable selection method. Statistical analyses were computed using SAS software, version 9.4 (SAS Inc., Cary, NC, USA). *P* values <0.05 were considered statistically significant.

Results

A total of 108 patients (69 males and 39 females) with NDD or ID underwent clinical exome sequencing. From this cohort, 68.3% had an IQ level less than 50, 32.4% had congenital anomalies in other organ systems, and 43.5% had an abnormal head circumference. Clinical characteristics are provided in Table 1. For exome sequencing, a total of 30 million reads were generated for each sample. Median sequence coverage was 210, with an average of 99.1% of all targeted exons covered by at least 30 sequence reads.

Our overall molecular diagnostic rate was 41.7% (45 of 108), with 51 pathogenic or likely pathogenic variants identified in 45 cases. A total of 38 sequence variations were detected in 33 (30.6%) cases, and 13 chromosomal CNVs were detected in 12 (11.1%) patients (Table 2).

Our diagnostic rate increased from 26.9% (29 of 108) to 41.7% (45 of 108) after following steps. When only sequence variations were analyzed, our diagnostic yield was 26.9% (29 of 108). Analyzing CNVs in addition to sequence variation analysis increased the diagnostic yield to 38.0% (12 of 108). Thus the initial diagnostic rate was 38.0% (41 of 108) when proband-only clinical exome sequencing was performed without additional parental testing. Four sequence variants were reclassified as likely pathogenic after parental testing, representing an additional 3.7% (four of 108) of the diagnostic yield.

Genetic heterogeneity was prominent, and identified genes were different in all cases, excluding 10 patients who had mutations in *ANKRD11, ARID1B, DNM1L, NSD1*, or *SLC2A1*. The roles of identified genes were diverse and included transcription regulation (21.1%), solute transportation through cell membrane (13.2%), intracellular trafficking (7.9%), chromatin remodeling (5.3%), signal transduction (5.3%), and other (15.8%).

For the 38 cases with pathogenic or likely pathogenic sequence variations, inherited modes were autosomal dominant in 25 (65.8%), autosomal recessive in 10 (26.3%), and X-linked in three (7.9%) male patients (Table 3). Of these 25 variants in autosomal dominant disease-causing genes, four variants could be confirmed as likely pathogenic after confirming as *de novo* mutations using parent genetic tests. One variant in *SLC2A1* was inherited from a paternal somatic mosaic mutation. For the remaining 20 variants, pathogenicity could be determined based on the nature of variations, population frequency, and previous reports.

Chromosomal CNVs included 11 chromosomal microdeletions and two microduplications; these were detected in 12 patients, in chromosomes 1, 2, 3, 5, 8, 15, 17, 18, 20, and 22 (Table 2). The size of involved regions ranged from 1.6 to 23 Mb. Large deletions or duplications involving more than 3 Mb or CNVs occurring in regions associated with well-known syndromic disorders, such as Phelan-McDermid syndrome involving the 22q13 deletion, were classified as pathogenic. For relatively small microdeletions, pathogenicity was assumed if a critical gene (e.g., *NSD1* gene in the 5q35 deletion) was involved. CMA was performed in three patients who showed pathogenic or likely pathogenic CNVs by exome sequencing. In all three patients, results of CMA and exome sequencing were consistent. There was a small difference in estimated size variation between the tests, although the differences did not exceed 0.6 Mb.

Initially, 67 patients had VOUSs. Of 67 patients, 29 had VOUSs that required parental test for confirmation of *de novo* origin. Of the 29 patients, 27 had novel missense variants of probable causal NDD or ID genes and two patients had chromosomal copy number variations requiring parental testing for confirmation of *de novo* origin. Among the 29 patients, parental testing could be performed in

TABLE 2
Patients With Pathogenic or Likely Pathogenic Sequence Variants or Copy Number Variations

ID	Gene	Variants	Protein Sequence Variants	Zygosity	Inheritance	Variant Type	Previous Reports	*de novo*
P1	*ALDH3A2*	c.1157A>G	p.Asn386Ser	Homo	AR	Missense		
P2	*ANKRD11*	c.1903_1907delAAACA	p.Lys635GlnfsTer26	Hetero	AD	Frameshift	25424714, 27667800	
P3	*ANKRD11*	c.5350_5351delTC	p.Ser1784HisfsTer12	Hetero	AD	Frameshift		
P4	*ARID1B*	c.5561delG	p.Gly1854ValfsTer21	Hetero	AD	Frameshift		
P5	*ARID1B*	c.5947delC	p.His1983ThrfsTer38	Hetero	AD	Frameshift		
P6	*BCL11A*	c.568C>T	p.Gln190Ter	Hetero	AD	Nonsense		
P7	*CHD7*	c.4667dupC	p.Arg1557LysfsTer16	Hetero	AD	Frameshift	23024289	
P8	*CIC*	c.3795+1G>C		Hetero	AD	Splice site		
P9	*CLN6*	c.307C>T	p.Arg103Trp	Homo	AR	Missense		
P10	*CTNNB1*	c.1603C>T	p.Arg535Ter	Hetero	AD	Missense	23033978, 27915094	
P11	*DEAF1*	c.842G>A	p.Cys281Tyr	Hetero	AD	Missense		*de novo*
P12	*DLG3*	c.113_117delCTTAC	p.Pro38ArgfsTer49	Hemi	XL	Frameshift		
P13	*DNM1L*	c.1949T>G	p.Leu650Arg	Hetero	AD	Missense		*de novo*
P14	*DNM1L*	c.1247T>C	p.Leu416Pro	Hetero	AD	Missense		*de novo*
P15	*DYRK1A*	c.1400G>A	p.Arg467Gln	Hetero	AD	Missense	28053047	
P16	*FOXG1*	c.667C>T	p.Gln223Ter	Hetero	AD	Nonsense	24836831, 21441262	
P17	*GRIN2A*	c.2449A>G	p.Met817Val	Hetero	AD	Missense	24903190, 28126851	
P18	*KCNMA1*	c.492delG	p.Cys165AlafsTer26	Hetero	AD	Frameshift		
P19	*KDM5C*	c.2427_2430delTGAG	p.Glu810CysfsTer5	Hemi	XL	Frameshift		
P20	*KIF1A*	c.806C>A	p.Ala269Asp	Hetero	AD	Missense		*de novo*
P21	*KIF21A*	c.387dupA	p.His130ThrfsTer5	Hetero	AD	Frameshift		
P22	*MED13L*	c.6579delC	p.Val2194SerfsTer14	Hetero	AD	Frameshift		*de novo*
P23	*NSD1*	c.3549dupT	p.Glu1184Ter	Hetero	AD	Frameshift	12464997	
P24	*NSD1*	c.4417C>T	p.Arg1473Ter	Hetero	AD	Nonsense	12464997	
P25	*POGZ*	c.2517_2518delTC		Hetero	AD	Frameshift		
P26	*PPT1*	c.413C>T	p.Ser138Leu	Homo	AR	Missense	21990111	
P27	*RAB39B*	c.117dupC	p.Thr40HisfsTer51	Hemi	XL	Frameshift		
P28	*SCN2A*	c.4886G>T	p.Arg1629Leu	Hetero	AD	Missense	23935176	
P29	*SLC19A3*	c.1264A>G	p.Thr422Ala	Homo	AR	Missense	15871139	
P30	*SLC2A1*	c.377G>A	p.Arg126His	Hetero	AD	Missense	11603379, 18387950	
P31	*SLC2A1*	c.1156_1157dupAT	p.Pro387SerfsTer122	Hetero	AD	Frameshift		Paternal mosaicism
P32	*SMARCA2*	c.3562G>A	p.Ala1188Thr	Hetero	AD	Missense		
P33	*STRADA*	c.369_370delAC	p.Lys123AsnfsTer19	Homo	AR	Frameshift		

ID	Deletion/Duplication	Location	Known Syndromes	Estimated Size (Mb)	Zygosity	OMIM#	Reported Key Genes
P34	Deletion	1q43q44	1q43q44 and 1q44 microdeletion syndrome	6.2	Hetero	612337	*AKT3, ZBTB18, HNRNPU, FAM36A*
P35	Deletion	2q36.3	-	4.3	Hetero	-	*COL4A3, COL4A3*
P36	Duplication	3p26p24	-	23.9	Hetero	-	-
P37	Deletion	3p21.2p21.1	-	2.6	Hetero	-	-
P38	Deletion	5q35.2	Sotos syndrome	1.6	Hetero	117550	*NSD1*
P39	Deletion	8p23.3p23.2	-	4.5	Hetero	-	-
P39	Duplication	8p22	-	5.3	Hetero	-	-
P40	Deletion	22q11.2	22q11.2 deletion syndrome	2.5	Hetero	602054	*TBX1*
P41	Deletion	17q11.2	Smith-Magenis syndrome	3.1	Hetero	610883	*RAI1*
P42	Deletion	18q12.1q21.1	Del(18) (q12.2q21.1) syndrome	9.6	Hetero	601808	*SETBP1*
P43	Deletion	20p13p12.2	Alagille syndrome 1	7.6	Hetero	118450	*JAG1*
P44	Deletion	22q13.31q13.33	Phelan-McDermid syndrome	6.9	Hetero	606232	*SHANK3*
P45	Deletion	22q13	Phelan-McDermid syndrome	8.2	Hetero	606232	*SHANK3*

Abbreviations:
AD = Autosomal dominant
AR = Autosomal recessive
Hetero = Heterozygous
Hemi = Hemizygous
Homo = Homozygous
OMIM = Online Mendelian Inheritance in Man
XL = X-linked

eight additional patients, leaving 21 patients with VOUSs. After test, four (3.7%) patients who had novel missense variants of probable causal genes could be reclassified to have pathogenic or likely pathogenic genetic abnormalities, resulting in 63 patients with VOUSs (Fig 3). The other four patients had maternally or paternally inherited variants.

Features such as male gender (odds ratio [OR], 7.8; 95% confidence interval [CI], 2.4 to 25.7; $P = 0.0007$), microcephaly (OR, 7.7; 95% CI, 2.3 to 25.1; $P = 0.0008$), macrocephaly (OR, 4.9; 95% CI, 1.3 to 19.2; $P = 0.02$), and presence of other anomalies (OR, 4.9; 95% CI, 1.7 to 14.0; $P = 0.003$) were more prevalent among patients with identified pathogenic mutations (Table 1) than in the genetic-abnormality-negative group. A greater number of individuals with full-scale IQs between 36 and 49 were found in the group of patients with identified CNV abnormalities than in the CNV-negative group (six of 12 versus 13 of 96, respectively, $P = 0.0376$). Male gender (27 of 33 versus 42 of 75, $P = 0.101$) and microcephaly (15 of 33 versus 18 of 75, respectively, $P = 0.0281$) were more prevalent in patients with SNV abnormalities than in the SNV-negative group.

TABLE 3
Association Between Disease Inheritance and Clinical Variables

Characteristics	CNV (n = 12)	AD (n = 25)	AR (n = 5)	X-Linked (n = 3)	*P* value
Sex					
Male	8 (66.7)	21 (84.0)	3 (60.0)	3 (100.0)	0.3345
Female	4 (33.3)	4 (16.0)	2 (40.0)	0 (0.0)	
IQ level					
50-70	2 (16.7)	7 (28.0)	1 (20.0)	0 (0.0)	0.1997
36-49	6 (50.0)	3 (12.0)	0 (0.0)	0 (0.0)	
20-35	1 (8.3)	6 (24.0)	2 (40.0)	2 (66.7)	
<20	3 (25.0)	9 (36.0)	2 (40.0)	1 (33.3)	
Head circumference					
Normocephaly	4 (40.0)	9 (39.1)	0 (0)	1 (33.3)	0.5565
Microcephaly	4 (40.0)	11 (47.8)	3 (60)	1 (33.3)	
Macrocephaly	2 (20.0)	3 (13.0)	2 (40)	1 (33.3)	

Abbreviations:
AD = Autosomal dominant
AR = Autosomal recessive
CNV = Copy number variation
IQ = Intelligence quotient

Discussion

The etiology could be diagnosed in one-third of NDD or ID cases using proband-only clinical exome sequencing. This result suggests that proband-only clinical exome sequencing is an effective diagnostic tool that can be implemented for uncovering the etiology of NDD of ID in the clinical setting.

Recent data suggest that whole exome sequencing may be the most effective diagnostic tool for NDD or IDs. Notably, the diagnostic yield of whole exome sequencing surpasses that of other diagnostic tools for IDs, including metabolic tests, karyotype, and even CMA.[9–12] However, as whole exome sequencing is usually performed on family trios, high cost and unavailability of parental DNA limit its application. Here, we sought to identify the underlying genetic defects in a population of patients with NDD or ID using proband-only clinical exome sequencing. We found that additional parental testing was required only in one-fifth of our patients. Our study suggests that current knowledge of NDD or ID genetics can provide sufficient information that can be used to identify pathogenic variants in about 40% patients. Of note, a previous study has indicated that whole genome sequencing could provide yields higher than that of exome sequencing.[19] However, whole genome sequencing would be more challenging to interpret and is currently difficult to perform in the clinical setting.

Reported diagnostic yields of exome sequencing in NDD or ID have ranged from 16% to 57% depending on the study population.[9–14] In these studies, diagnostic yield and genetic characteristics were different depending on comorbidities, race, and severity of the condition. Study populations were diverse and included probands, parents and probands, or families. Overall, the diagnostic yield of proband-only exome sequencing was slightly lower than that of proband-parents trio studies and reached only about 25% in a recent study.[12]

However, in our study, the diagnostic yield of proband-only exome sequencing increased to 40%. This increase can likely be explained by a number of factors. Owing to our growing understanding of NDD or ID genetics, pathogenicity of several past VOUS could now be readily determined. In the last five years, about 200 additional NDD or ID genes have been identified.[16] In addition, because this study was performed in-house, at one center, in-depth phenotyping and genotyping correlation analysis could be performed. The detection of CNVs using next-generation sequencing technologies also increased our diagnostic yield, as we could identify CNVs using relative depth comparisons.

Currently, CMA is designated as a first-line test for individuals with NDD or ID.[23] An important role for CNVs in this condition is well known. Here, our data suggest that CNVs can be detected by using next-generation sequencing technologies. CNVs were

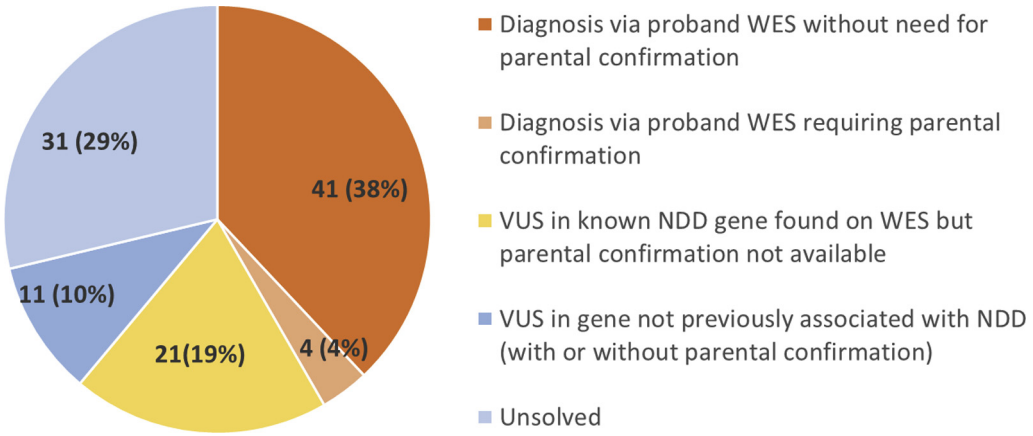

FIGURE 3. Proportion of patients with pathogenic variants and variants of unknown significance. NDD, neurodevelopmental disorder; VUS, variant of unknown significance; WES, whole exome sequencing. The color version of this figure is available in the online edition.

detected in one-tenth of our patients by using proband-only clinical exome sequencing. This result correlates well with the previously known diagnostic yields of CMA.[3–5] Thus we speculate that, in some cases, exome sequencing may provide a substitute for CMA. Exome sequencing may even be considered before CMA in the future if CNVs can be reliably detected using exome sequencing.

The main challenge of using proband-only clinical exome sequencing for NDD of IDs is in the interpretation of *de novo* variants. Some of these variants were missense mutations of disease-causing genes. In these cases, effects of mutations on protein function could not be determined based on current knowledge. Other variants included nonsense mutations in unknown genes. Because some encode proteins that are highly expressed in the brain, there is a good possibility that these genes are pathogenic. However, because of our limited knowledge, the effects of these variants could not presently be determined.

Several studies have reported higher numbers of *de novo* variants in subjects with NDD or ID than in controls, although the significance of individual variants remained undetermined.[28,29] Critically, to determine pathogenicity of new genetic variants, multiple factors must be considered, including recurrence, previous evidence, position of the mutation in the protein, absence of the mutation among healthy individuals, and involvement in disease-implicated protein networks.[29,30] In the clinical setting, however, such extensive evaluations are difficult to perform.

It is well known that male gender is associated with increased risk of NDD or ID. More than 100 genes on the X chromosome can cause this condition.[16,24,25] However, we found that genetic abnormalities on the X chromosome were found rarely in our patients. This finding shows that increased risk of NDD or ID in males cannot be simply explained by genetic abnormalities on the X chromosome. Complex inheritance has been suggested for neurodevelopmental disorders.[26] Diagnostic yields for genetic abnormalities were particularly high in our patients with microcephaly and macrocephaly. Our results correlate well with previous reports.[27] Various CNVs and SNVs are known to cause microcephaly or macrocephaly. We could achieve a high diagnostic yield because we could detect CNVs as well as SNVs with exome sequencing.

At present, the best diagnostic approach to NDDs remain controversial. Although proband-only exome sequencing would cost two-thirds less than trio analysis, it would require longer time to get results and additional work for clinicians. Additional work would include explaining nondiagnostic results to families, tracking down parental samples, and then reinterpreting results with that additional information. Cost-effectiveness of trio analysis, proband-only exome sequencing followed by trio analysis in selected candidates, and even whole genome sequencing would be different depending on the health care system of each center and country. However, we expect that the cost-effectiveness of proband-only clinical exome sequencing would surpass that of trio analysis in the clinical setting in the near future with our rapidly increasing knowledge of neurogenetics. For academic purposes, trio analysis will continue to be pursued.

Because this study was done in a tertiary care referral center, caution is required in interpreting these data. We included a highly selected group of patients with severe NDDs who failed prior evaluations. Also, our monthly, multidisciplinary consensus meeting could have increased the diagnostic yield. The diagnostic yield would be different if data were obtained in a regular developmental pediatrics clinic. A prospective study with a large sample size will be needed to confirm these data.

Despite the study's limitations and challenges, over one-third of our patients with NDD or IDs could be etiologically diagnosed using exome sequencing. Thus this study demonstrates that clinical exome sequencing is a highly efficient and practical diagnostic tool that may be implemented at present in the clinical setting for patients with NDD or IDs. In addition, we predict that the diagnostic yields of this technique will further increase over time with the accumulation of genetic data and improvements to genetic analysis techniques, thereby enhancing its utility.

References

1. Hochstenbach R, van Binsbergen E, Engelen J, et al. Array analysis and karyotyping: workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherlands. Eur J Med Genet. 2009;52:161–169.
2. Schoumans J, Ruivenkamp C, Holmberg E, Kyllerman M, Anderlid BM, Nordenskjöld M. Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). J Med Genet. 2005;42:699–705.
3. de Vries BB, Pfundt R, Leisink M, et al. Diagnostic genome profiling in mental retardation. Am J Hum Genet. 2005;77:606–616.
4. Shaw-Smith C, Redon R, Rickman L, et al. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. J Med Genet. 2004;41:241–248.
5. Vissers LE, de Vries BB, Osoegawa K, et al. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. Am J Hum Genet. 2003;73:1261–1270.
6. Henderson HE, Goodman R, Schram J, Diamond E, Daneel A. Biochemical screening for inherited metabolic disorders in the mentally retarded. S Afr Med J. 1981;60:731–733.
7. Reinecke CJ, Mienie LJ, Hitzeroth HW, Op't Hof J. Screening for inborn errors of metabolism among mentally retarded patients. Outcome of a surgery at the Witrand Care and Rehabilitation Centre. S Afr Med J. 1983;63:14–16.
8. van Karnebeek CD, Jansweijer MC, Leenders AG, Offringa M, Hennekam RC. Diagnostic investigations in individuals with mental retardation: a systematic literature review of their usefulness. Eur J Hum Genet. 2005;13:6–25.
9. Najmabadi H, Hu H, Garshasbi M, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011;478:57–63.
10. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012;367:1921–1929.
11. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet. 2012;380:1674–1682.
12. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA. 2014;312:1870–1879.
13. Anazi S, Maddirevula S, Salpietro V, et al. Expanding the genetic heterogeneity of intellectual disability. Hum Genet. 2017;136:1419–1429.
14. Xiao B, Qiu W, Ji X, et al. Marked yield of re-evaluating phenotype and exome/target sequencing data in 33 individuals with intellectual disabilities. Am J Med Genet A. 2018;176:107–115.
15. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014;511:344–347.
16. Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. Nat Rev Genet. 2016;17:9–18.
17. Schalock R, Borthwick-Duffy S, Bradley V, et al. Intellectual disability: definition, classification, and systems of supports. 11th ed. Washington DC: American Association on Intellectual and Developmental Disabilities; 2010.
18. Li H. Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics. 2014;30:2843–2851.
19. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.
20. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. 2010;26:2069–2070.
21. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29:24–26.
22. Plagnol V, Curtis J, Epstein M, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics. 2012;28:2747–2754.
23. Moeschler JB, Shevell M. Comprehensive evaluation of the child with intellectual disability or global developmental delays. Pediatrics. 2014;134:e903–e918.
24. Raymond FL. X linked mental retardation: a clinical guide. J Med Genet. 2006;43:193–200.
25. Lubs HA, Stevenson RE, Schwartz CE. Fragile X and X-linked intellectual disability: four decades of discovery. Am J Hum Genet. 2012;90:579–590.
26. Jacquemont S, Coe BP, Hersch M, et al. A higher mutational burden in females supports a “female protective model” in neurodevelopmental disorders. Am J Hum Genet. 2014;94:415–425.

27. Fan Y, Wu Y, Wang L, et al. Chromosomal microarray analysis in developmental delay and intellectual disability with comorbid conditions. BMC Med Genomics. 2018;11:49.
28. Kosmicki JA, Samocha KE, Howrigan DP, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. Nat Genet. 2017;49:504–510.
29. Heyne HO, Singh T, Stamberger H, et al. De novo variants in neurodevelopmental disorders with epilepsy. Nat Genet. 2018;50:1048–1053.
30. Hoischen A, Krumm N, Eichler EE. Prioritization of neurodevelopmental disease genes by discovery of new mutations. Nat Neurosci. 2014;17:764–772.