Check for updates

Construction of copy number variation landscape and characterization of associated genes in a Bangladeshi cohort of neurodevelopmental disorders

Hosneara Akter[1,2†], Muhammad Mizanur Rahman[3†],
Shaoli Sarker[4,5], Mohammed Basiruzzaman[4,6],
Md. Mazharul Islam[4,6], Md. Atikur Rahaman[1],
Md. Ashiquir Rahaman[7], Tamannyat Binte Eshaque[1],
Nushrat Jahan Dity[1], Shouvik Sarker[8], Md. Robed Amin[9],
Mohammad Monir Hossain[10], Maksuda Lopa[7], Nargis Jahan[7],
Shafaat Hossain[11], Amirul Islam[1,12], Ashaduzzaman Mondol[7],
Md Omar Faruk[7], Narayan Saha[10], Gopen kumar Kundu[13],
Shayla Imam Kanta[5], Rezaul Karim Kazal[14], Kanij Fatema[3],
Md. Ashrafur Rahman[15], Maruf Hasan[16],
Md. Abid Hossain Mollah[17], Md. Ismail Hosen[2],
Noushad Karuvantevida[18], Ghausia Begum[18], Binte Zehra[18],
Nasna Nassir[18], A. H. M. Nurun Nabi[2], K. M. Furkan Uddin[1,19] and
Mohammed Uddin [iD] [18,20]*

[1]Genetics and Genomic Medicine Centre, NeuroGen Healthcare, Dhaka, Bangladesh, [2]Department of
Biochemistry and Molecular Biology, University of Dhaka, Dhaka, Bangladesh, [3]Department of Paediatric
Neurology, Bangabandhu Sheikh Mujib Medical University, Dhaka, Bangladesh, [4]Department of Child
Neurology, NeuroGen Healthcare, Dhaka, Bangladesh, [5]Department of Paediatric Neuroscience, Dhaka
Shishu Hospital, Dhaka, Bangladesh, [6]Department of Neurology, National Institute of Neurosciences and
Hospital, Dhaka, Bangladesh, [7]Centre for Precision Therapeutics, NeuroGen Healthcare, Dhaka,
Bangladesh, [8]Institute of Plant Genetics, Department of Plant Biotechnology, Leibniz University Hannover,
Hanover, Germany, [9]Department of Medicine, Dhaka Medical College, Dhaka, Bangladesh, [10]Department
of Paediatric Neurology, National Institute of Neuroscience and Hospital, Dhaka, Bangladesh,
[11]Department of Biology and Biochemistry, University of Houston, Houston, TX, United States, [12]Cellular
Intelligence Lab, GenomeArc Inc, Toronto, ON, Canada, [13]Department of Child Neurology, Bangabandhu
Sheikh Mujib Medical University, Dhaka, Bangladesh, [14]Department of Obstetrics and Gynaecology,
Bangabandhu Sheikh Mujib Medical University, Dhaka, Bangladesh, [15]Department of Pharmaceutical
Sciences, Wilkes University, Pennsylvania, PA, United States, [16]Department of Biomedical Engineering,
Military Institute of Science and Technology, Dhaka, Bangladesh, [17]Department of Paediatrics, BIRDEM
General Hospital, Dhaka, Bangladesh, [18]College of Medicine, Mohammed Bin Rashid University of
Medicine and Health Sciences, Dubai, United Arab Emirates, [19]Department of Biochemistry, Holy Family
Red Crescent Medical College, Dhaka, Bangladesh, [20]Cellular Intelligence (Ci) Lab, GenomeArc Inc,
Toronto, ON, Canada

Introduction: Copy number variations (CNVs) play a critical role in the
pathogenesis of neurodevelopmental disorders (NDD) among children. In this
study, we aim to identify clinically relevant CNVs, genes and their phenotypic
characteristics in an ethnically underrepresented homogenous population of
Bangladesh.

Methods: We have conducted chromosomal microarray analysis (CMA) for 212 NDD patients with male to female ratio of 2.2:1.0 to identify rare CNVs. To identify candidate genes within the rare CNVs, gene constraint metrics [i.e., "Critical-Exon Genes (CEGs)"] were applied to the population data. Autism Diagnostic Observation Schedule-Second Edition (ADOS-2) was followed in a subset of 95 NDD patients to assess the severity of autism and all statistical tests were performed using the R package.

Results: Of all the samples assayed, 12.26% (26/212) and 57.08% (121/212) patients carried pathogenic and variant of uncertain significance (VOUS) CNVs, respectively. While 2.83% (6/212) patients' pathogenic CNVs were found to be located in the subtelomeric regions. Further burden test identified females are significant carriers of pathogenic CNVs compared to males (OR = 4.2; p = 0.0007). We have observed an increased number of Loss of heterozygosity (LOH) within cases with 23.85% (26/109) consanguineous parents. Our analyses on imprinting genes show, 36 LOH variants disrupting 69 unique imprinted genes and classified these variants as VOUS. ADOS-2 subset shows severe social communication deficit (p = 0.014) and overall ASD symptoms severity (p = 0.026) among the patients carrying duplication CNV compared to the CNV negative group. Candidate gene analysis identified 153 unique CEGs in pathogenic CNVs and 31 in VOUS. Of the unique genes, 18 genes were found to be in smaller (<1 MB) focal CNVs in our NDD cohort and we identified *PSMC3* gene as a strong candidate gene for Autism Spectrum Disorder (ASD). Moreover, we hypothesized that *KMT2B* gene duplication might be associated with intellectual disability.

Conclusion: Our results show the utility of CMA for precise genetic diagnosis and its integration into the diagnosis, therapy and management of NDD patients.

Introduction

Neurodevelopmental disorders (NDDs) are a group of developmental deficits that disrupt the normal physiology and function of the brain. These disorders are referred to as a collection of early-onset developmental delay (DD) conditions that include autism spectrum disorders (ASDs), intellectual disabilities (IDs), epilepsy encephalopathy, attention-deficit hyperactive disorders (ADHDs), obsessive compulsive disorder (OCD), and cognitive skill disorders (Zoghbi, 2003; Levitt et al., 2004; Dennis et al., 2009; Bale et al., 2010; Krakowiak et al., 2012). When isolated, such disorders are termed non-syndromic, while these are referred to as syndromic when associated with dysmorphisms or apparent congenital anomalies (CAs) (Abou Jamra et al., 2011). The incidence of DD/ID is 3% in the general population (Shevell et al., 2003), while the statistics from the United States shows that ASD affects 1 in 54 live births (Knopf, 2020). Individuals affected with NDDs usually present reduced adaptive skills, limited intellectual ability, motor difficulties, CAs, and problems with social interaction. Phenotypically, there are major overlaps among ASDs with epilepsy encephalopathy, ADHD, Fragile X syndrome (FXS), motor abnormality, and intellectual disability (Matson and Shoemaker, 2009; Hampson et al., 2011).

The etiology of NDDs is principally genetic. Advancements in genomic techniques such as CMA and next-generation sequencing have yielded significant insights into the genetic etiology of NDDs

(Zhang et al., 2013). For decades, structural genomic variation has been a major contributor to the etiology of a proportion of children diagnosed with NDDs (Lionel et al., 2011; Ahn et al., 2014; Liu et al., 2015b; Oskoui et al., 2015; Begum et al., 2021). In the last decade, many large international genomic consortiums have profiled NDD cases, mostly of European ancestry, to identify genomic alterations and NDD-associated genes. More than 100 genes and genomic loci (Cooper et al., 2011) have been consistently found to be involved in the etiology of NDDs. Studies based on ASD cohorts have identified an increased burden of rare genetic copy number variations (CNVs) and have characterized rare, usually *de novo*, recurrent CNV loci that contribute to genetic risk (Pinto et al., 2014). Specific genes within these CNV regions implicated in the etiology of ASD and other NDDs include *SHANK3*, *SYNGAP1*, *NRXN1*, *GRM7*, and *DLGAP2*. (Moessner et al., 2007; Hamdan et al., 2011; Chien et al., 2013; Liu et al., 2015a; Woodbury-Smith et al., 2017). As the number of candidate genes and loci has increased, a striking recurrence of candidates identified in multiple disorders has been uncovered, which may account for a proportion of the significant comorbidity that has been noted among neurodevelopmental disorders (Chen et al., 2014; Cristino et al., 2014). The availability of microarray-related technologies and the contribution of structural variations to NDD enabled CMA as one of the first-tier diagnostic tests for NDD cases in developing countries. In 2010, Miller et al. (2010) demonstrated the utility of

CMA as a first-tier clinical diagnostic test to enable early diagnosis of individuals with NDDs. This technology can precisely detect 10%–30% of NDD cases (Miller et al., 2010; Ho et al., 2016; Uddin et al., 2016; Chaves et al., 2019).

In Bangladesh, child mental health is considered a significant health problem, with around 5 million Bangladeshi children and adolescents having psychiatric disorders (Mullick and Goodman, 2005). In a well-characterized NDD cohort, a gold-standard observational assessment tool ADOS-2 confirmed 73.85% (209/283) as ASD cases and the remaining 26.15% (74/283) as broader NDD cases (Rahaman et al., 2020). Here, the diagnosis of NDDs was mostly carried out by observing the clinical conditions of the patients and using psychological assessment tools (DSM-IV and ADOS/ADOS2) (Rahaman et al., 2020). Due to the overlapping and complex clinical presentation of NDDs, it is difficult to confirm NDD diagnosis by psychological assessment. Therefore, early diagnosis of NDD cases in children may lead to better outcomes through expeditious educational planning and therapeutics (Filipek et al., 1999). In Bangladesh, the genetic cause of breast cancer, intellectual disability, and rare diseases was uncovered by whole-genome sequencing, whole-exome sequencing, targeted sequencing, chromosomal microarray analysis, and quantitative PCR (Uddin K. M. F. et al., 2018; Akter et al., 2019; Rahman et al., 2019; Uddin et al., 2019; Akter et al., 2021). However, there is no comprehensive genetic study carried out with a large NDD cohort. Our study analyzed a cohort of 212 NDD patients in Bangladesh who underwent microarray testing to identify copy number variations (CNVs) from 2017 to 2020 for diagnostic purpose. To our knowledge, this is the first cohort of NDD patients reporting a significant number of clinically relevant variants and genes from the Bangladeshi population.

Materials and methods

Cohort description

The cohort comprised 212 neurodevelopmental disorder patients with ASD ($n = 95$), DD ($n = 48$), epilepsy/seizure ($n = 19$), ID ($n = 15$), hypotonia ($n = 5$), speech and language disabilities ($n = 144$), ADHD (13), and rare unexplained cases with syndromic features ($n = 25$). Almost all the patients have more than one phenotype (Supplementary Table S1). Of the 212 patients, 68.40% ($n = 145$) and 31.60% ($n = 67$) were men and women, respectively. Their detailed phenotype is provided in Supplementary Table S1. We collected occipitofrontal head circumference (OFC) data, measured in centimeter using a non-stretchable plastic tape, and body weight data, measured in kilogram using a calibrated weight machine. From the available data, we have shown the distribution of age, OFC, weight, and height of 202 patients (Supplementary Figures S1, S2). All patients are from the local Bangladeshi population. Around 43% patients are city-based, so they have access to resources of modern healthcare. The rest of the patients are from rural areas that lack modern healthcare resources. All these patients were referred for microarray testing after clinical evaluation by different neurologists and pediatricians from different parts of the country. From the available medical records of some patients, we have prepared a summary table keeping information

such as mode of delivery, perinatal complication, percentage of premature and intrauterine dystrophic newborns, percentage of specific instrumental examinations (EEG, etc.), and ultrasonography and MRI of the CNS (Supplementary Table S2). From the available medical records, we found the birth weight of 76 patients and produced a birth weight distribution plot (Supplementary Figure S3). Moreover, a subset of patients ($n = 95$) was evaluated for autism spectrum disorders by using the ADOS-2 method. Among these 95 participants (70 men and 25 women), 71 met the criteria of autism-positive (51 men and 20 women), and the remaining 24 were classified as autism-negative. In addition to the 71 patients , 24 more autism patients were diagnosed by using different assessment tools, including the DSM-V ($n = 14$) and ADOS ($n = 10$). Therefore, in total, there were 95 ASD patients (44.81% = 95/212) diagnosed using different psychological assessment tools, DSM-V ($n = 14$), ADOS ($n = 10$), and ADOS-2 ($n = 71$), with a male-to-female ratio of 2.65:1.00 and age ranging from 1.5 years to 19 years (Supplementary Table S1).

Autism assessment

Autism Diagnostic Observation Schedule-Second Edition (ADOS-2) (Lord et al., 2012) is a semi-structured, standardized assessment of communication, social interaction, imaginative use of materials, and restricted and repetitive behaviors. ADOS-2 is a gold-standard observational assessment tool for diagnosing ASD. ADOS-2 has five modules, and each module offers standard activities designed to elicit behaviors relevant to diagnosing ASD at different chronological ages and language abilities. In this study, autism characteristics were measured using the Toddler module, Module 1, and Module 2. Each module uses communication, reciprocal social interaction, and restricted and repetitive behaviors to generate a total score. Elevated scores classify an individual in the autism spectrum or autism diagnostic range based on the severity or the frequency displayed.

DNA extraction and quantitation

DNA was extracted from peripheral blood samples using the ReliaPrep™ Blood gDNA isolation kit (Promega, United States) following the manufacturer's protocol. Extracted DNA samples were checked for quality using a NanoDrop spectrophotometer. Samples were electrophoresed on agarose gels, and samples with intact genomic DNA showing no smearing on agarose gel were selected for the experiment. Intact genomic DNA was diluted to 50 ng/μL concentration based on Quant-iTPicoGreen (Invitrogen) quantitation. The whole-genome amplification process requires 200 ng of input gDNA.

Chromosomal microarray analysis

We conducted CMA to identify copy number variations and investigated the changes in fluorescence intensities between the test specimen and the controls. Illumina Global Screening Array-24 + v1.0 was used applying Illumina SNP genotyping technology and Illumina CNVPartition 3.2.1 plug-in of GenomeStudio to detect

chromosomal abnormalities using the reference genome GRCh38/hg38. This microarray uses 642,824 probes spread across the genome to detect genetic abnormalities (includes >60 loci in the DECIPHER database reported for neurodevelopmental disorders) greater than 30 kb (for deletion and duplication CNV) and targets sub-telomeric regions vulnerable to chromosomal abnormalities. We analyzed loss of heterozygosity (LOH) within a 2–10 MB range for further analysis due to low-resolution array and false-positive calls. We used rigorous multiple algorithmic techniques (MatLab and Java) and manual curation of the data to pinpoint genomic variation based on the normalized log2 intensities of the probes. Our algorithm excluded all common CNVs found in the in-house (NeuroGen) control population samples (9,689 samples) from analysis and only rare (frequency <0.01) CNVs to infer their contribution to human diseases. Sample digestion, ligation, PCR, labeling, hybridization, and scanning were performed following the standard protocols.

Frequency determination

We used a commercially available tool, GenomeArc (GenomeArc Inc.), an annotation software that integrates clinical, genomics, and OMICs data. To determine the frequency, >50% reciprocal overlap within the same chromosome was considered, and we used 9,689 unrelated population control samples from European ancestry (Uddin et al., 2016). These samples were collected from multiple major population-scale studies that used high-resolution microarray platforms. These included 4,347 control samples assayed by Illumina 1 M from the Study of Addiction Genetics and Environment (SAGE) (Bierut et al., 2010) and the Health, Aging, and Body Composition (HABC) (Coviello et al., 2012); 2,988 control samples were assayed by Illumina Omni 2.5 M from the Collaborative Genetic Study of Nicotine Dependence (COGEND) (Bierut et al., 2007) and Cooperative Health Research in the Region of Augsburg KORA projects (Verhoeven et al., 2013); and 2,357 control samples were assayed by Affymetrix 6.0 from the Ottawa Heart Institute (Stewart et al., 2009) and the PopGen project (Krawczak et al., 2006).

CNV classification criteria

To classify and assess the clinical relevance of a particular CNV, at first, all common CNVs (frequency >1%) found in the control population samples (9,689 samples) and Database of Genomic Variants (DGV) (http://dgv.tcag.ca/dgv/app/home) were excluded from further analysis. Then, rare CNVs (frequency<1%) were classified based on their type (gain or loss or LOH), size, location, gene content, and patients' clinical data. By searching genomic databases, including Online Mendelian Inheritance in Man (OMIM) (https://omim.org/), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), ClinGen Dosage Sensitivity Map (https://dosage.clinicalgenome.org), ISCA database (http://dbsearch.clinicalgenome.org/search/), DECIPHER (https://decipher.sanger.ac.uk/), and published literature (http://www.ncbi.nlm.nih.gov/pubmed/), CNVs were finally classified as pathogenic, variants of uncertain significance (VOUS), and benign following the ACMG guidelines (Kearney et al., 2011).
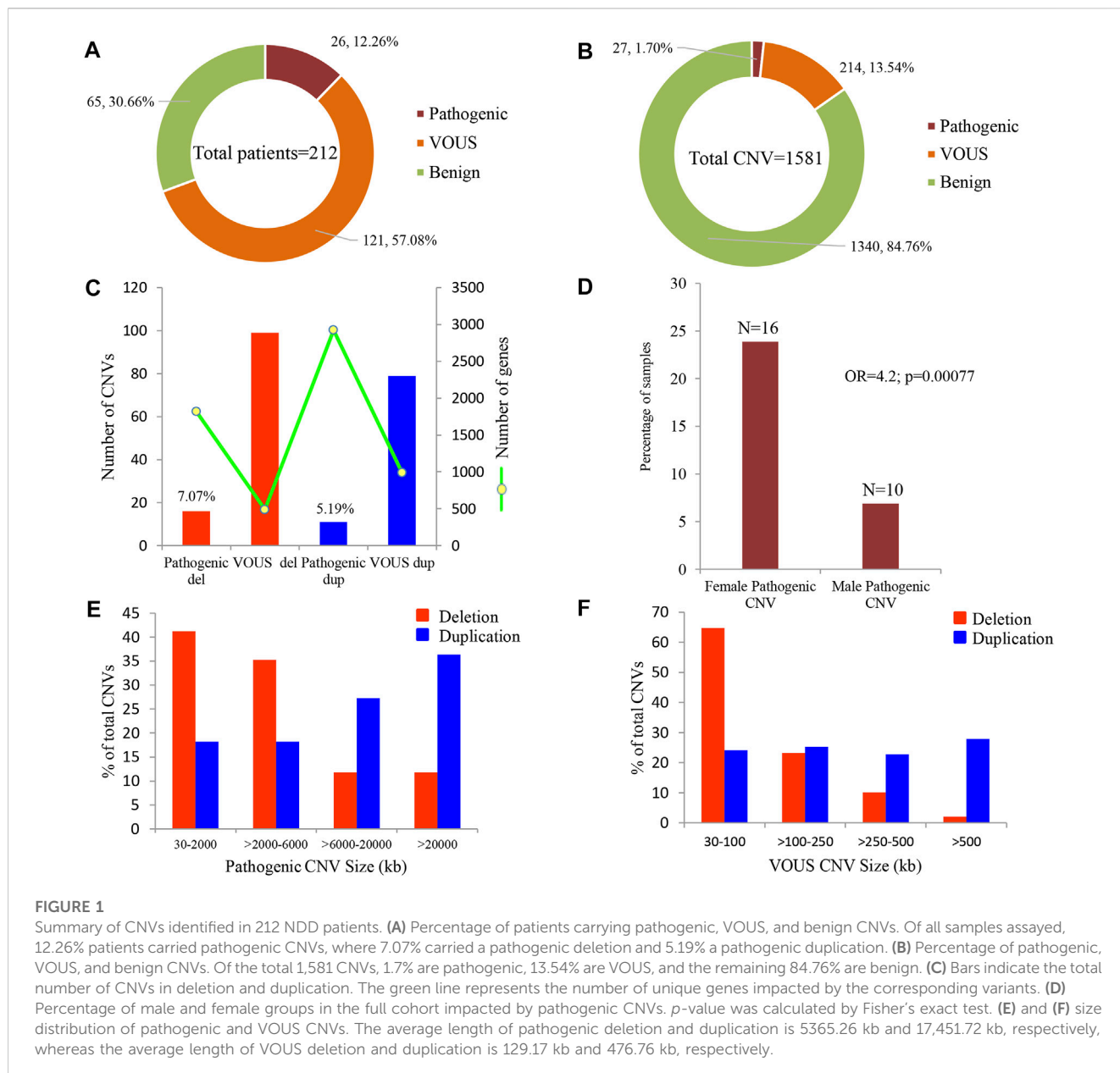
Droplet digital PCR

Copy number assays were performed using the Droplet Digital PCR (ddPCR) System (Bio-Rad Laboratories, Inc.). The GeneAssist™ Copy Number Assay Workflow Builder (Thermo Fischer, United States) was used to design TaqMan assays on Chr11, *PKNOX2* (Hs03289418_cn), Chr15, *SNHG14* (Hs05375107_cn), and Chr21, *TSPEAR* (Hs02835328_cn) with FAM dye. TaqMan™ Copy Number Reference Assay, human, RNase P with VIC dye was used as a reference assay. A total of 22 μL reaction mix was prepared, containing 3.5 μL of template DNA (20 ng/μL) without restriction digestion, 10 μL of 2X ddPCR supermix for probes (No UDP) (Bio-RAD Laboratories Inc.), 1 μL each 20X TaqMan target probe (FAM) and 20X TaqMan reference probe (VIC) (Applied Biosystems, United States), and 6.5 μL of RNase-/DNase-free water. All reactions were prepared in triplicates with one negative control. The reaction mixtures were partitioned using the QX200 Droplet Generator™ and then transferred to a 96-well plate and amplified using the C1000 Touch thermal cycler, as per the manufacturer's protocol. Then, the samples were read using the QX200 Droplet Reader™. Data acquisition and analysis were performed using QuantaSoft Version 1.7.4.0917, and the Poisson algorithm was used to determine the concentrations of the targets as copies/μL.

Candidate gene and pathway analysis

We used pathogenic CNVs for enrichment analysis (https://github.com/MBRULab/2021_Nassir-etal/blob/main/geneoverlap.R) and mapping. We scanned the KEGG pathway database, which comprises an assembly of the up-to-date interactions, reactions, and relations of molecular networks (https://www.genome.jp/kegg/pathway.html), and the GO database (http://geneontology.org/) to identify all the pathways in which five or more genes were expressed. Only the pathways having more than 50 genes and less than 1,000 genes were considered for this analysis (Cline et al., 2007; Bader et al., 2014; Nassir et al., 2021). The pathways were identified by their unique KEGG ID and name. A pathway was considered significant if the false discovery rate (FDR) and *p*-value cut-off were <0.01 and 0.001, respectively. Then, the network was built using the enrichment map and the auto-annotate Cytoscape (https://cytoscape.org/) application. *P*-values are denoted using a color gradient (low *p*-values with darker colors). Furthermore, the "critical-exon" method (Uddin et al., 2014; Uddin et al., 2016) was applied to identify candidate genes within the pathogenic and VOUS breakpoints. "Critical exons" are highly expressed in the brain and have low population rare mutation burden. Genes with critical exons are called critical exon genes.

Statistical analysis

Welch's *t*-test was used to determine the significant difference between the means of the two groups. A significance level of 0.05 was determined to test differences. For qualitative data, Fisher's exact test was applied. All analyses were performed using the R package.

FIGURE 1
Summary of CNVs identified in 212 NDD patients. **(A)** Percentage of patients carrying pathogenic, VOUS, and benign CNVs. Of all samples assayed, 12.26% patients carried pathogenic CNVs, where 7.07% carried a pathogenic deletion and 5.19% a pathogenic duplication. **(B)** Percentage of pathogenic, VOUS, and benign CNVs. Of the total 1,581 CNVs, 1.7% are pathogenic, 13.54% are VOUS, and the remaining 84.76% are benign. **(C)** Bars indicate the total number of CNVs in deletion and duplication. The green line represents the number of unique genes impacted by the corresponding variants. **(D)** Percentage of male and female groups in the full cohort impacted by pathogenic CNVs. *p*-value was calculated by Fisher's exact test. **(E)** and **(F)** size distribution of pathogenic and VOUS CNVs. The average length of pathogenic deletion and duplication is 5365.26 kb and 17,451.72 kb, respectively, whereas the average length of VOUS deletion and duplication is 129.17 kb and 476.76 kb, respectively.

Results

The cohort comprised 212 neurodevelopmental disorder patients with a male-to-female ratio of 2.2:1.0 (Supplementary Table S1). In this cohort, 95 patients were evaluated for autism spectrum disorder using the ADOS-2 method, and 71 met the ADOS-2 cut-off criteria of autism-positive (51 men and 20 women, male/female ratio 2.55:1.00), while the remaining 24 were classified as autism-negative (non-spectrum). We observed that patients carrying duplication CNVs showed severe social communication deficit ($p = 0.014$) and overall ASD symptom severity ($p = 0.026$) compared to the CNV-negative group (patients with no clinically rare variants). Moreover, a trend of increased number of CNVs in autism patients was observed in comparison to non-spectrum individuals (OR = 2.29; $p = 0.06$).

Of all samples assayed, 12.26% (26/212) and 57.08% (121/212) patients carried pathogenic and variant of uncertain significance (VOUS) CNVs, respectively (Figure 1A). We identified 1,581 CNVs (excluding <30 kb deletion,<50 kb duplication, <2 Mb, and >10 Mb LOH) that included 395 duplications, 658 deletions, and 528 LOH interpreted and classified according to the ACMG guidelines (Kearney et al., 2011) into pathogenic ($n = 27$), variant of uncertain clinical significance (VOUS) ($n = 214$) and benign ($n = 1,340$) CNVs (Figure 1B). Of the 27 pathogenic CNVs, recurrent deletions and duplications were identified in the 15q11.2q13.1 and 21q11.1q22.3 loci (Table 1). In this cohort, 2.83% (6/212) and 20.28% (43/212) patients carried pathogenic and VOUS subtelomeric CNVs, respectively (Supplementary Table S1). Of the 26 samples carrying pathogenic CNVs, one patient was carrying double terminal pathogenic deletion impacting chromosome 18 (Table 1), and 11 patients with a pathogenic

TABLE 1 Details of pathogenic CNVs found in the cohort.

ID	Sex	Type	Chromosomal band and position (hg38)	Size (Mb)	Number of affected genes	Critical exon genes	Known syndrome/genes associated with developmental disorder within the breakpoints	Clinical conditions	BD frequency (n = 212)	Canada control frequency (n = 9,692)
1	M	Del[a]	4p16.3p16.2 chr4: 35,031-4,878,160	4.843	155	CTBP1, PCGF3, FAM193A, and HTT	Wolf–Hirschhorn Syndrome; PIGG, UVSSA3, IDUA, FGFR3, LRPAP1, SH3BP2, and MSX1	Autism, developmental delay (walking at 5 years), behavioral problem, intellectual disability, frequent falling, and dysmorphic face	0.00472	0
19	M	Del[b]	15q11.2q13.1 chr15: 22,838,642-28,314,382	5.476	170	GABRB3, CYFIP1, UBE3A, HERC2, and GABRA5	PWS/AS; UBE3A, GABRB3, and MAGEL2	Global developmental delay, cannot stand spontaneously, fatty body, and syndromic child	0.01415	0
20	F	Del	20p12.3 chr20: 8,117,650-8,593,664	0.476	1	PLCB1	PLCB1	Autism, speech delay, temper tantrum, and restlessness	0.00472	1.03E-04
35	F	Del[b]	15q11.2q13.1 chr15: 22,777,709-28,736,935	5.959	181	GABRB3, CYFIP1, UBE3A, HERC2, and GABRA5	PWS/AS; UBE3A; GABRB3; and MAGEL2	Hypotonic	0.01415	0
40	F	Del	17p11.2 chr17: 16,858,825-20,387,706	3.529	136	LLGL1, SREBF1, COPS3, and AKAP10	Smith–Magenis syndrome; Potocki–Lupski_syndrome_(17p11.2_duplication_syndrome); TNFRSF13B, RAI1, B9D1, and ALDH3A2	Behavioral problem, speech delay, and low hairline	0.00472	0
53	F	Del	1q24.2q31.2 chr1: 168,503,684-191,056,769	22.553	360	TPR, NMNAT2, DHX9, RGL1, SMG7, PRRC2C, GLUL, CACNA1E, IVNS1ABP, RNF2, ASTN1, CACYBP, ATP1B1, VAMP4, TNR, and XPR1	GORAB, TSEN15, LHX4, ACBD6, GLUL, PRRX1, NPHS2, DARS2, and MYOC	Global developmental delay, high ammonia and lactate levels, and EEG-normal	0.00472	0
59	M	Del	17q12 chr17: 36,201,251-37,889,808	1.689	51	GGNBP2 and ACACA	RCAD (renal_cysts_and_diabetes); HNF1B, PIGW	Global developmental delay (delayed walking and speech), salivation, restlessness, and H/O pneumonia at 15 days of age	0.00472	0
86	M	Del[a]	18p11.32 chr18:45,680-2,276,243	2.231	38	COLEC12	Null	Developmental delay and complete cleft lip with cleft palate	0.00472	0
			18q21.31q23 chr18: 58,298,202-80,227,529	21.929	280	PHLPP1, KIAA1468, NEDD4L, ZNF532, NETO1, and TSHZ1	CTDP1, RTTN, NEDD4L, TSHZ1, PIGN, TXNL4A, RAX, CCBE1, and CTDP1		0.00472	0
87	F	Del		0.102	RIMS2	Null	RIMS2	Autism, impaired language development,	0.00472	0

(Continued on following page)

TABLE 1 (*Continued*) Details of pathogenic CNVs found in the cohort.

ID	Sex	Type	Chromosomal band and position (hg38)	Size (Mb)	Number of affected genes	Critical exon genes	Known syndrome/genes associated with developmental disorder within the breakpoints	Clinical conditions	BD frequency (*n* = 212)	Canada control frequency (*n* = 9,692)
			8q22.3 chr8: 103,458,677-103,560,657					restlessness and hyperactivity, and delayed walking		
107	F	Del	7q11.23 chr7: 73,294,506-74,715,504	1.421	40	*BAZ1B, LIMK1, STX1A, EIF4H,* and *CLIP2*	Williams–Beuren syndrome (WBS); 7q11.23 duplication syndrome; *ELN*	Developmental delay (unable to sit without support), feeding problem, jerky movement, and hypotonic	0.00472	0
144	F	Del[a]	2q37.1q37.3 chr2: 233,505,099-242,065,217	8.560	219	*KIF1A, ATG4B, AGAP1, UBE2F,* and *HDLBP*	2q37 monosomy; *KIF1A, COL6A3, NDUFA10, D2HGDH, CAPN10, TWIST2, HDAC4, UGT1A1,* and *AGXT*	Syndromic child, developmental delay, unable to sit without support, poor neck control, no speech, sleeping problem at night, and lack of social response. Eyes closed for 2 days after birth, breathing with oxygen support and needed blood transfusion. Pneumonia occurred on day 28	0.00472	0
159	M	Del	2q24.1 chr2: 154,077,277-158,596,423	4.519	70	*KCNJ3, NR4A2,* and *ACVR1*	*ACVR1*	Intellectual disability, ADHD, speech delay, involuntary movement, abnormal behavior, self-injurious behavior, and aggressive behavior. History of epilepsy two times. At 6 months of age, patient suffered from pneumonia. Facial appearance: normal	0.00472	0
178	M	Del	15q11.2q13.1 chr15: 22,600,363-28,760,485	6.160	188	*GABRB3, CYFIP1, UBE3A, HERC2,* and *GABRA5*	PWS/AS; *UBE3A, GABRB3,* and *MAGEL2*	Spastic quadriplegia/CP, global developmental delay, speech problem, and cannot sit without support. CT scan of the brain: atrophy ventricular dilation. Dysmorphic face	0.01415	0

TABLE 1 (*Continued*) Details of pathogenic CNVs found in the cohort.

ID	Sex	Type	Chromosomal band and position (hg38)	Size (Mb)	Number of affected genes	Critical exon genes	Known syndrome/genes associated with developmental disorder within the breakpoints	Clinical conditions	BD frequency (n = 212)	Canada control frequency (n = 9,692)
182	F	Del	2q24.3 chr2: 165,992,388-166,079,596	0.087	1	Null	*SCN1A*	Global developmental delay (poor neck control, cannot sit or walk, and speech delay), cerebral palsy, epilepsy starts at the age of 5 months, history of frequent infantile spasm, flexor spasm, sleeping problem, eating problem (only liquid food), restlessness, crying nonstop, and drooling of saliva. Her elder brother deceased at the age of 8 months after suffering from pneumonia. He also had delayed development; his neck was also weak, anemic during pregnancy. Facial appearance: abnormal	0.00472	0
184	M	Del	17q12 chr17: 34,815,552-36,249,430	1.434	76	*AP2B1*	*PEX12*	Low-level autism, epilepsy, ADHD, club foot, temper tantrum, speech delay, self-smiling, aggressive, poor response, and poor peer relationship. Admitted to the NICU at the age of 4 days due to convulsion for 2 days, and re-hospitalized at 15 days of age due to cold attack and sleeping problem. Family history: paternal cousin was hyperactive and father had seizure during childhood	0.00472	0
5	F	Dup[a]	3p26.3p26.2 chr3: 2,299,060-3,322,758	1.024	5	null	*CRBN*	Expressive language disorder	0.00472	0
77	F	Dup	9p24.3p12 chr9:48,828-39,297,860	39.249	619	*PTPRD, MLLT3, SH3GL2, CDC37L1, SMARCA2, NOL6, KLHL9, NFIB, PSIP1, SMU1, ELAVL2, UBAP1, RUSC2, UBE2R2, CLTA, CNTFR, VCP,*	*PIGO, IL11RA, NPR2, TPM2, GBA2, GALT, FANCG, EXOSC3, RMRP, MPDZ, GLIS3, DDX58, VLDLR, KANK1, APTX, FREM1, DOCK8, TEK; SMARCA2, GLIS3, PLAA, TYRP1,* and *GLDC*	Developmental delay, speech delay, intellectual disability, and anorexia	0.00472	0

TABLE 1 (*Continued*) Details of pathogenic CNVs found in the cohort.

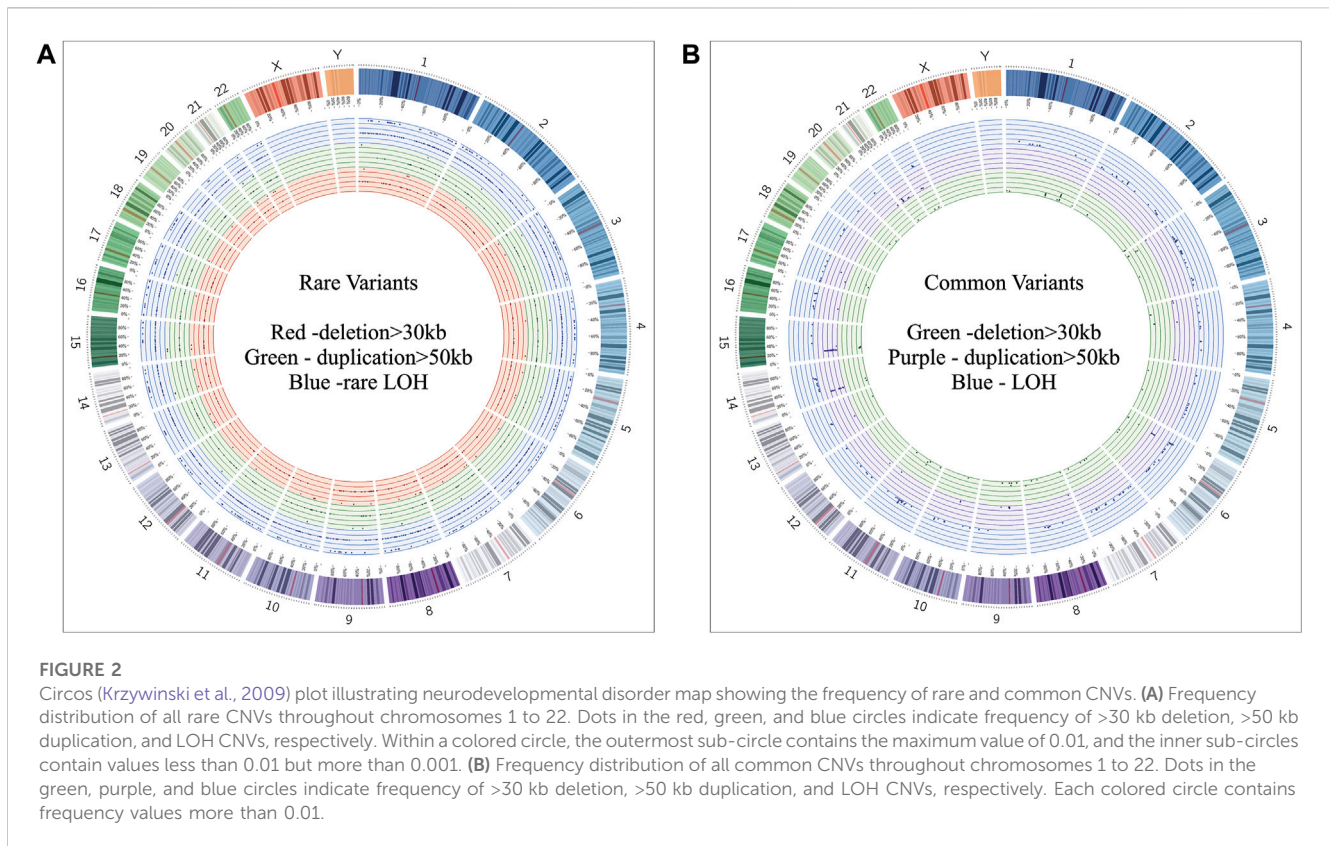| ID	Sex	Type	Chromosomal band and position (hg38)	Size (Mb)	Number of affected genes	Critical exon genes	Known syndrome/genes associated with developmental disorder within the breakpoints	Clinical conditions	BD frequency (n = 212)	Canada control frequency (n = 9,692)
						NPR2, TESK1, TLN1, and RNF38				
14	M	Dup[b]	11q23.2q24.2 chr11: 114,071,597-125,513,912	11.442	307	BCL9L, BACE1, IFT46, DDX6, C2CD2L, GRAMD1B, UBE4A, CADM1, HSPA8, HYOU1, ARHGEF12, SIK3, ARCN1, PKNOX2, and PAFAH1B2	ZBTB16, NECTIN1, ARCN1, ROBO3, DPAGT1, SC5D, CLMP, CBL, KMT2A, and MFRP	Developmental delay (speech delay), poor eye contact, history of seizure, cerebral palsy, facial dysmorphism present, and corpus callosum agenesis	0.00472	0
15	F	Dup[a]	14q32.11q32.33 chr14: 90,576,836-106,816,816	16.240	626	DICER1, ITPK1, TTC7B, TRAF3, PAPOLA, TECPR2, AKT1, DYNC1H1, EVL, WARS, YY1, CDC42BPB, BCL11B, PPP2R5C, INF2, EIF5, PACS2, and MTA1	TRIP11, UBR7, SLC24A4, CCDC88C, APOPT1, TECPR2, VRK1, DYNC1H1, AKT1, YY1, and PACS2	High level of salivation, lack of self-care, and developmental delay	0.00472	0
2	M	Dup[b]	15q11.2q13.1 chr15: 23,443,797-28,289,373	4.846	147	GABRB3, UBE3A, HERC2, and GABRA5	PWS/AS; UBE3A, GABRB3, and MAGEL2	Autism, syndromic child	0.009434	0
101	F	Dup[b]	15q11.2q13.1 chr15: 23,370,969-28,371,148	5.000	157	GABRB3, UBE3A, HERC2, and GABRA5	PWS/AS; UBE3A, GABRB3, and MAGEL2	Autism, impaired language development, abnormal behavior, stereotypical hand movement, restrictive and repetitive activity, restlessness and hyperactivity, and unable to play with peers. High lactic acid and ammonia levels. Facial appearance: normal	0.009434	0
79	M	Dup[a]	17p13.3p12 chr17: 151,597-11,392,231	11.241	414	PRPF8, PITPNA, ABR, YWHAE, MNT, SMG6, TNFSF12, SENP3, RAP1GAP2, CHD3, CAMTA2, KDM6B, DLG4, NLGN2, ANKFY1, PHF23, FXR2, POLR2A, MINK1, ZBTB4, CTDNEP1, RABEP1, TNFSF12-TNFSF13, EIF4A1, PAFAH1B1,PITPNM3, NEURL4, RPL26, MYH10, and NDEL1	Miller–Dieker syndrome (MDS); INPP5K, KANSL1, WDR81, BHLHA9, WRAP53, DLG4, TRPV3, ACADVL, ASPA, C1QBP, MPDU1, KDM6B, SLC13A5, CTNS, AIPL1, AIPL1, MYH10, CTC1, MYH3, MYH8, and SCO1	Speech delay, hyperactivity, restlessness, behavioral problem, history of seizure, seizure-free for >5 years, sleeping problem, and presence of facial dysmorphism	0.00472	0

TABLE 1 (*Continued*) Details of pathogenic CNVs found in the cohort.

ID	Sex	Type	Chromosomal band and position (hg38)	Size (Mb)	Number of affected genes	Critical exon genes	Known syndrome/genes associated with developmental disorder within the breakpoints	Clinical conditions	BD frequency (*n* = 212)	Canada control frequency (*n* = 9,692)
134	F	Dup[b]	21q11.1q22.3 chr21: 12,987,574-46,679,698	33.692	683	*NRIP1, ETS2, COL6A1, ITSN1, PDXK, CCT8, SON, AGPAT3, NCAM2, BRWD1, DYRK1A*, and *TIAM1*	Down's syndrome; *C21orf59, SON, HLCS, KCNE1, DYRK1A, C21orf2, COL18A1, PCNT, RSPH1, RIPK4, CBS, FTCD, SIK1, COL18A1, CRYAA, AIRE*, and *CSTB*	Global developmental delay, feeding problem, no toilet training, speech delay, vomiting tendency, hyperactivity, and facial dysmorphism present (flat nose, Down's syndrome -like eyes and body)	0.01415	0
136	F	Dup[b]	21q11.1q22.3 chr21: 12,987,574-46,679,698	33.692	683	*NRIP1, ETS2, COL6A1, ITSN1, PDXK, CCT8, SON, AGPAT3, NCAM2, BRWD1, DYRK1A*, and *TIAM1*	Down's syndrome; *C21orf59, SON, HLCS, KCNE1, DYRK1A, C21orf2, COL18A1, PCNT, RSPH1, RIPK4, CBS, FTCD, SIK1, COL18A1, CRYAA, AIRE*, and *CSTB*	Developmental delay, dysmorphic facial appearance, nystagmus, speech delay, low eye contact, and eating problem. Family history: low IQ in the mother	0.01415	0
152	F	Dup[b]	21q11.1q22.3 chr21: 12,987,574-46,679,698	33.692	683	*NRIP1, ETS2, COL6A1, ITSN1, PDXK, CCT8, SON, AGPAT3, NCAM2, BRWD1, DYRK1A*, and *TIAM1*	Down's syndrome; *C21orf59, SON, HLCS, KCNE1, DYRK1A, C21orf2, COL18A1, PCNT, RSPH1, RIPK4, CBS, FTCD, SIK1, COL18A1, CRYAA, AIRE*, and *CSTB*	Developmental delay and ventricular septal defect (VSD) surgery. He had CHD in the early neonatal period and was diagnosed as having large aortic VSD with inlet extension, large PDA, mild TR, and severe pulmonary hypertension. He had a history of fast breathing, feeding difficulty with failure to thrive, excessive sweating, and repeated LRTIs since the neonatal period	0.01415	0
6	F	Dup	22q11.22q11.23 chr22: 22,749,561-24,600,663	1.851	114	*BCR, SMARCB1*, and *SPECC1L*	22q11.2 distal deletion syndrome; *SPECC1L* and *SMARCB1*	Restlessness, low memory, and a tendency of tearing body clothes. CT scan: normal	0.00472	1.03E-04

[a]Subtelomeric CNV; CNV, copy number variation; CP, cerebral palsy; BD, Bangladesh; del, deletion; dup, duplication; F, female; M, male.
[b]Each CNV validated by droplet digital PCR (ddPCR) with TaqMan assay.

FIGURE 2
Circos (Krzywinski et al., 2009) plot illustrating neurodevelopmental disorder map showing the frequency of rare and common CNVs. **(A)** Frequency distribution of all rare CNVs throughout chromosomes 1 to 22. Dots in the red, green, and blue circles indicate frequency of >30 kb deletion, >50 kb duplication, and LOH CNVs, respectively. Within a colored circle, the outermost sub-circle contains the maximum value of 0.01, and the inner sub-circles contain values less than 0.01 but more than 0.001. **(B)** Frequency distribution of all common CNVs throughout chromosomes 1 to 22. Dots in the green, purple, and blue circles indicate frequency of >30 kb deletion, >50 kb duplication, and LOH CNVs, respectively. Each colored circle contains frequency values more than 0.01.

CNV were also carrying VOUS CNV (Supplementary Table S3). The 27 pathogenic CNVs comprised 16 deletions and 11 duplications (Table 1; Figure 1C). It was observed that women are the more substantial carriers of pathogenic CNVs than men (OR = 4.2; p = 0.0007) (Figure 1D). The average length of deletion and duplication was 5365.26 kb and 17,451.72 kb, respectively, and the highest frequency group for pathogenic CNV deletion and duplication was 30–2000 kb and >20000 kb, respectively (Figure 1E). To exclude false CNV calls, we randomly chose nine pathogenic/ likely pathogenic CNVs for ddPCR validation, and it yielded an (8/9) 88.89% validation rate (Supplementary Figure S4, Supplementary Table S4). In this validation, a single 775.77 kb deletion variant (#case 7) did not show accurate log intensity data due to the algorithm's false call on the low resolution of the microarray. A large number of genes were found to be disrupted in pathogenic deletion (n = 1846) and duplication CNVs (n = 2,925) compared to VOUS deletions (n = 490) and duplications (n = 991) (Figure 1C). The average length of VOUS deletion and duplication was 129.17kb and 476.76 kb, respectively, and the highest frequency group for VOUS CNV deletion and duplication was 30–100 kb and >500 kb, respectively (Figure 1F). Furthermore, we determined the frequency of all 1,581 deletion, duplication, and LOH CNVs within 212 Bangladeshi NDD patients and 9,689 control CNVs from the Ontario population, which is predominantly European in ancestry. Using the frequency distribution of both rare and common CNVs of NDD patients from Bangladesh, a Circos map was constructed (Krzywinski et al., 2009) (Figure 2). Consanguinity increases the possibility of inheriting recessive diseases, and in our Bangladeshi cohort, 23.85% (26/109) families were found to be

consanguineous. In the children born to consanguineous parents, we found an increased number (12.95 LOH/patient) and larger size (>6 Mb) of LOH compared to that found in the children of non-consanguineous parents (Supplementary Figures S5, S6). The pathogenesis of LOH includes homozygous mutation of recessive diseases and imprinting effects caused by uniparental disomy (UPD). However, without whole-genome sequencing data of the genes associated with recessive diseases, it is difficult to address the severity of LOH-harboring recessive genes. We have thoroughly checked the LOH for imprinted genes and found 36 LOH-disrupting 69 unique imprinted genes and classified these variants as VOUS (Supplementary Table S5).

Our custom pathway enrichment (comprising Gene Ontology and KEGG databases) analysis of the impacted genes within the CNV breakpoints of all pathogenic deletions identified "ubiquitin-like protein transferase activity (GO:0019787)," "negative regulation of cell death (GO:0060548)," and "vesicle-mediated transport in synapse (GO:0099003)" pathways to be highly significant (FDR p < 6.53×10^{-7}, FDR p < 7.2×10^{-6}, and FDR p < 7.1×10^{-5}) after correction for multiple tests (Figure 3A). Pathway enrichment on the genes impacted by pathogenic duplications identified "interspecies interaction between organisms (GO:0044419)," "dependent protein catabolic process (GO:0030163)," "regulation of neuronal death (GO:1901214)," and "synaptic signaling (GO: 0099536)" pathways to be highly significant (FDR p < 1.51×10^{-5}, p < 1.85×10^{-5}, p < 5.34×10^{-7}, and p < 1.72×10^{-4}) after correction for multiple tests (Figure 3B).

Analysis of both pathogenic and VOUS CNVs for critical exon genes (CEGs) yielded 153 unique CEGs in pathogenic CNVs and
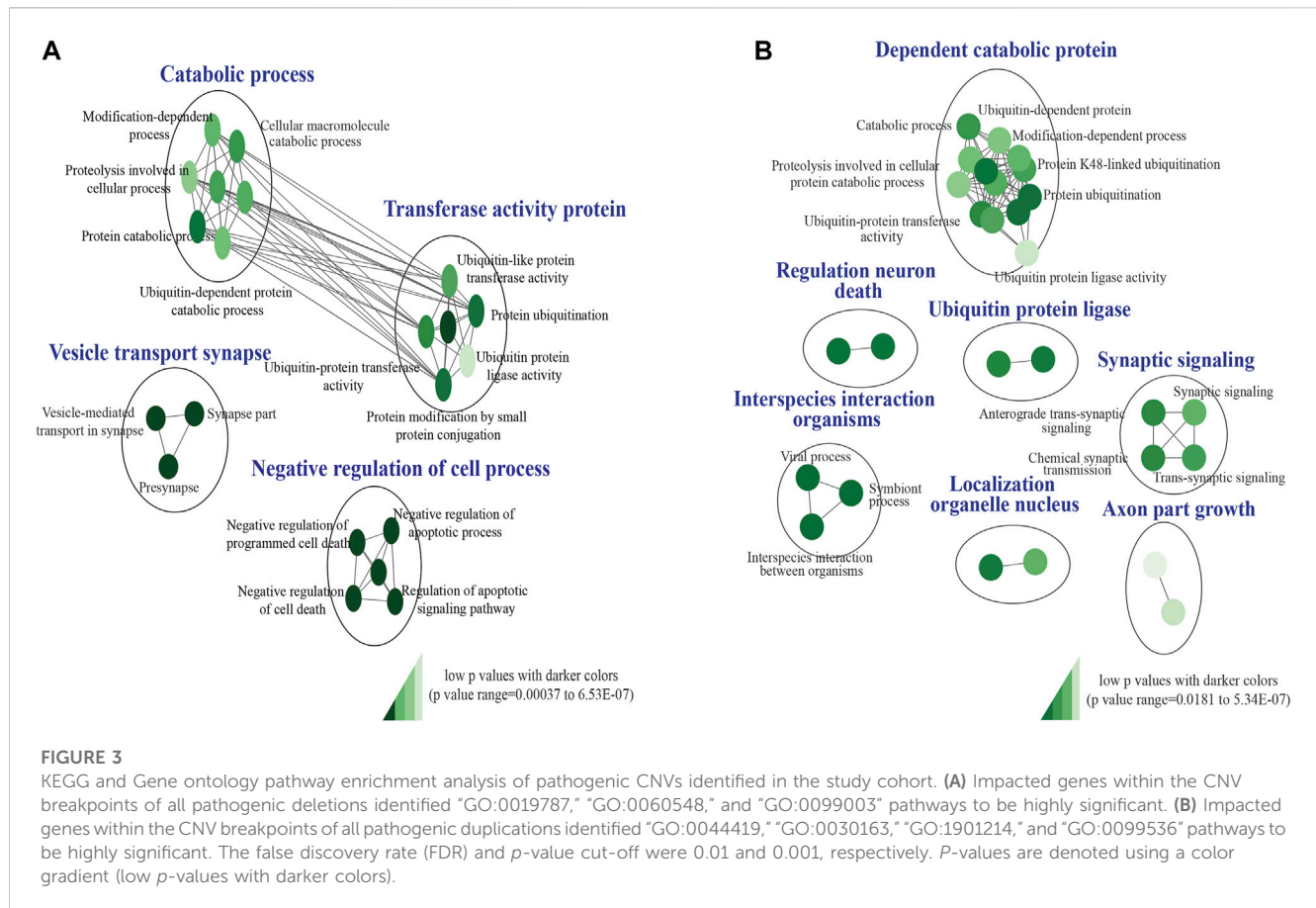
FIGURE 3
KEGG and Gene ontology pathway enrichment analysis of pathogenic CNVs identified in the study cohort. **(A)** Impacted genes within the CNV breakpoints of all pathogenic deletions identified "GO:0019787," "GO:0060548," and "GO:0099003" pathways to be highly significant. **(B)** Impacted genes within the CNV breakpoints of all pathogenic duplications identified "GO:0044419," "GO:0030163," "GO:1901214," and "GO:0099536" pathways to be highly significant. The false discovery rate (FDR) and *p*-value cut-off were 0.01 and 0.001, respectively. *P*-values are denoted using a color gradient (low *p*-values with darker colors).

TABLE 2 Critical exon genes (CEGs) identified in CNVs from neurodevelopmental disorder.

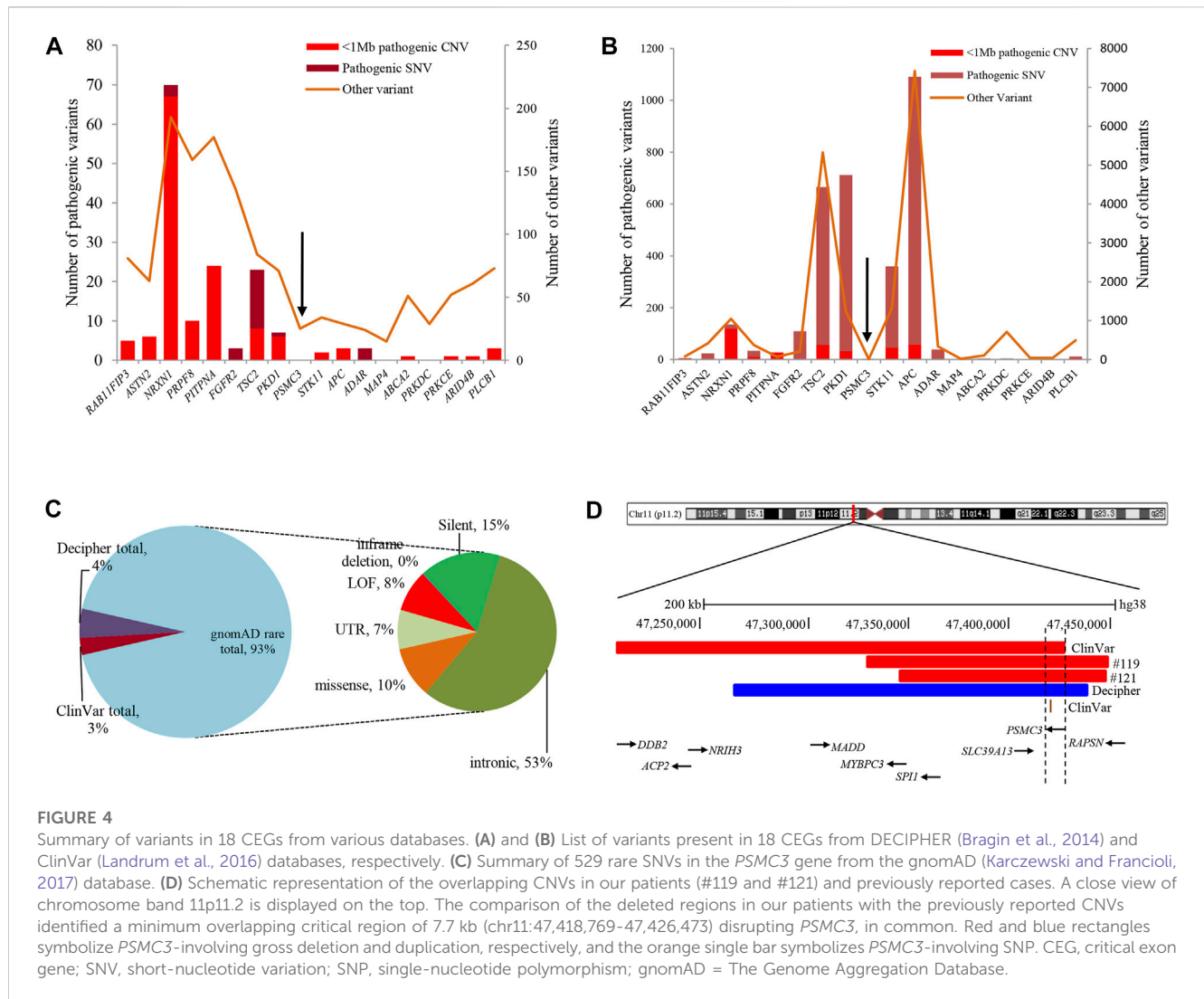|	Pathogenic	VOUS
Total CNVs	26	178
Total number of CEGs identified	195	33
Unique CEGs identified	153	31
Average number of CEG/CNV	7.50	0.19
*t (p)	4.33 (0.0002)	

*Welch's *t*-test.

31 in VOUS, including three genes found in both pathogenic CNVs and VOUS. On average, there were significantly more CEGs per pathogenic variant (7.5) than those in VOUS (0.19) (*p* = 0.0002) (Table 2). Through "critical-exon" analysis, we found 24 focal CNVs of <1 Mb that disrupt 22 unique CEGs in 22 NDD patients. We excluded four CEG genes: *PTPRD*, *RBFOX1*, *PCDH9*, and *LRRC4C* from further analysis, as only the intronic regions of the genes were disrupted by corresponding CNVs. After excluding these genes, the remaining 18 genes (*RAB11FIP3, ASTN2, NRXN1, PRPF8, PITPNA, FGFR2, TSC2, PKD1, PSMC3, STK11, APC, ADAR, MAP4, ABCA2, PRKDC, PRKCE, ARID4B,* and *PLCB1*) were analyzed further (Supplementary Table S6; Figures 4A, B). Of the 18 genes, we found the *PSMC3* gene as a possible candidate for ASD (Supplementary Table S7; Figure 4). We also found that a patient

(#106) carrying 2.01 Mb duplication CNVs had the most similar phenotypes to the *KMT2B*-related disorder patients previously described (Zech et al., 2016; Meyer et al., 2017; Zech et al., 2017; Faundes et al., 2018; Zech et al., 2019; Cif et al., 2020) (Figure 5; Supplementary Table S8).

Discussion

This is the first Bangladeshi NDD-cohort-conducted CMA analysis. Using the GenomeArc Analytics tool, we detected a diagnostic yield of 12.26% (26/212), which is in strong agreement with other ethnically diverse neurodevelopmental disorder cohorts (Miller et al., 2010; Uddin et al., 2016; Chaves et al., 2019). Of the 12.26% (26/212), five patients were carrying different lengths of deletion and duplication CNVs in the 15q11.2q13.1 region of chromosome 15 (Table 1; Supplementary Figure S7). The proximal long arm 'q' of human chromosome 15q11-q13 harbors a cluster of imprinted genes, the expressions of which are controlled by the imprinting center (Buiting et al., 1995). Disruptions of this region due to deletion or duplication CNVs are associated with three distinct neurodevelopmental disorders: Prader–Willi syndrome (PWS, #176270), Angelman syndrome (AS, #105830), and 15q11-q13 duplication syndrome (Dup15q syndrome, #608636) (Bragin et al., 2014; Kalsner and Chamberlain, 2015). We also found three patients carrying an extra copy of chromosome 21q (dup21q), which is associated with Down's syndrome (DS#190685). Moreover, there

FIGURE 4
Summary of variants in 18 CEGs from various databases. **(A)** and **(B)** List of variants present in 18 CEGs from DECIPHER (Bragin et al., 2014) and ClinVar (Landrum et al., 2016) databases, respectively. **(C)** Summary of 529 rare SNVs in the *PSMC3* gene from the gnomAD (Karczewski and Francioli, 2017) database. **(D)** Schematic representation of the overlapping CNVs in our patients (#119 and #121) and previously reported cases. A close view of chromosome band 11p11.2 is displayed on the top. The comparison of the deleted regions in our patients with the previously reported CNVs identified a minimum overlapping critical region of 7.7 kb (chr11:47,418,769-47,426,473) disrupting *PSMC3*, in common. Red and blue rectangles symbolize *PSMC3*-involving gross deletion and duplication, respectively, and the orange single bar symbolizes *PSMC3*-involving SNP. CEG, critical exon gene; SNV, short-nucleotide variation; SNP, single-nucleotide polymorphism; gnomAD = The Genome Aggregation Database.
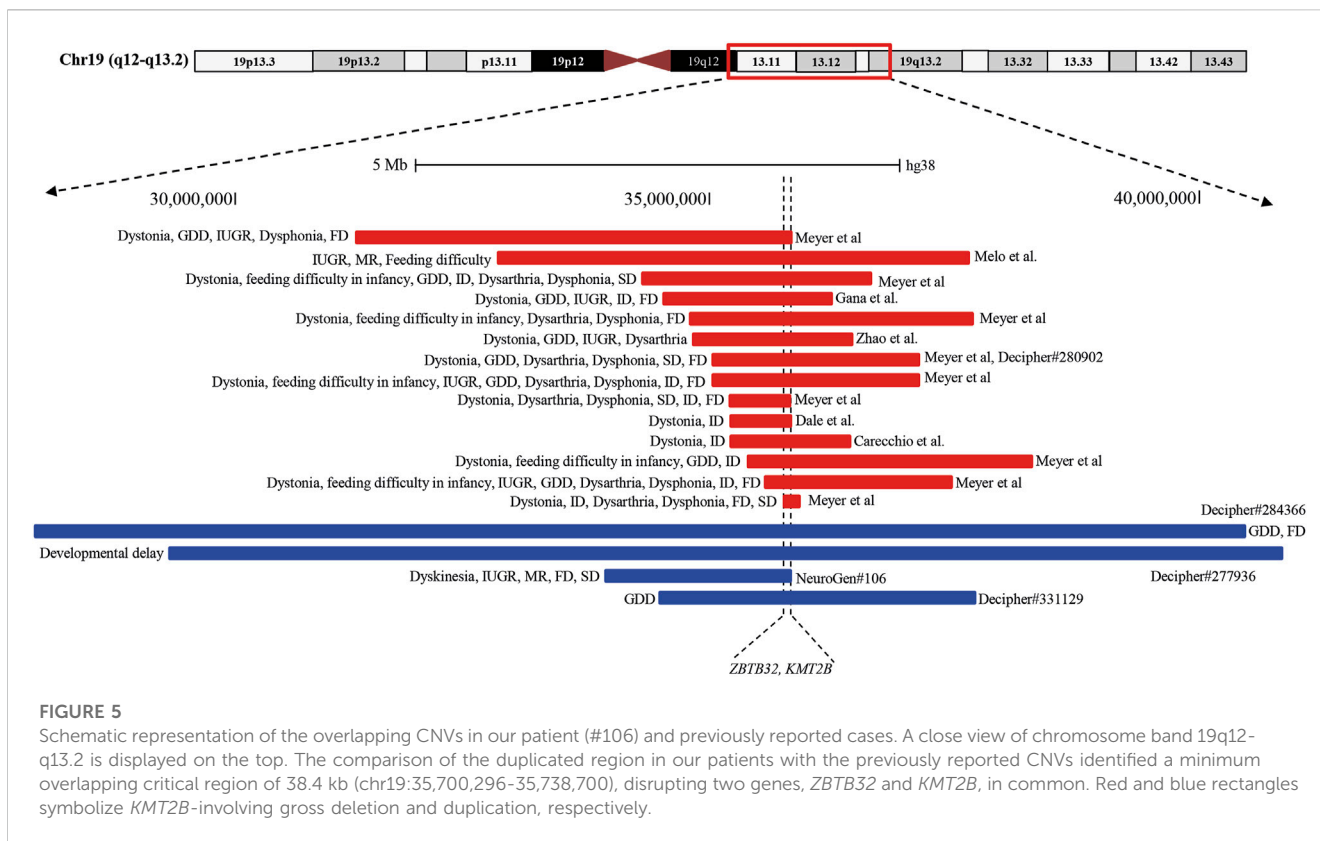
were four other patients carrying 3.5 Mb deletion in the 17p11.2 (#40), 1.4 Mb deletion in the 7q11.23 (#107), 1.69 Mb deletion in the 17q12 (#59), and 1.9 Mb duplication in the 22q11.22q11.23 (#6) chromosomal regions (Table 1) that overlapped with the critical regions implicated with some well-characterized syndromes (Bragin et al., 2014), including Smith–Magenis syndrome (SMS# 182290), Williams–Beuren syndrome (WS# 194050), renal cysts and diabetes syndrome (RCAD#137920), and 22q11.2 distal deletion syndrome (OMIM#611867), respectively.

Interestingly, 2.83% (6/212) pathogenic CNVs in this cohort were found to be in the subtelomeric region (Table 1), which is within the range of the previously reported (2.4%–4.4%) diagnostic rate in patients with DD/ID and congenital anomaly (Ravnan et al., 2006; Shao et al., 2008). Of the six CNVs, we found 4.8 Mb deletion in the 4p16.3p16.2 (#1), 11.2 Mb duplication in the 17p13.3p12 (#79), and 8.6 Mb deletion in the 2q37.1q37.3 (#144) subtelomeric chromosomal regions that overlapped with the critical regions associated with the three well-characterized syndromes (Bragin et al., 2014) of Wolf–Hirschhorn syndrome (WHS#194190), Miller–Dieker syndrome (MDS#247200), and 2q37 monosomy (#600430), respectively. Subtelomeric regions are gene-rich and

particularly prone to genomic instability due to the repeat sequences found in these areas. Mutation in these regions causes a variety of phenotypic abnormalities, particularly intellectual disability (ID) (Linardopoulou et al., 2005; Ruiter et al., 2007; Wu et al., 2010; Uddin et al., 2011; Rahman et al., 2019). Rahman et al. (2019) identified two siblings affected with ID who are both impacted by large terminal duplication (18.82 Mb) and deletion (3.90 Mb) CNVs. The remaining eight patients carried different lengths of deletion and duplication CNVs that disrupt the vital regions of the chromosome containing *bona fide* genes (*SCN1A, KANK1, DOCK8, etc.*) associated with different types of neurodevelopmental disorders (Nassir et al., 2021) (Table 1).

Identifying overlapping genes and pathways across disorders (Iourov et al., 2019; Zelenova et al., 2019) is critical to improving the understanding of their potential shared genetic etiology. Gene Ontology and KEGG pathway enrichment analysis of the impacted genes within the CNV breakpoints of all pathogenic deletions and duplication identified "ubiquitin-like protein transferase activity (GO:0019787)," "vesicle-mediated transport in synapse (GO:0099003)," "dependent protein catabolic process (GO: 0030163)," "regulation of neuronal death (GO:1901214)," and

FIGURE 5
Schematic representation of the overlapping CNVs in our patient (#106) and previously reported cases. A close view of chromosome band 19q12-q13.2 is displayed on the top. The comparison of the duplicated region in our patients with the previously reported CNVs identified a minimum overlapping critical region of 38.4 kb (chr19:35,700,296-35,738,700), disrupting two genes, *ZBTB32* and *KMT2B*, in common. Red and blue rectangles symbolize *KMT2B*-involving gross deletion and duplication, respectively.

"synaptic signaling (GO:0099536)" pathways to be highly significant (Figures 3A, B). Aberrations in autophagy-related (a major cellular catabolic process) gene mutations and signaling have been implicated in several neurodevelopmental disorders, including autism, tuberous sclerosis, Fragile X syndrome, and neurofibromatosis type 1 (Lee et al., 2013; Morice-Picard et al., 2016). Loss of function in the ubiquitin ligase gene *HERC2* has been associated with a severe neurodevelopmental phenotype (Morice-Picard et al., 2016). Further mutations in presynaptic genes have been linked to various neurodevelopmental disorders, including autism, ID (Zelenova et al., 2019), and epilepsy (Bonnycastle et al., 2020), and synaptic signaling has been identified as one of the principal molecular pathways affected in neurodevelopmental disorders (Parenti et al., 2020). Perturbations in the apoptotic signaling pathway have also been identified in various NDDs, including autism, Fragile X syndrome, and schizophrenia (Rudin and Thompson, 1997; Wei et al., 2014).

Analyses using the "critical-exon" method (Uddin et al., 2014, 2016) to identify constraint candidate genes discerned a high number of CEGs within the pathogenic variants (7.5) compared to VOUS (0.19) (p = 0.0002). Moreover, CEGs per pathogenic CNV were also previously reported (Wassman et al., 2019) as reflective of length bias and gene density of pathogenic variants. From the previous studies (Uddin et al., 2014; Uddin et al., 2016; Wassman et al., 2019), it is observed that critical exon genes are significantly enriched for *de novo* mutations in autism probands but not in unaffected siblings or in population control subjects. Identifying critical exon genes has shown (Wassman et al., 2019; Safizadeh Shabestari et al., 2022) improvement in clinical interpretations of rare CNVs. In this study, CEG analysis of short

focal CNVs identified 18 unique CEGs highly expressed in the neurotypical human brain and which have a low burden of non-synonymous variants in the control population. To discover the role of these genes within the context of neurodevelopmental disorders and our cohort, we conducted a comprehensive literature search. For example, in our cohort, one autism patient (#20) was carrying a 476-kb pathogenic deletion disrupting the *PLCB1* gene. Girirajan et al. (2013) found an enrichment of microdeletions and duplications involving the *PLCB1* gene in individuals with autism. The other 17 genes were located within VOUS CNVs. Of the 17, we tried to find a common overlapping genomic breakpoint shared by multiple patients with similar clinical condition to identify candidate genes for the disrupted locus. In this cohort, we found two autism patients (#119 and #121) who harbored a common deletion breakpoint disrupting four genes (*SLC39A13*, *PSMC3*, *SPI1*, and *RAPSN*), including one critical exon gene *PSMC3*. Autosomal recessive mutations in the *SLC39A13* gene are associated with a well-defined disease, Ehlers–Danlos syndrome, spondylodysplastic type 3 (OMIM# 612350). Mutation in the *SPI1* gene was previously reported in acute lymphoblastic leukemia (Seki et al., 2017). Autosomal recessive mutation in the *RAPSN* gene is associated with two other well-developed diseases, fetal akinesia deformation sequence 2 (#601592) and myasthenic syndrome, congenital, 11, associated with acetylcholine receptor deficiency (#616326). The other gene of the common breakpoint is *PSMC3* which encodes the 26S regulatory subunit 6A, also known as the 26S proteasome AAA-ATPase subunit (Rpt5) of the 19S proteasome complex responsible for recognition, unfolding, and translocation of substrates into the 20S proteolytic cavity of the proteasome (Tanaka, 2009). This suggests that *PSMC3* plays an

essential role in the ubiquitin–proteasome system (UPS), which includes morphogenesis, dendritic spine structure, synaptic activity, and regulation of synaptic strength in neurons (Hegde, 2010; Tai et al., 2010; Bingol and Sheng, 2011; Hamilton and Zito, 2013). Recently, Kröll-Hermi et al. (2020) demonstrated that homozygous single-nucleotide variants in the *PSMC3* cause neurosensory syndrome combining deafness, cataract, and autism/neurodevelopmental delay due to proteotoxic stress. Furthermore, we found one 217-kb deletion (Landrum et al., 2016) (ClinVar_VCV000151005) and one 173-kb duplication (Bragin et al., 2014) (Decipher_412053) CNVs in two previously reported NDD patients in whom the *PSMC3* gene was also disrupted (Figures 4C, D; Supplementary Table S7). Although our patients did not show syndromic appearance at the age of 1.6 and 3.1 years, except ASD, we hypothesize on the basis of previous studies in UPS (Hegde, 2010; Tai et al., 2010; Bingol and Sheng, 2011; Hamilton and Zito, 2013), including cases in the DECIPHER (Bragin et al., 2014) database that heterozygous mutations in the *PSMC3* (size ~7.7 kb) gene might be associated with NDDs without causing neurosensory syndrome.

In this study, we found a 20-year-old girl (#106), born in a consanguineous marriage with healthy parents, who was delivered preterm after an eventful pregnancy (IUGR). She had a history of delayed development. She had intellectual disability with dysmorphic features of an elongated face and long fingers (Supplementary Table S1). Her cognitive and physical conditions progressively worsened, starting at 15 years. At 18 years, she developed dyskinesia and speech and swallowing difficulties. Her MRI finding was normal, and she had no issues with walking or speech-related deficits. Analyzing the overlapping region of 14 previously reported deletions (length 0.19–4.91 Mb) (Dale et al., 2012; Gana et al., 2012; Bragin et al., 2014; Meyer et al., 2017; Zhao et al., 2018; Carecchio et al., 2019) and three duplications (length 3.31–12.63 Mb) (Bragin et al., 2014), our patient (#106) had a shorter 2.01 Mb duplication with a 38.4 kb (chr19: 35,700,296-35,738,700) common overlapping region among the CNVs disrupting two genes, *ZBTB32* and *KMT2B* (Figure 5; Supplementary Table S8). Furthermore, CEG and GenomeArc analytics also identified *UBA2*, *USF2*, *SCN1B*, *KMT2B*, *COX6B1*, *LGI4*, and *ZNF599* genes.

SCN1B is known to be associated with atrial fibrillation, familial, 13 (#615377), Brugada syndrome 5 (#612838), developmental and epileptic encephalopathy 52 (#617350), and epilepsy with febrile seizures plus and type 1 (#604233) due to the presence of pathogenic mutations. Our patient had no history of seizures, epilepsy, or cardiac problems, indicating that *SCN1B* duplication might not be associated with the patient's clinical condition. *KMT2B* is another interesting gene within this breakpoint where pathogenic mutations are associated with dystonia 28, childhood-onset (#617284) with core phenotypes described as limb-onset childhood dystonia that tends to spread progressively, resulting in generalized dystonia with craniocervical involvement. Co-occurring signs such as distinct facial dysmorphism and intellectual disability are most common (Meyer et al., 2017; Zech et al., 2019). A distinct group of *KMT2B* patients presented with neurodevelopmental disorders in the absence of dystonia or related movement disorders (Faundes et al., 2018; Zech et al., 2019; Cif et al., 2020). We also found some DECIPHER (Bragin et al., 2014) patients carrying a duplication CNV containing the *KMT2B* gene whose common phenotype was GDD in the absence of dystonia. Most of the clinical conditions of our patient (#106) in the form of global developmental delay, intrauterine growth retardation, intellectual

disability, facial dysmorphism, dyskinesia, and speech and swallowing problems match with the *KMT2B*-related disorder. So, we hypothesize that *KMT2B* gene duplication in our patient might be associated with the *KMT2B*-related disorder. Although *KMT2B* haploinsufficiency due to frameshift (small insertion/deletion), nonsense, splice-site, missense, and large deletion mutations is the primary cause of disease mechanism (Meyer et al., 2017; Zech et al., 2019; Cif et al., 2020), it is also reported that the penetrance for *KMT2B*-related disease is high, with almost complete penetrance for protein-truncating variants and chromosomal deletions and reduced penetrance for missense variants (Zech et al., 2019; Cif et al., 2020). Within the common overlapping region, *ZBTB32* is another new candidate gene in our cohort with no previous report of association with dystonia or neurodevelopmental disorder patients.

In our cohort, a series of overlapping rare clinically relevant variants were identified in multiple patients (Supplementary Table S9). For example, a 96.8-Kb deletion was found in two unrelated ASD patients (#8 and #127) at 6p21.33, which disrupts the major histocompatibility complex (MHC) class I gene *MICA*, which was not previously implicated with broader neurodevelopmental disorder patients. Three other unrelated patients (#13, #15, and #207) carried 48 Kb duplication at 15q13.3, which disrupts the previously reported (Gillentine and Schaaf, 2015; Gillentine et al., 2017; Uddin M. et al., 2018) broader NDD genes *CHRNA7* and *OTUD7A*. We found a 381-Kb duplication in two unrelated ASD patients (#34 and #104) at 8q21.2, disrupting carbonic anhydrases II (*CA2*), a gene associated with the disease osteopetrosis, autosomal recessive 3, OPTB3 (OMIM# 259730). We found two siblings (#42 and #43) affected with variable NDD phenotypes and one unrelated patient (#152) carrying 233 kb–340 kb duplication at 3p26.3, disrupting the previously reported (Hu et al., 2015; Repnikova et al., 2020) broader NDD-associated gene *CNTN6*. We found two other unrelated ASD patients carrying 130 Kb duplication CNV, disrupting the previously reported (Yasuda et al., 2014) NDD gene *NPHP1*. Three unrelated patients (#57, #159, and #200) carried a 236-Kb overlapping focal duplicated region, disrupting the two constraint genes *NPY4R* and *GPRIN2*. In addition, we found two unrelated ASD patients (#65 and #99) carrying 51 Kb and 59 Kb duplication at 4p16.3, disrupting alpha-L-iduronidase, the *IDUA* gene.

Conclusion

In this paper, we used multiple annotation tools (the "critical-exon" method and GenomeArc) to identify disease-associated candidate genes within rare CNVs. Although these tools can identify possible candidate genes, they may still miss some critical genes due to the limitations of the methods. We have shown the utility of chromosomal microarray analysis as a first-tier diagnostic technology for neurodevelopmental disorder patients. Without a proper genetic test, the clinical complexity alone may not be enough to identify the cause and often lead to a diagnostic odyssey. Although the price of microarray is reducing, it is still not affordable for most underdeveloped countries. This study shows how in near future developing countries can integrate such technology within their existing healthcare system. Despite the advantages of CMA, there are some notable limitations. Because

only unbalanced CNVs are detected, arrays cannot identify balanced inversions/insertions or reciprocal translocations. Likewise, because of the overall resolution, arrays also will miss low-level mosaicism (typically below 20%) and point mutations or small insertions or deletions in single genes. Although there are limitations of CMA, it is true that this technology can precisely detect 10%–30% of NDD cases in a cost-effective way and will enable the detection of novel variants and genes from underrepresented populations.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Ethics statement

The study protocol was approved by the Institutional Review Board of Holy Family Red Crescent Medical College and Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

Conceptualization: MU, HA, and MMR; preparation of the draft manuscript: MU, HA, MMR, SLS, SKS, MB, MMI, MAR(1), MAR(2), MAR(3), MMH, ML, NJ, AM, MOF, TBE, NS, GK, SIK, RKK, KF, MH, MAHM, NK, BZ, NN, ANN, MRA and KMFU; figure preparation: MU, HA, SH, and MAR(2); microarray detection: MU, HA, MAR(1), and NJD; analysis: MU, GB, AI, HA, NN, MIH, MAR(1), MAR(2), and NJD; interpretation of variants: MU, HA, MAR(1), NJD and AI; comprehensive literature search: MU, HA, ANN, NJD, MAR(1), NN, BZ, SH, TBE, and SKS. All authors have read and agreed to the published version of the manuscript.

Funding

Acknowledgments

Conflict of interest

Authors AI and MU are employees of GenomeArc Inc. Authors HA, SLS, MB, MMI, MAR(1), MAR(2), TBE, NJD, ML, NJ, AI, AM, MOF, and KMFU are employees of NeuroGen Healthcare. The authors declare that this study received funding from NeuroGen Healthcare. The funder was not involved in the study design, collection, analysis, interpretation of data, and the writing of this article or the decision to submit it for publication.

The remaining 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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.955631/full#supplementary-material

References

Abou Jamra, R., Wohlfart, S., Zweier, M., Uebe, S., Priebe, L., Ekici, A., et al. (2011). Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity. *Eur. J. Hum. Genet.* 19, 1161–1166. doi:10.1038/ejhg.2011.98

Ahn, K., Gotay, N., Andersen, T. M., Anvari, A. A., Gochman, P., Lee, Y., et al. (2014). High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol. Psychiatry* 19, 568–572. doi:10.1038/mp.2013.59

Akter, H., Sultana, N., Martuza, N., Siddiqua, A., Dity, N. J., Rahaman, M. A., et al. (2019). Novel mutations in actionable breast cancer genes by targeted sequencing in an ethnically homogenous cohort. *BMC Med. Genet.* 20, 150–158. doi:10.1186/S12881-019-0881-0

Akter, H., Hossain, M. S., Dity, N. J., Rahaman, M. A., Furkan Uddin, K. M., Nassir, N., et al. (2021). Whole exome sequencing uncovered highly penetrant recessive mutations for a spectrum of rare genetic pediatric diseases in Bangladesh. *npj Genomic Med.* 6, 14–19. doi:10.1038/s41525-021-00173-0

Bader, G. D., Isserlin, R., Merico, D., and Voisin, V. (2014). Enrichment map – A Cytoscape app to visualize and explore OMICs pathway enrichment results. *F1000Research* 3, 141. doi:10.12688/F1000RESEARCH.4536.1

Bale, T. L., Baram, T. Z., Brown, A. S., Goldstein, J. M., Insel, T. R., McCarthy, M. M., et al. (2010). Early life programming and neurodevelopmental disorders. *Biol. Psychiatry* 68, 314–319. doi:10.1016/j.biopsych.2010.05.028

Begum, G., Albanna, A., Bankapur, A., Nassir, N., Tambi, R., Berdiev, B. K., et al. (2021). Long-read sequencing improves the detection of structural variations impacting complex non-coding elements of the genome. *Int. J. Mol. Sci.* 22, 2060–2114. doi:10.3390/IJMS22042060

Bierut, L. J., Madden, P. A. F., Breslau, N., Johnson, E. O., Hatsukami, D., Pomerleau, O. F., et al. (2007). Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum. Mol. Genet.* 16, 24–35. doi:10.1093/hmg/ddl441

Bierut, L. J., Agrawal, A., Bucholz, K. K., Doheny, K. F., Laurie, C., Pugh, E., et al. (2010). A genome-wide association study of alcohol dependence. *Proc. Natl. Acad. Sci. U. S. A.* 107, 5082–5087. doi:10.1073/pnas.0911109107

Bingol, B., and Sheng, M. (2011). Deconstruction for reconstruction: The role of proteolysis in neural plasticity and disease. *Neuron* 69, 22–32. doi:10.1016/j.neuron.2010.11.006

Bonnycastle, K., Davenport, E. C., and Cousin, M. A. (2020). Presynaptic dysfunction in neurodevelopmental disorders: Insights from the synaptic vesicle life cycle. *J. Neurochem.* 157, 179–207. doi:10.1111/jnc.15035

Bragin, E., Chatzimichali, E. A., Wright, C. F., Hurles, M. E., Firth, H. V., Bevan, A. P., et al. (2014). DECIPHER: Database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Res.* 42, D993–D1000. doi:10.1093/nar/gkt937

Buiting, K., Saitoh, S., Gross, S., Dittrich, B., Schwartz, S., Nicholls, R. D., et al. (1995). Inherited microdeletions in the Angelman and Prader–Willi syndromes define an imprinting centre on human chromosome 15. *Nat. Genet.* 9, 395–400. doi:10.1038/ng0495-395

Careccio, M., Invernizzi, F., Gonzàlez-Latapi, P., Panteghini, C., Zorzi, G., Romito, L., et al. (2019). Frequency and phenotypic spectrum of KMT2B dystonia in childhood: A single-center cohort study. *Mov. Disord.* 34, 1516–1527. doi:10.1002/MDS.27771

Chaves, T. F., Baretto, N., de Oliveira, L. F., Ocampos, M., Barbato, I. T., Anselmi, M., et al. (2019). Copy number variations in a cohort of 420 individuals with neurodevelopmental disorders from the south of Brazil. *Sci. Rep.* 9, 17776. doi:10.1038/s41598-019-54347-z

Chen, E. S., Gigek, C. O., Rosenfeld, J. A., Diallo, A. B., Maussion, G., Chen, G. G., et al. (2014). Molecular convergence of neurodevelopmental disorders. *Am. J. Hum. Genet.* 95, 490–508. doi:10.1016/j.ajhg.2014.09.013

Chien, W. H., Gau, S. S. F., Liao, H. M., Chiu, Y. N., Wu, Y. Y., Huang, Y. S., et al. (2013). Deep exon resequencing of DLGAP2 as a candidate gene of autism spectrum disorders. *Mol. Autism* 4, 26. doi:10.1186/2040-2392-4-26

Cif, L., Demailly, D., Lin, J. P., Barwick, K. E., Sa, M., Abela, L., et al. (2020). KMT2B-related disorders: Expansion of the phenotypic spectrum and long-term efficacy of deep brain stimulation. *Brain* 143, 3242–3261. doi:10.1093/brain/awaa304

Cline, M. S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* 2, 2366–2382. doi:10.1038/NPROT.2007.324

Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. A., Vu, T. H., Baker, C., et al. (2011). A copy number variation morbidity map of developmental delay. *Nat. Genet.* 43, 838–846. doi:10.1038/ng.909

Coviello, A. D., Haring, R., Wellons, M., Vaidya, D., Lehtimäki, T., Keildson, S., et al. (2012). A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple loci implicated in sex steroid hormone regulation. *PLoS Genet.* 8, e1002805. doi:10.1371/journal.pgen.1002805

Cristino, A. S., Williams, S. M., Hawi, Z., An, J. Y., Bellgrove, M. A., Schwartz, C. E., et al. (2014). Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. *Mol. Psychiatry* 19, 294–301. doi:10.1038/mp.2013.16

Dale, R. C., Grattan-Smith, P., Nicholson, M., and Peters, G. B. (2012). Microdeletions detected using chromosome microarray in children with suspected genetic movement disorders: A single-centre study. *Dev. Med. Child. Neurol.* 54, 618–623. doi:10.1111/J.1469-8749.2012.04287.X

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., and Fletcher, J. M. J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J. Int. Neuropsychol. Soc.* 15, 331–343. doi:10.1017/S1355617709090481

Faundes, V., Newman, W. G., Bernardini, L., Canham, N., Clayton-Smith, J., Dallapiccola, B., et al. (2018). Histone lysine methylases and demethylases in the landscape of human developmental disorders. *Am. J. Hum. Genet.* 102, 175–187. doi:10.1016/j.ajhg.2017.11.013

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *J. Autism Dev. Disord.* 29, 439–484. doi:10.1023/A:1021943802493

Gana, S., Veggiotti, P., Sciacca, G., Fedeli, C., Bersano, A., Micieli, G., et al. (2012). 19q13.11 cryptic deletion: Description of two new cases and indication for a role of WTIP haploinsufficiency in hypospadias. *Eur. J. Hum. Genet.* 20, 852–856. doi:10.1038/ejhg.2012.19

Gillentine, M. A., and Schaaf, C. P. (2015). The human clinical phenotypes of altered CHRNA7 copy number. *Biochem. Pharmacol.* 97, 352–362. doi:10.1016/j.bcp.2015.06.012

Gillentine, M. A., Yin, J., Bajic, A., Zhang, P., Cummock, S., Kim, J. J., et al. (2017). Functional consequences of CHRNA7 copy-number alterations in induced pluripotent

stem cells and neural progenitor cells. *Am. J. Hum. Genet.* 101, 874–887. doi:10.1016/j.ajhg.2017.09.024

Girirajan, S., Dennis, M. Y., Baker, C., Malig, M., Coe, B. P., Campbell, C. D., et al. (2013). Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. *Am. J. Hum. Genet.* 92, 221–237. doi:10.1016/j.ajhg.2012.12.016

Hamdan, F. F., Daoud, H., Piton, A., Gauthier, J., Dobrzeniecka, S., Krebs, M. O., et al. (2011). De novo syngap1 mutations in nonsyndromic intellectual disability and autism. *Biol. Psychiatry* 69, 898–901. doi:10.1016/j.biopsych.2010.11.015

Hamilton, A. M., and Zito, K. (2013). Breaking it down: The ubiquitin proteasome system in neuronal morphogenesis. *Neural Plast.* 2013, 196848. doi:10.1155/2013/196848

Hampson, D. R., Adusei, D. C., and Pacey, L. K. K. (2011). The neurochemical basis for the treatment of autism spectrum disorders and Fragile X Syndrome. *Biochem. Pharmacol.* 81, 1078–1086. doi:10.1016/j.bcp.2011.02.005

Hegde, A. N. (2010). The ubiquitin-proteasome pathway and synaptic plasticity. *Learn. Mem.* 17, 314–327. doi:10.1101/lm.1504010

Ho, K. S., Wassman, E. R., Baxter, A. L., Hensel, C. H., Martin, M. M., Prasad, A., et al. (2016). Chromosomal microarray analysis of consecutive individuals with autism spectrum disorders using an ultra-high resolution chromosomal microarray optimized for neurodevelopmental disorders. *Int. J. Mol. Sci.* 17 (17), 2070. doi:10.3390/IJMS17122070

Hu, J., Liao, J., Sathanoori, M., Kochmar, S., Sebastian, J., Yatsenko, S. A., et al. (2015). CNTN6 copy number variations in 14 patients: A possible candidate gene for neurodevelopmental and neuropsychiatric disorders. *J. Neurodev. Disord.* 7, 26. doi:10.1186/S11689-015-9122-9

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2019). Pathway-based classification of genetic diseases. *Mol. Cytogenet* 12, 4. doi:10.1186/s13039-019-0418-4

Kalsner, L., and Chamberlain, S. J. (2015). Prader-Willi, Angelman, and 15q11-q13 duplication syndromes. *Pediatr. Clin. North. Am.* 62, 587–606. doi:10.1016/j.pcl.2015.03.004

Karczewski, K. J., and Francioli, L. (2017). *The genome aggregation database (gnomAD)*. MacArthur Lab. Available at: https://macarthurlab.org/2017/02/27/the-genome-aggregation-database-gnomad/.

Kearney, H. M., Thorland, E. C., Brown, K. K., Quintero-Rivera, F., and South, S. T.Working Group of the American College of Medical Genetics Laboratory Quality Assurance Committee (2011). American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet. Med.* 13, 680–685. doi:10.1097/GIM.0b013e3182217a3a

Knopf, A. (2020). Autism prevalence increases from 1 in 60 to 1 in 54: CDC. *Brown Univ. Child. Adolesc. Behav. Lett.* 36, 4. doi:10.1002/CBL.30470

Krakowiak, P., Walker, C. K., Bremer, A. A., Baker, A. S., Ozonoff, S., Hansen, R. L., et al. (2012). Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. *Pediatrics* 129, e1121–e1128. doi:10.1542/peds.2011-2583

Krawczak, M., Nikolaus, S., Von Eberstein, H., Croucher, P. J. P., El Mokhtari, N. E., and Schreiber, S. (2006). PopGen: Population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet.* 9, 55–61. doi:10.1159/000090694

Kröll-Hermi, A., Ebstein, F., Stoetzel, C., Geoffroy, V., Schaefer, E., Scheidecker, S., et al. (2020). Proteasome subunit PSMC3 variants cause neurosensory syndrome combining deafness and cataract due to proteotoxic stress. *EMBO Mol. Med.* 12, e11861. doi:10.15252/emmm.201911861

Krzywinski, M., Schein, J., Inanc, X., Birol, I., Connors, J., Gascoyne, R., et al. (2009). Circos: An information aesthetic for comparative genomics. *genome.cshlp.Org.* 19, 1639–1645. doi:10.1101/gr.092759.109

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., et al. (2016). ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 44, D862–D868. doi:10.1093/NAR/GKV1222

Lee, K.-M., Hwang, S.-K., and Lee, J.-A. (2013). Neuronal autophagy and neurodevelopmental disorders. *Exp. Neurobiol.* 22, 133–142. doi:10.5607/en.2013.22.3.133

Levitt, P., Eagleson, K. L., and Powell, E. M. (2004). Regulation of neocortical interneuron development and the implications for neurodevelopmental disorders. *Trends Neurosci.* 27, 400–406. doi:10.1016/j.tins.2004.05.008

Linardopoulou, E. V., Williams, E. M., Fan, Y., Friedman, C., Young, J. M., and Trask, B. J. (2005). Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* 437, 94–100. doi:10.1038/nature04029

Lionel, A. C., Crosbie, J., Barbosa, N., Goodale, T., Thiruvahindrapuram, B., Rickaby, J., et al. (2011). Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. *Sci. Transl. Med.* 3, 95ra75. doi:10.1126/scitranslmed.3002464

Liu, Y., Zhang, Y., Zhao, D., Dong, R., Yang, X., Tammimies, K., et al. (2015a). Rare de novo deletion of metabotropic glutamate receptor 7 (GRM7) gene in a patient with autism spectrum disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 168, 258–264. doi:10.1002/AJMG.B.32306

Liu, Y., Zhao, D., Dong, R., Yang, X., Zhang, Y., Tammimies, K., et al. (2015b). De novo exon 1 deletion of AUTS2 gene in a patient with autism spectrum disorder and developmental delay: A case report and a brief literature review. *Am. J. Med. Genet. A* 167, 1381–1385. doi:10.1002/AJMG.A.37050

Lord, C., Rutter, M., Dilavore, P. C., Risi, S., Gotham, K., and Bishop, S. L. (2012). *Autism diagnostic observation schedule.* 2nd Edn. (ADOS-2). Manual (Part I) (Los Angeles: Western Psychological Services).

Matson, J. L., and Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Res. Dev. Disabil.* 30, 1107–1114. doi:10.1016/j.ridd.2009.06.003

Meyer, E., Carss, K. J., Rankin, J., Nichols, J. M. E., Grozeva, D., Joseph, A. P., et al. (2017). Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. *Nat. Genet.* 49, 223–237. doi:10.1038/ng.3740

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., et al. (2010). Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* 86, 749–764. doi:10.1016/j.ajhg.2010.04.006

Moessner, R., Marshall, C. R., Sutcliffe, J. S., Skaug, J., Pinto, D., Vincent, J., et al. (2007). Contribution of SHANK3 mutations to autism spectrum disorder. *Am. J. Hum. Genet.* 81, 1289–1297. doi:10.1086/522590

Morice-Picard, F., Benard, G., Rezvani, H. R., Lasseaux, E., Simon, D., Moutton, S., et al. (2016). Complete loss of function of the ubiquitin ligase HERC2 causes a severe neurodevelopmental phenotype. *Eur. J. Hum. Genet.* 25, 52–58. doi:10.1038/ejhg.2016.139

Mullick, M. S. I., and Goodman, R. (2005). The prevalence of psychiatric disorders among 5-10 year olds in rural, urban and slum areas in Bangladesh: An exploratory study. *Soc. Psychiatry Psychiatr. Epidemiol.* 40, 663–671. doi:10.1007/s00127-005-0939-5

Nassir, N., Bankapur, A., Samara, B., Ali, A., Ahmed, A., Inuwa, I. M., et al. (2021). Single-cell transcriptome identifies molecular subtype of autism spectrum disorder impacted by de novo loss-of-function variants regulating glial cells. *Hum. Genomics* 15, 68. doi:10.1186/S40246-021-00368-7

Oskoui, M., Gazzellone, M. J., Thiruvahindrapuram, B., Zarrei, M., Andersen, J., Wei, J., et al. (2015). Clinically relevant copy number variations detected in cerebral palsy. *Nat. Commun.* 6, 7949. doi:10.1038/ncomms8949

Parenti, I., Rabaneda, L. G., Schoen, H., and Novarino, G. (2020). Neurodevelopmental disorders: From genetics to functional pathways. *Trends Neurosci.* 43, 608–621. doi:10.1016/j.tins.2020.05.004

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., et al. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* 94, 677–694. doi:10.1016/j.ajhg.2014.03.018

Rahaman, M. A., Lopa, M., Uddin, K. M. F., Baqui, M. A., Keya, S. P., Faruk, M. O., et al. (2020). An exploration of physical and phenotypic characteristics of Bangladeshi children with autism spectrum disorder. *J. Autism Dev. Disord.* 51, 2392–2401. doi:10.1007/s10803-020-04703-0

Rahman, M. M., Uddin, K. F., Al Jezawi, N. K., Karuvantevida, N., Akter, H., Dity, N. J., et al. (2019). Gonadal mosaicism of large terminal de novo duplication and deletion in siblings with variable intellectual disability phenotypes. *Mol. Genet. Genomic Med.* 7, e00954. doi:10.1002/mgg3.954

Ravnan, J. B., Tepperberg, J. H., Papenhausen, P., Lamb, A. N., Hedrick, J., Eash, D., et al. (2006). Subtelomere FISH analysis of 11 688 cases: An evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *J. Med. Genet.* 43, 478–489. doi:10.1136/jmg.2005.036350

Repnikova, E. A., Lyalin, D. A., McDonald, K., Astbury, C., Hansen-Kiss, E., Cooley, L. D., et al. (2020). CNTN6 copy number variations: Uncertain clinical significance in individuals with neurodevelopmental disorders. *Eur. J. Med. Genet.* 63, 103636. doi:10.1016/J.EJMG.2019.02.008

Rudin, C. M., and Thompson, C. B. (1997). Apoptosis and disease: Regulation and clinical relevance of programmed cell death. *Annu. Rev. Med.* 48, 267–281. doi:10.1146/annurev.med.48.1.267

Ruiter, E. M., Koolen, D. A., Nillesen, W. M., Pfundt, R., de Leeuw, N., Hamel, B. C. J., et al. (2007). Pure subtelomeric microduplications as a cause of mental retardation. *Clin. Genet.* 72, 362–368. doi:10.1111/j.1399-0004.2007.00874.x

Safizadeh Shabestari, S. A., Nassir, N., Sopariwala, S., Karimov, I., Tambi, R., Zehra, B., et al. (2022). Overlapping pathogenic de novo CNVs in neurodevelopmental disorders and congenital anomalies impacting constraint genes regulating early development. *Hum. Genet.*, 1–13. doi:10.1007/s00439-022-02482-5

Seki, M., Kimura, S., Isobe, T., Yoshida, K., Ueno, H., Nakajima-Takagi, Y., et al. (2017). Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat. Genet.* 49, 1274–1281. doi:10.1038/ng.3900

Shao, L., Shaw, C. A., Lu, X. Y., Sahoo, T., Bacino, C. A., Lalani, S. R., et al. (2008). Identification of chromosome abnormalities in subtelomeric regions by microarray analysis: A study of 5,380 cases. *Am. J. Med. Genet. Part A* 146, 2242–2251. doi:10.1002/ajmg.a.32399

Shevell, M., Ashwal, S., Donley, D., Flint, J., Gingold, M., Hirtz, D., et al. (2003). Practice parameter: Evaluation of the child with global developmental delay: Report of

the quality standards subcommittee of the American academy of neurology and the practice committee of the child neurology society. *Neurology* 60, 367–380. doi:10.1212/01.WNL.0000031431.81555.16

Stewart, A. F. R., Dandona, S., Chen, L., Assogba, O., Belanger, M., Ewart, G., et al. (2009). Kinesin family member 6 variant Trp719Arg does not associate with angiographically defined coronary artery disease in the Ottawa Heart genomics study. *J. Am. Coll. Cardiol.* 53, 1471–1472. doi:10.1016/j.jacc.2008.12.051

Tai, H. C., Besche, H., Goldberg, A. L., and Schuman, E. M. (2010). Characterization of the brain 26S proteasome and its interacting proteins. *Front. Mol. Neurosci.* 3, 12. doi:10.3389/fnmol.2010.00012

Tanaka, K. (2009). The proteasome: Overview of structure and functions. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 85, 12–36. doi:10.2183/pjab.85.12

Uddin, M., Sturge, M., Peddle, L., O'Rielly, D. D., and Rahman, P. (2011). Genome-wide signatures of "rearrangement hotspots" within segmental duplications in humans. *PLoS One* 6, 28853. doi:10.1371/journal.pone.0028853

Uddin, M., Tammimies, K., Pellecchia, G., Alipanahi, B., Hu, P., Wang, Z., et al. (2014). Brain-expressed exons under purifying selection are enriched for de novo mutations in autism spectrum disorder. *Nat. Genet.* 46, 742–747. doi:10.1038/ng.2980

Uddin, M., Pellecchia, G., Thiruvahindrapuram, B., D'Abate, L., Merico, D., Chan, A., et al. (2016). Indexing effects of copy number variation on genes involved in developmental delay. *Sci. Rep.* 6, 28663–28712. doi:10.1038/srep28663

Uddin, K. M. F., Amin, R., Majumder, S. N., Aleem, M. A., Rahaman, A., Dity, N. J., et al. (2018). An ANKRD26 nonsense somatic mutation in a female with epidermodysplasia verruciformis (Tree Man Syndrome). *Clin. Case Rep.* 6, 1426–1430. doi:10.1002/ccr3.1595

Uddin, M., Unda, B. K., Kwan, V., Holzapfel, N. T., White, S. H., Chalil, L., et al. (2018). OTUD7A regulates neurodevelopmental phenotypes in the 15q13.3 microdeletion syndrome. *Am. J. Hum. Genet.* 102, 278–295. doi:10.1016/j.ajhg.2018.01.006

Uddin, K. F., Amin, M. R., Sultana, N., Aleem, M. A., Sarker, S. K., Banu, L. A., et al. (2019). Detection of clinically relevant copy number variation of SEZ6L2 gene in a Bangladeshi autism spectrum disorder cohort. *Bangladesh J. Med.* 30, 24–29. doi:10.3329/BJMED.V30I1.39919

Verhoeven, V. J. M., Hysi, P. G., Wojciechowski, R., Fan, Q., Guggenheim, J. A., Höhn, R., et al. (2013). Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat. Genet.* 45, 314–318. doi:10.1038/ng.2554

Wassman, E. R., Ho, K. S., Bertrand, D., Davis, K. W., Martin, M. M., Page, S., et al. (2019). Critical exon indexing improves clinical interpretation of copy number variants in neurodevelopmental disorders. *Neurol. Genet.* 5, e378. doi:10.1212/nxg.0000000000000378

Wei, H., Alberts, I., and Li, X. (2014). The apoptotic perspective of autism. *Int. J. Dev. Neurosci.* 36, 13–18. doi:10.1016/j.ijdevneu.2014.04.004

Woodbury-Smith, M., Nicolson, R., Zarrei, M., Yuen, R. K. C., Walker, S., Howe, J., et al. (2017). Variable phenotype expression in a family segregating microdeletions of the NRXN1 and MBD5 autism spectrum disorder susceptibility genes. *npj Genomic Med.* 21 (2), 17–18. doi:10.1038/s41525-017-0020-9

Wu, Y., Ji, T., Wang, J., Xiao, J., Wang, H., Li, J., et al. (2010). Submicroscopic subtelomeric aberrations in Chinese patients with unexplained developmental delay/mental retardation. *BMC Med. Genet.* 11, 72. doi:10.1186/1471-2350-11-72

Yasuda, Y., Hashimoto, R., Fukai, R., Okamoto, N., Hiraki, Y., Yamamori, H., et al. (2014). Duplication of the NPHP1 gene in patients with autism spectrum disorder and normal intellectual ability: A case series. *Ann. Gen. Psychiatry* 13, 22. doi:10.1186/S12991-014-0022-2

Zech, M., Boesch, S., Maier, E. M., Borggraefe, I., Vill, K., Laccone, F., et al. (2016). Haploinsufficiency of KMT2B, encoding the lysine-specific histone methyltransferase 2B, results in early-onset generalized dystonia. *Am. J. Hum. Genet.* 99, 1377–1387. doi:10.1016/j.ajhg.2016.10.010

Zech, M., Jech, R., Havránková, P., Fečíková, A., Berutti, R., Urgošík, D., et al. (2017). KMT2B rare missense variants in generalized dystonia. *Mov. Disord.* 32, 1087–1091. doi:10.1002/mds.27026

Zech, M., Lam, D. D., and Winkelmann, J. (2019). Update on KMT2B-related dystonia. *Curr. Neurol. Neurosci. Rep.* 19, 92–11. doi:10.1007/s11910-019-1007-y

Zelenova, M. A., Yurov, Y. B., Vorsanova, S. G., and Iourov, I. Y. (2019). Laundering CNV data for candidate process prioritization in brain disorders. *Mol. Cytogenet.* 12, 54–56. doi:10.1186/s13039-019-0468-7

Zhang, Y., Haraksingh, R., Grubert, F., Abyzov, A., Gerstein, M., Weissman, S., et al. (2013). Child development and structural variation in the human genome. *Child. Dev.* 84, 34–48. doi:10.1111/cdev.12051

Zhao, K., van der Spoel, A., Castiglioni, C., Gale, S., Fujiwara, H., Ory, D. S., et al. (2018). 19q13.12 microdeletion syndrome fibroblasts display abnormal storage of cholesterol and sphingolipids in the endo-lysosomal system. *Biochim. Biophys. acta. Mol. Basis Dis.* 1864, 2108–2118. doi:10.1016/J.BBADIS.2018.03.020

Zoghbi, H. Y. (2003). Postnatal neurodevelopmental disorders: Meeting at the synapse? *Science* 80-302, 826–830. doi:10.1126/science.1089071