Detection of Clinically Relevant Copy Number Variations and Genes in a Bangladeshi Cohort of Neurodevelopmental Disorders

Hosneara Akter
Genetics and Genomic medicine Centre, NeuroGen Healthcare

Muhammad Mizanur Rahman
Bangabandhu Sheikh Mujib Medical University

Shaoliarker
Department of Child Neurology, NeuroGen Healthcare, Dhaka

Mohammed Basinuzzaman
Department of Child Neurology, NeuroGen Healthcare

Mazharul Islam
Department of child Neurology, NeuroGen Healthcare

Md. Atikur Rahaman
Genetics and Genomic Medicine Centre, NeuroGen Healthcare

Md. Ashiquir Rahaman
Centre for Precision Therapeutics, NeuroGen Healthcare

Tamannyat Binte Eshaque
Genetics and Genomic Medicine Centre, NeuroGen Healthcare

Nushrat Jahan Dity
Genetics and Genomic Medicine Centre, NeuroGen Healthcare

Shouvik Sarker
Leibniz University Hanover: Leibniz Universitat Hannover

Md. Robed Amin
Dhaka Medical College and Hospital

Md Monir Hossain
National Institute of Neurosciences & Hospital

Maksuda Lopa
Centre for Precision Therapeutics, NeuroGen Healthcare, Dhaka

Nargis Jahan
Centre for Precision Therapeutics, NeuroGen Healthcare, Dhaka

Shafaat Hossain
University of Houston

Aminul Islam
Genetics and Genomic Medicine Centre, NeuroGen Healthcare, Dhaka

Ashaduzzaman Mondol
Centre for Precision Therapeutics, NeuroGen Healthcare, Dhaka

Md Omar Faruk
Centre for Precision Therapeutics, NeuroGen Healthcare, Dhaka

Narayan Saha
National Institute of Neurosciences & Hospital

Gopenkumar Kundu
Bangabandhu Sheikh Mujib Medical University

Shayla Imam Kanta
Department of Paediatric Neuroscience, Dhaka Shishu Hospital

Rezaul Karim Kazal
Bangabandhu Sheikh Mujib Medical University

Kanil Fatema
Bangabandhu Sheikh Mujib Medical University

Md. Ashrafur Rahman
North South University

Manuf Hasan
Military Institute of Science and Technology

Md. Abid Hossain Mollah
Keywords: Neurodevelopmental Disorders (NDD), Chromosomal Microarray Analysis (CMA), Copy number variation (CNV), Autism Diagnostic Observation Schedule-Second Edition (ADOS-2), Variant of Uncertain Significance (VOUS), Critical Exon Gene (CEG), Autism Spectrum Disorder (ASD)

Posted Date: October 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-948388/v1

License: ©  This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: Copy number variations (CNVs) play a critical role into 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 genome-wide chromosomal microarray analysis (CMA) for 212 NDD patients with male to female ratio of 2.2:1.0 to identify rare chromosomal abnormalities (deletion /duplication/ rearrangements). To identify candidate genes within the rare CNVs, multiple 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 R package.

Results: In our cohort, the head circumference of males are significantly greater than females (p=0.0002). Of all samples assayed, 12.26% (26/212) and 47.17% (100/212) patients carried pathogenic and variant of uncertain significance (VOUS) CNVs, respectively. 2.83% (6/212) pathogenic CNVs are located at the subtelomeric regions. Further burden test identified females are significant carriers of pathogenic CNVs in comparison to males (OR=4.2; p=0.0007).

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

Introduction

Neurodevelopmental disorders (NDDs) are a group of developmental deficits that disrupt normal physiology and function of the brain. These disorders are referred as a collection of early onset conditions that include autism spectrum disorders (ASDs), intellectual disability (ID), epilepsy encephalopathy, attention deficit hyperactive disorders (ADHD), obsessive compulsive disorder (OCD), and cognitive skills disorders [1–5]. Such disorders, when isolated, are termed non-syndromic; when associated with the presence of dysmorphisms or apparent congenital anomalies (CA), are termed syndromic [6]. The incidence of DD/ID is 3% in the general population [7] while the statistics from USA shows that ASD affects 1 in 54 live births [8]. Individuals affected with NDDs usually present reduced adaptive skills, limited intellectual ability, motor difficulties, CA and problems with social interaction. Phenotypically, there are major overlaps among ASDs with epilepsy encephalopathy, ADHD, Fragile X syndrome (FXS), motor abnormalities and intellectual disability [9, 10].

The etiology of NDDs is principally genetic. Advancement of genomic techniques such as high-resolution microarrays and next-generation sequencing have yielded significant insights into the genetic etiology of NDDs [11]. For decades, structural genomic variation known to be a major contributor into the etiology of a proportion of children diagnosed with NDDs [12–14]. In the last decade, many international genomic consortiums have profiled NDD cases mostly from European ancestry to identify genomic alterations and NDD associated genes. More than 100 genes and genomic loci [15] have been consistently found to be involved in the etiology of NDDs. Studies based on ASD cohorts have identified an increased burden of rare genic copy-number variations (CNVs) and have characterized rare, usually de novo, recurrent CNV loci that are thought to contribute to the genetic risk [16]. Specific genes within these CNV regions that are implicated in the etiology of ASD and other NDDs include SHANK3, SYNGAP1, NRXN1, GRM7 and DLGAP2, etc. [17–21]. As the number of candidate genes and loci have 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 [22, 23]. The availability of microarray related technologies and the contribution of structural variations to NDD enabled whole genome chromosomal microarray (CMA) as one of the first tier diagnostic tests for NDD cases in developing countries. In 2010, Miller et al. demonstrated the utility of CMA as a first-tier clinical diagnostic test to enable early diagnosis of individuals with NDDs [24].

In Bangladesh, autism is referred as a great economic burden indicating a significant health problem. A community based study reported the increase in incidence of ASD from 0.2 (2005) to 0.84 (2009) per 1000 children in Bangladesh [25, 26]. In a well characterized NDD suspected cohort, a gold standard observational assessment tool ADOS-2 [27] confirmed 73.85% (209/283) ASD cases rest 26.15% (74/283) are broader NDD cases [28]. The prevalence of NDDs including cerebral palsy, developmental delay and ASD among rural community is 5.6/1000, 2.6/1000 and 0.75/1000etc [29]. Here, the diagnosis of NDDs is mostly done by clinical conditions of the patients and psychological assessment tools like (DSM-IV & ADOS/ADOS2) [28, 30]. Due to overlapping and complex clinical presentation of NDDs, it is difficult to confirm diagnosis by psychological assessment. Therefore, early diagnosis of NDD cases in children may lead to better outcomes through expeditious educational planning and therapeutics [31]. In Bangladesh, the genetic cause of breast cancer, intellectual disability and rare diseases were uncovered by whole genome sequencing [32], whole exome sequencing [33], targeted sequencing [34, 35], whole genome microarray [36] and quantitative PCR [37]. But there is no comprehensive genetic study with large NDD cohort. Our study analyzed a cohort of 212 NDD patients of Bangladesh that underwent microarray testing to identify copy number variations (CNVs) from 2017 to 2020 for diagnostic purpose. To our knowledge, this is a first cohort of NDD patients reporting a significant number of clinically relevant variants and genes from Bangladeshi population.

Methods And Materials

Cohort description

The cohort comprised of 212 neurodevelopmental disorder patients with autism spectrum disorders (ASDs), developmental delay (DD), seizures/epilepsy, intellectual disability (ID), syndromic features, psychiatric or behavioral issues, hypotonia, speech and language disability, attention deficit hyperactivity...
disorder (ADHD). Almost all the patients have more than one phenotype. The age of the patients ranged from 9 days to 31 years with 54.25% (115/212) categorized in the range of 1-5 years while 3.8% (8/212) patients were of less than 1 year of age (Table 1). Around 68.40% (145/212) of the cohort were males. Of the 212 patients, ten (7 male and 3 female) were not included in the statistical analysis of age, weight, OFC and BMI as it was not possible to obtain corresponding data (Table 1 and Table 2). Moreover, a subset of patients (n=95) were evaluated for autism spectrum disorders by ADOS-2 method. Among these 95 participants (70 male, 25 female), 71 meet the criteria of autism positive (51 male, 20 female) and the remaining 24 were classified as autism negative. Besides the 71, there were more 24 autism patients who were diagnosed by different assessment tool DSM-V (n=14) and ADOS (n=10). Therefore, in total there are 95 ASD patients (44.81% =95/212) diagnosed by different types of psychological assessment tools, DSM-V (n =14), ADOS (n=10) and ADOS-2 (n=71), with the male to female ratio of 2.65:1 (Table 2) and age ranging from 1.5 years to 19 years. We have also collected occipito frontal head circumference, measured in centimeters using a non-stretchable plastic tape and body weight measured in kilograms using a calibrated weight machine.

**Ethics statement**

The study protocol was approved by Institutional Review Board of Holy Family Red Crescent Medical College and Hospital. Before participant enrolment, written informed consent including the use of peripheral blood and clinical data for research use and publication was obtained from the parents.

**ADOS-2**

Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) [27] 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 for diagnosing ASD. ADOS-2 has five modules and each module offers standard activities designed to elicit behaviors that are relevant to the diagnosis of ASD at different chronological ages and language abilities. For the present study, autism characteristics were measured using 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 quantification**

DNA was extracted from peripheral blood sample using ReliaPrep™ Blood gDNA isolation kit (Promega, USA) followed by the protocols provided by the manufacturer. Extracted DNA samples were checked for the quality using Nanodrop Spectrophotometer. Then samples were electrophoresed on agarose gels and samples with intact genomic DNA showing no smearing on agarose gel were selected for experiment. Intact genomic DNA was diluted to 50 ng/µL concentration based on Quant-ITPicogreen (Invitrogen) quantitation. The whole-genome amplification process requires 200ng of input gDNA.

**Whole-genome microarray**

We conducted genome-wide microarray to identify chromosomal abnormalities (deletion/duplication/translocation and rearrangements) and investigated the changes in fluorescence intensities between the test specimen and the controls. Illumina Global Screening Array-24+ v1.0 was used applying illumine SNP genotyping technology and Illumina cnv Partition 3.2.1 plug-in of GenomeStudio was used to detect chromosomal abnormalities. This microarray uses 642,824 probes spread across the genome to detect genetic abnormalities (includes >60 loci in DECIPHER database reported for neurodevelopmental disorders) greater than 30kb and also targets sub-telomeric regions that are vulnerable to chromosomal abnormalities. We have used rigorous multiple algorithmic techniques (MathLab 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 (9689 samples) from analysis and only analyzed rare CNVs to infer their contribution to human diseases. Digestion, ligation, PCR, labeling, hybridization, and scanning were performed following the standard protocols.

**Frequency determination**

We have used commercially available tool GenomeArc (GenomeArc Inc.) annotation software that integrates clinical, genomics and OMICs data. To determine the frequency >50% reciprocal overlap within the same chromosome was considered and we have used 9689 unrelated population control samples from European ancestry [38]. 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) [39]and the Health, Aging, and Body Composition (HABC) [40]; 2,988 control samples assayed by Illumina Omni 2.5 Mfrom the Collaborative Genetic Study of Nicotine Dependence (COGEND) [41] and Cooperative Health Research in the Region of Augsburg KORA projects [42]; 2,357 control samples assayed by Affymetrix 6.0 from the OttawaHeart Institute [43] and the PopGen project [44].

**CNV classification criteria**

To classify and assess the clinical relevance for a particular CNV, at first, all common CNVs (frequency >1%) found in the control population samples (9689 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), size, location, gene content as well as the 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 [45].

**Droplet Digital PCR**

Copy Number Assays were performed using Droplet Digital PCR (ddPCR) System (Bio-Rad Laboratories, Inc.). GeneAssist™ Copy Number Assay Workflow Builder (ThermoFischer, USA) 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 22µl reaction mix was prepared, containing 3.5µl of template DNA (20ng/µl) without restriction digestion, 10 µl of 2X ddPCRsupermix for probes (No UDP) (Bio-RAD Laboratories Inc.), 1µl each 20X TaqMan Target Probe (FAM) and 20X TaqMan Reference probe (VIC) (Applied Biosystems, USA) and 6.5µl of RNase-/Dnase Free water. All reactions were prepared in triplicates with one negative control. The reaction mixtures were partitioned using QX200 Droplet Generator™ and then transferred to a 96 well plate and amplified using C1000 Touch thermal cycler, as per the manufacturer's protocol, then the samples were read by QX200 Droplet Reader™. Data acquisition and analysis were performed using QuantaSoft Version 1.7.4.0917. Poisson algorithm was used to determine the concentrations of the targets as copies/µl.

**Candidate gene and pathway analysis**

Gene set enrichment analysis (GSEA) [46] was performed using the gene overlap package of R and cytoscape [47] was used to analyze the pathways interrupted by pathogenic CNVs. The enriched set of genes along with their GMT file (consisting of KEGG and GO[48] pathways) were uploaded in the Cytoscape. A pathway was considered to be significant if the false discovery rate (FDR) and p-value cut off was <0.01 and 0.001, respectively. Then, the network was built using the enrichment map and the auto-annotate cytoscape application. P-values are denoted using color gradient (low p values with darker colors). Further, “critical-exon” method [38, 49] was applied to identify candidate genes within the pathogenic and VOUS breakpoints.

**Statistical analysis**

Welch t-test and one way ANOVA were used to test the physical and phenotypic characteristics difference based on sex, autism diagnosis, and mutation. Significance level of 0.05 was determined to test differences. The analysis was done using statistical analysis in social science (SPSS) version 23. For qualitative data we applied Fisher's exact test.

**Results**

The cohort comprised of 212 neurodevelopmental disorder patients with male to female ratio of 2.2:1.0 (Table 1). Among 212, 202 were included for descriptive statistical analysis and showed significant differences between the male and female groups with respect to weight (p=0.001), head circumference (p=0.0002) and BMI (p=0.001) (Table 1 and Figure 1). Patients carrying rare CNVs (pathogenic/VOUS) and patients carrying no rare CNVs are defined as CNV positive and CNV negative group respectively. Significant differences with respect to head circumference (p=0.0001), weight (p=0.002) and BMI (p=0.007) were also found in the male CNV positive group compared to female CNV positive group (Table 1, Figure S1) while no significant differences were observed in terms of head circumference and weight in the CNV negative groups (Table S2). It was observed that females are the substantial carriers of pathogenic CNV than males (OR=4.2; p=0.00077) (Figure 3d).

**Table 1:** Physical and phenotypic characteristics difference based on sex and CNV types
In this cohort, 95 patients were evaluated for autism spectrum disorders by ADOS-2 method and 71 met the ADOS-2 cut-off criteria of autism positive (51 male, 20 female, male/female ratio 2.55:1) while the remaining 24 were classified as autism negative (non-spectrum). The entire cohort (p=0.0096) and subset of ASD positive cases (p=0.019) show head circumference of males was significantly greater than females (Table 2, Figure 2a and 2b) while no significant differences were observed among male and female in terms of age, weight, height, BMI, social affect score, and ADOS-2 total score. We also observed that patients carrying duplication CNV showed severe social communication deficit (p=0.014) and overall ASD symptoms severity (p=0.026) compared to the CNV negative group (Table 2, Figure 2d and 2e). 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) (Figure 2c). The detailed statistical analysis of the subset was embedded in Table S3.

Table 2: Physical and phenotypic characteristics difference based on sex, autism and CNV types

| General description of full cohort | Male | Female | Total | Significance of group difference (p) |
|-----------------------------------|------|--------|-------|-------------------------------------|
| Number (Frequency in %)           | 145 (68.40%) | 67 (31.60%) | 212 | 0.0001 |
| Mean Age in years (age range)     | 5.828 (6M to 30.8Y) | 4.794 (8M to 18.9Y) | 212 | 0.077 |
| Male to female ratio              | 2.16:1 | | | |
| 1 to 5 years old patients         | 75 (65.22%) | 40 (34.78%) | 115 | 0.0019 |
| < 1 year old patients             | 6 (75.00%) | 2 (25.00%) | 8 | 0.2367 |

| Physical characteristics vs gender | Physical characteristics vs male/female CNV status |
|-----------------------------------|-----------------------------------------------|
| **Physical characteristics**      | **Male (138) Male CNV positive (78)**          |
| **Female (64) Female CNV positive (41)** | Mean | Mean | t (p) | Mean | Mean | t (p) | Mean | Mean |
| OFC (cm)                          | 50.098 | 47.891 | 0.0002* | 50.558 | 49.500 | 47.732 | 48.174 |
| Weight (kg)                       | 22.088 | 16.547 | 0.001*  | 23.133 | 20.730 | 16.354 | 16.891 |
| Height (m)                        | 1.066 | 1.022 | 0.214  | 1.066 | 1.065 | 0.994 | 1.073 |
| BMI (kg/m²)                       | 18.809 | 15.304 | 0.001*  | 20.308 | 16.860 | 15.601 | 14.775 |

| **Physical characteristics vs male/female CNV status** |
|-----------------------------------------------|
| Phenotype vs M/F CNV status                  | Male CNV pos (78) | Female CNV pos (41) | Male CNV neg (60) | Female CNV negative (23) | Significance of difference | Male CNV positive (78) | Male CNV negative (60) |
|-----------------------------------------------|-------------------|----------------------|-------------------|--------------------|--------------------------|------------------------|------------------------|
| Mean | Mean | t (p) | Mean | Mean | t (p) | Mean | Mean | t (p) | Mean | Mean |
| OFC (cm) | 50.558 | 47.732 | 0.0001*  | 49.500 | 48.174 | 1.276 (0.211) | 50.558 | 49.500 |
| Weight(kg) | 23.133 | 16.354 | 0.002*  | 20.730 | 16.891 | 1.713 (0.09) | 23.133 | 20.730 |
| Height (m) | 1.066 | 0.994 | 0.087  | 1.065 | 1.073 | 0.134 (0.894) | 1.066 | 1.065 |
| BMI (kg/m²) | 20.308 | 15.601 | 0.007*  | 16.860 | 14.775 | 2.162 (0.038) | 20.308 | 16.860 |

* p values are Bonferroni corrected
General description of autism cohort

|                                   | Male                      | Female                   | Significance of group difference (p) |
|-----------------------------------|---------------------------|--------------------------|--------------------------------------|
| Patient's number and frequency (N=95) | 69 (72.63%)               | 26 (27.37%)              | 0.0001                               |
| Mean Age in months                | 65.343                    | 58.92                    | 0.42                                 |
| Autism male to female ratio       |                           |                          | 2.55:1                               |

Physical and phenotypic characteristics vs gender

|                                   | Male (70) | Female (25) | Significance of difference | Male ASD positive (51) | Female ASD positive (20) | Significance of difference |
|-----------------------------------|-----------|-------------|---------------------------|------------------------|-------------------------|---------------------------|
| Mean OFC (cm)                     | 50.957    | 49.600      | 0.0096                    | 51.216                 | 49.800                  | 0.019                     |
| Weight (kg)                       | 21.964    | 18.600      | 0.106                     | 23.118                 | 18.850                  | 0.100                     |
| Height (m)                        | 1.094     | 1.105       | 0.842                     | 1.109                  | 1.116                   | 0.914                     |
| BMI (kg/m²)                       | 17.122    | 15.640      | 0.094                     | 17.247                 | 15.734                  | 0.168                     |
| Social affect score               | 11.229    | 11.920      | 0.462                     | 13.392                 | 13.100                  | 0.728                     |
| ADOS 2 total score                | 13.000    | 13.720      | 0.539                     | 15.745                 | 15.300                  | 0.662                     |

Physical and phenotypic characteristics vs ASD male-female

|                                   | Del CNV (27) | Dup CNV (24) | Significance of difference | Dup CNV (24) | CNV negative (37) | Significance of difference |
|-----------------------------------|--------------|--------------|---------------------------|--------------|-------------------|---------------------------|
| Mean OFC (cm)                     | 50.333       | 51.250       | 0.264                     | 51.250       | 50.486            | 0.310                     |
| Weight (kg)                       | 22.241       | 24.667       | 0.530                     | 24.667       | 19.108            | 0.046                     |
| Height (m)                        | 1.082        | 1.154        | 0.204                     | 1.154        | 1.095             | 0.258                     |
| BMI (kg/m²)                       | 17.216       | 17.671       | 0.682                     | 17.671       | 15.927            | 0.071                     |
| Social affect score               | 10.926       | 13.208       | 0.041                     | 13.208       | 10.541            | 0.014                     |
| ADOS 2 total score                | 12.667       | 15.333       | 0.051                     | 15.333       | 12.216            | 0.026                     |

Phenotype and physical characteristics vs CNV types

Our microarray analysis from the cohort identified 1053 CNVs (excluding <30kb deletion and <50kb duplication) that includes 395 duplications and 658 deletions that were interpreted and classified according to the ACMG guideline [45] into benign, pathogenic and variants of uncertain clinical significance (VOUS) (Figure 3b). We have also determined the frequency of all 1053 CNVs within 212 Bangladeshi NDD patients and 9689 control CNVs from Ontario population with European ancestry. Using the frequency distribution of both rare and common CNVs of NDD patients of Bangladesh, circos map was constructed[50] (Figure 4). Of all samples assayed, 12.26% (26/212) (Table 3, Figure 3a) and 47.17% patients carried pathogenic and variant of uncertain significance (VOUS) CNVs (Figure 3a). Recurrent deletions and duplications were identified in the 15q11.2q13.1 and 21q11.1q22.3 loci (Table 3). All the CNVs in these regions were further confirmed by ddPCR (Table S4 and Figure S2). In this cohort, 2.83% (6/212) and 20.28% (43/212) patients were carrying pathogenic and VOUS subtelomeric CNVs respectively (Table 3 and Table S1). Of the pathogenic 26 samples, one patient was carrying double terminal deletion impacting chromosome 18 (Table 3) and 11 along with a pathogenic CNV also presented VOUS (Table S5). The 27 pathogenic CNVs comprised of 16 deletions, and 11 duplications (Table 3 and Figure 3C). The average length of deletion and duplication are 5365.26kb and 17451.72kb respectively and the highest frequency group for pathogenic CNV deletion and duplication are 30-2000kb and >20000kb respectively (Figure 3e). To exclude false CNV calls, we have randomly chosen 9 pathogenic CNV for ddPCR validation and yielded (8/9) 88.89% validation rate (Table S4 and Figure S2).

Larger numbers of genes were found to be disrupted in pathogenic deletion (n=1846) and duplication CNVs (n=2925) compared to VOUS deletions (n=490) and duplications (n=991) (Figure 3C). The average length of VOUS deletion and duplication is 129.17kb and 476.76kb respectively and the highest frequency group for VOUS CNV deletion and duplication are 30-100kb and >500kb respectively (Figure 3f). In this cohort, we have also found 1591 loss of heterozygosity (LOH) variants in almost all samples of the cohort (210/212). Of the 1591, 62.10% (988/1591) and 37.90% (603/1591) LOH CNVs are rare and common, respectively. The distribution of frequency has been shown in Figure 4a and 4b.

Our custom pathway enrichment (comprised of 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:00060548)" and "vesicle-mediated transport in synapse (GO:0099003)" pathways to be highly significant (FDR P <6.53 x 10^{-7}), (FDR P <7.2 x 10^{-6}) and (FDR P <7.1 x 10^{-5}) after correction for multiple tests (Figure 5a). Pathway enrichment on the genes impacted by pathogenic duplications identified "interspecies interaction between organisms (GO:0044419), "dependent protein catabolic process (GO:0030163)", "regulation of neuron death (GO:1901214)" and synaptic signaling...
(GO:0099536)* pathways to be highly significant (FDR P <1.51 × 10^{-5}), (FDR P <1.85 × 10^{-5}), (FDR P < 5.34 × 10^{-7}) and (FDR P <1.72× 10^{-4}) after correction for multiple tests (Figure 5b).

Analysis of both pathogenic and VOUS CNVs for critical exon genes (CEGs) yielded 153 unique CEGs in pathogenic CNVs and 31 in VOUS, including 3 genes found in both pathogenic CNVs and VOUS alike. On average, there were significantly more CEGs per pathogenic variant (7.5) compared with VOUS (0.19) \((p=0.0002)\) (Table 4). In "critical-exon" analysis, we have found total 24 focal CNVs of <1Mb that disrupt 22 unique CEGs in 22 NDD patients. We have excluded 4 CEG genes; *PTPRD*, *RBFOX1*, *PCDH9* and *LRRC4C* from further analysis as only intronic region of the genes were disrupted by corresponding CNVs.

After excluding these genes, rest 18 genes; *RAB11FIP3, ASTN2, NRXN1, PRPF8, PITPNA, FGFR2, TSC2, PKD1, PSMC3, STK11, APC, ADAR, MAP4, ABCA2, PRKDC, PRKCE, ARID4B, PLCB1*, were taken for further analysis (Table S6 and Figure 6). Of the 18, we have found *PSMC3* gene as a possible candidate gene for ASD. We have also found a patient (#106), carrying 2.01Mb and 891Kb duplication CNVs, had most similar phenotypes to the *KMT2B* related disorder patients previously described [51–56] (Figure 7 and Table S8).

**Table 4:** 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 CEG/CNV   | 7.50       | 0.19 |

\* Welch t test

**Discussion**

In this study, we have found significantly greater head circumference in the male patients compared to female patients in the overall NDD and autism positive patients. This is a replication of previous studies in other ethnically diverse population that also shown similar association with abnormal acceleration of head growth among children of ASD compared to neurotypical children [57–59]. This observation hints that aberrant brain cell proliferation may be a key neurobiological mechanism in the disorder [60]. Some studies also suggest that certain mutations underlying neurodevelopmental disorders may also lead to changes in brain volume, microcephaly or macrocephaly [61, 62]. Significant difference in head circumference was not observed between CNV positive and negative group as well as CNV deletion and duplication group (Table S2 and S3). But larger head circumference and weight were found in male carrying rare CNV compared to female carrying rare CNV (Table 1, Figure 1).

This is the first Bangladeshi NDD cohort conducted whole genome microarray analysis using GenomeArc annotation tool and identified a diagnostic yield of 12.26% (26/212) that is in strong agreement with other ethnically diverse neurodevelopmental disorder cohorts [24, 38, 63]. Of the 12.26% (26/212), 5 patients were carrying different lengths of deletion and duplication CNVs in the 15q11.2q13.1 region of chromosome 15 (Table 3 and Figure S3). The proximal long arm 'q' of human chromosome 15q11-q13 harbors a cluster of imprinted genes, which expressions are controlled by the imprinting centre (IC) [64]. 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) [65, 66]. We have also found 3 patients carrying extra copy of chromosome 21q (dup21q) that is associated with Down Syndrome (DS#190685). Moreover, there were 4 other patients carrying 3.5Mb deletion in the 17p11.2 (#40), 1.4Mb deletion in the 7q11.23 (#107), 1.69Mb deletion in the 17q12 (#59) and 1.9Mb duplication in the 22q11.22q11.23 (#6) chromosomal regions (Table 3) that overlapped with the critical regions implicated with some well characterized syndrome [66], 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.
Table 3
Details of pathogenic CNVs found in the cohort

| ID | Sex | Type | Chromosomal band and position | Size (Mb) | Number of affected genes | Critical exon genes | GenomeArc analysis | Clinical features |
|----|-----|------|-------------------------------|----------|------------------------|---------------------|-------------------|-------------------|
| 1  | M   | Del  | chr4:35,031-4,878,160         | 4.843    | 155                    | CTBP1, PCGF3, FAM193A, HTT | Wolf Hirschhorn Syndrome; PIGG, UVSSA3, IDUA, FGFR3, LRPAP1, SH3BP2, MSX1 | Autism, delay (w: behavior, intellect, frequent Dysmorphic features) |
| 19 | M   | Del  | chr15:22,838,642-28,314,382   | 5.476    | 170                    | GABRB3, CYFIP1, UBE3A, HERC2, GABRA5 | PWS/AS; UBE3A, GABRB3, MAGEL2 | Global delay, ca sponta body, Synencephaly |
| 20 | F   | Del  | chr20:8,117,650-8,593,664     | 0.476    | 1                      | PLCB1               | PLCB1             | Autism, temper tsp restless |
| 35 | F   | Del  | chr15:22,777,709-28,736,935  | 5.959    | 181                    | GABRB3, CYFIP1, UBE3A, HERC2, GABRA5 | PWS/AS; UBE3A,GABRB3;MAGEL2 | Hypotonia |
| 40 | F   | Del  | chr17:16,858,825-20,387,706  | 3.529    | 136                    | LLGL1, SREBF1, COP5A, AKAP10 | Smith Magenis Syndrome; Potocki-Lupski syndrome; (17p11.2 duplication syndrome); TNFRSF13B, RAI1, B9D1, ALDH3A2 | Behavior speech development |
| 53 | F   | Del  | chr1:168,503,684-191,056,769 | 22.553   | 360                    | TPR, MNMAT2, DHX9, RGL1, SMG7, PRRC2C, GLUL, CACNA1E, IVNS1ABP, RNF2, ASTN1, CACYBP, ATP1B1, VAMP4, TNR, XPR1 | GORAB, TSEN15, LHX4, ACBD6, GLUL, PRRX1, NPHS2, DARS2, MYOC | Global delay, Ar act lactate level normal |
| 59 | M   | Del  | chr17:36,201,251-37,889,808  | 1.689    | 51                     | GGNBP2, ACACA        | RCAD (renal cysts and diabetes); HNF1B, PIGW | Global delay (delay and speech delay restlessness pneumo of age) |
| 86 | M   | Del  | chr18:45,458-40,276,243      | 2.231    | 38                     | COLEC12             | null              | Develop complet clef palsy |
| 107 | F  | Del  | chr7:73,294,506-74,715,504   | 1.421    | 40                     | BAZ1B, LIMK1, STX1A, EIF4H, CLIP2 | Williams Beuren Syndrome (WBS); 7q11.23 duplication syndrome, ELN | Developor (unable t support) problem, moveme |
| 144 | F  | Del  | chr2:233,505,099-242,065,217 | 8.560    | 219                    | KIF1A, ATG4B, AGAP1, UBE2F, HDLBP | 2q37 monosomy, KIF1A, COL6A3, NDUFA10, D2HGDH, CAPN10, TWIST2, HDAC4, UGT1A1, AGXT | Syndrome develop yet cann neck con sleeping night, lac response days eye Breathing and nee transfus Pneumo 28 days. |

*a* = Each CNV validated by droplet digital PCR (ddPCR) with TaqMan assay; **b** = subtelomeric CNV; CNV, copy-number variation; CP, cerebral palsy; del, deletion; d, duplication.
| ID | Sex | Type | Chromosomal band and position | Size (Mb) | Number of affected genes | Critical exon genes | GenomeArc analysis | Clinical notes |
|----|-----|------|-------------------------------|----------|--------------------------|---------------------|--------------------|-----------------|
| 159 | M   | Del  | 2q24.1 chr2:154,077,277-158,596,423 | 4.519 | 70 | KCNJ3, NR4A2, ACVR1 | ACVR1 | Intellectual ADHD, self-injury, epilepsy; suffered Facial atrophy |
| 178 | M   | Del  | 15q11.2q13.1 chr15:22,600,363-28,760,485 | 6.160 | 188 | GABRB3, CYFIP1, UBE3A, HERC2, GABRA5 | PWS/AS; UBE3A, GABRB3, MAGEL2 | Spastic Quadriplegia; developmental delay; speech delay; CP; cerebral palsy; seizure; speech delay; atrophy; dilatation; facial atrophy |
| 182 | F   | Del  | 2q24.3 chr2:165,992,388-166,079,996 | 0.087 | 1 | null | SCN1A | Global developmental delay; speech delay; CP; cerebral palsy; seizure; speech delay; atrophy; dilatation; facial atrophy |
| 184 | M   | Del  | 17q12 chr17:34,815,552-36,249,430 | 1.434 | 76 | AP2B1 | PEX12 | Low level epilepsy; foot, tem speech delay; smiling; poor respiration; peer relations; difficulty with food, crying, drooling; elder brother had delay like his; normal pregnancy appearance |
| 5   | F   | Dup  | 3p26.3p26.2 chr3:2,299,060-3,322,758 | 1.024 | 5 | null | CRBN | Expressive disorder |
| 77  | F   | Dup  | 9p24.3p12 chr9:48,828-39,297,860 | 39.249 | 619 | PTPRD, MLLT3, SH3GL2, CDC37L1, SMARCA2, NOL6, KHLH9, NFIB, PSIP1, SMU1, ELAVL2, UBAP1, RUSC2, UBE2R2, CLTA, CNTFR, VCP, NPR2, TESK1, TLN1, RNF38 | PIGO, IL11RA, NPR2, TPM2, GBA2, GALT, FANCQ, EXOSC3, RMRP, MPD2, GLIS3, DDX58, VLDLR, KANK1, APTX, FREM1, DOCK8, TEK, SMARCA2, GLIS3, PLAA, TYRP1, GLDC | Developmental delay; speech delay; intellectual disorder; anorexia |

* Each CNV validated by droplet digital PCR (ddPCR) with TaqMan assay; **subtelomeric CNV; CNV, copy-number variation; CP, cerebral palsy; del, deletion;
| ID | Sex | Type | Chromosomal band and position | Size (Mb) | Number of affected genes | Critical exon genes | GenomeArc analysis | Clinical description |
|----|-----|------|-------------------------------|----------|--------------------------|---------------------|---------------------|---------------------|
| 14 | M   | Dup<sup>a</sup> | 11q23.2q24.2 chr11:114,071,597-125,513,912 | 11.442 | 307 | **BCL9L, BACE1, IFT46, DDX6, C2CD2L, GRAMD1B, UBE4A, CADM1, HSPAB, HYOU1, ARHGGEF12, SIK3, ARCN1, PKNOX2, PAFAH1B2** | **ZBTB16, NECTIN1, ARCN1, ROBO3, DPAGT1, SC5D, CLMP, CBL, KMT2A, MFRP** | Develop speech delay, eye contact, seizure, developmental delay, facial dysmorphism, callosum appearance |
| 15 | F   | Dup<sup>b</sup> | 14q32.11q32.33 chr14:40,576,836-106,816,816 | 16.240 | 626 | **DICER1, ITPK1, TTC7B, TRAF3, PAPOLA, TECPR2, AKT1, DYN1H1, EVL, WARS, YY1, CDC42BPB, BCL11B, PPP2R5C, INF2, EIF5, PACS2, MTA1** | **TRIP11, UBR7, SLC24A4, CCDC88C, APOPT1, TECPR2, VRK1, DYNC1H1, AKT1, YY1, PACS2** | High level of seizures, lack of speech development, developmental delay |
| 2  | M   | Dup<sup>a</sup> | 15q11.2q13.1 chr15:23,443,797-28,289,373 | 4.846 | 147 | **GABRB3, UBE3A, HERC2, GABRA5** | **PWS/AS; UBE3A, GABRB3, MAGEL2** | Autism, intellectual disability |
| 101 | F  | Dup<sup>a</sup> | 15q11.2q13.1 chr15:23,370,969-28,371,148 | 5.000 | 157 | **GABRB3, UBE3A, HERC2, GABRA5** | **PWS/AS; UBE3A, GABRB3, MAGEL2** | Autism, intellectual disability, speech delay, anxiety, hyperactivity, restless behavior, sleep problems |
| 79 | M   | Dup<sup>b</sup> | 17p13.3p12 chr17:151,597-11,392,231 | 11.241 | 414 | **PRPF8, PITPNAA, ABR, YWHAE, MNT, SMG6, TNFSF12, SNTN1, RAP1GAP2, CHD3, CAMTA2, KDM6B, DLG4, NGLN2, ANKFY1, PHF23, FXR2, POLR2A, MINK1, ZBTB4, CTDNEP1, RABEP1, TNFSF12-TNFSF13, EIF4A1, PAFAH1B1, PIP4K3M3, NEURL4, RPL26, MYH10, NDEL1** | **Miller Dieker syndrome (MDS); INPP5K, KANSL1, WDR91, BHLHA9, WRAP53, DLG4, TRPV3, ACDAVL, ASPA, C1QB, MPDUD1, KDM6B, SLCT3A5, CTNS, AIPL1, AIPL1, MYH10, NDEL1** | Speech delay, high level of seizures, developmental delay, hyperactivity, restless behavior, history of gastrointestinal problems |
| 134 | F  | Dup<sup>a</sup> | 21q11.1q22.3 chr21:12,987,574-46,679,698 | 33.692 | 683 | **NRIP1, ETS2, COL6A1, ITSN1, PDXK, CCT8, SON, AGPAT3, NCAM2, BRWD1, DYRK1A, TIAM1** | **Down Syndrome; C21orf59, SON, HLC5, KCN1, DYSK1A, C21orf2, COL18A1, PCNT, RSPH1, RIPK4, CBS, FTO, SIK1, COL18A1, CRYAA, AIRE, CTSB** | Global development delay, feeding problems, no toilet training, speech delay, hyperactivity, dysmorphic features (flat nose, and eyes) |
| 136 | F  | Dup<sup>a</sup> | 21q11.1q22.3 chr21:12,987,574-46,679,698 | 33.692 | 683 | **NRIP1, ETS2, COL6A1, ITSN1, PDXK, CCT8, SON, AGPAT3, NCAM2, BRWD1, DYRK1A, TIAM1** | **Down Syndrome; C21orf59, SON, HLC5, KCN1, DYSK1A, C21orf2, COL18A1, PCNT, RSPH1, RIPK4, CBS, FTO, SIK1, COL18A1, CRYAA, AIRE, CTSB** | Developmental delay, speech delay, hyperactivity, contact problems, behavior problems |

<sup>a</sup>= Each CNV validated by droplet digital PCR (ddPCR) with TaqMan assay; <sup>b</sup>= subtelomeric CNV; CNV, copy-number variation; CP, cerebral palsy; del, deletion; d
2.83% (6/212) pathogenic CNVs in this cohort found to be in the subtelomeric region (Table 3) that is within the range of previously reported (2.4–4.4%) diagnostic rate in patients with DD/DD and congenital anomaly [67, 68]. Of the six, we have found 4.8Mb deletion in the 4p16.3p16.2 (#1), 11.2Mb duplication in the 17p13.3p12 (#47) and 8.6Mb deletion in the 2q37.1q37.3 (#144) subtelomeric chromosomal regions that overlapped with the critical regions associated with the three well characterized syndrome [66], Wolf-Hirschhorn syndrome (WHS#194190), Miller-Dieker syndrome (MDS#247200) and 2q37 monosomy (#600430), respectively. Subtelomeric regions are particularly prone to genomic instability due to the repeat sequences found in these areas. Due to the high gene density, these subtelomeric loci reported to have a variety of phenotypic abnormalities particularly intellectual disability (ID)/mental retardation [69–73]. Rahman et al. [69] identified two siblings affected with ID who are both impacted by a large terminal duplication (18.82 Mb) and a deletion (3.90 Mb) CNVs. Rest 8 patients were carrying 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 (Table 3).

Identifying overlapping genes and pathways across disorders is critical to improve 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:0006990)"; "dependent protein catabolic process (GO:0009538)"; "regulation of neuron death (GO:0030163)" and synaptic signaling (GO:0099536) pathways to be highly significant (Figure 5A and 5B). Aberrations in autophagy (a major cellular catabolic process) related signaling and mutations in autophagy related genes have been implicated in several neurodevelopmental disorders including autism, Tuberous sclerosis, Fragile X syndrome and Neurofibromatosis type 1 [74, 75]. Loss of function in the ubiquitin ligase gene HERC2 has been associated with severe neurodevelopmental phenotype [75]. Mutations in presynaptic genes have been linked to various neurodevelopmental disorders including autism, intellectual disability and epilepsy [76] and synaptic signaling has been identified as one of the principle molecular pathways affected in neurodevelopmental disorders [77]. Perturbations in the apoptotic signaling pathway have also been identified in various NDDs including autism, Fragile X syndrome, and schizophrenia [78, 79].

Analysis of "critical-exon"method[38,49] to identify constraint candidate genes have found high number of CEGs within pathogenic variant (7.5) compared with VOUS (0.19) (p=0.0002). More CEGs per pathogenic CNV was also previously reported[80] and reflective of length bias and gene density of pathogenic variants. CEG analysis of short focal CNVs identified 18 unique CEGs that are highly expressed in brain and have low burden of non-synonymous variants. The role of these genes within the context of neurodevelopmental disorders and our cohort, we have conducted comprehensive literature search. For example, in our cohort, one autism patient (#20) was carrying a 476 kb pathogenic deletion disrupting PLCB1 gene. Girirajan et al. found an enrichment of microdeletions and duplications involving the PLCB1 gene in individuals with autism [81]. Rest 17 patients were carrying VOUS. Of the 17, we were trying to find out a common genetic breakpoint shared by multiple patients with same clinical condition to identify candidate gene for the disrupted locus. In this cohort, we have found 2 autism patients (#119 and #121) that harbored common deletion breakpoint disrupting 4 genes SLC39A13, PSMC3, SPI1 and RAPSN including one critical exon gene, PSMC3. Autosomal recessive mutations in the SLC39A13 gene is associated with a well defined disease Ehlers-Danlos syndrome, spondylopathyloplastic type, 3 (OMIM# 612350). Mutation in the SPI1 gene was previously reported in acute lymphoblastic leukemia [82]. Autosomal recessive mutation in another gene, RAPSN, 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 rest one gene of the common breakpoint is PSMC3. PSMC3 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 [83]. This is suggestive of PSMC3 plays an essential role in the ubiquitin–proteasome system (UPS) that includes morphogenesis, dendritic spine structure, synaptic activity, and the regulation of synaptic strength in neurons [84–87]. Recently, Ariane Kröll-Hermi et al. demonstrate that homozygous single nucleotide variant in the PSMC3 cause neurosensory syndrome combining deafness, cataract, autism/neurodevelopmental delay due to proteotoxic stress [88]. Further, we have found one 217kb deletion [89] (ClinVar_VCV000151005) and one 173kb duplication [90] (Decipher_412053) CNVs in two previously reported NDD patients in which PSMC3 gene was also disrupted (Figure 6c, 6d and Table S7). Although our patients did not show syndromic appearance at the age of 1.6 and 3.1 years except ASD we...
hypothesized on the basis of previous studies in UPS [84–87] and decipher [66] data that heterozygous mutation in the PSMC3 (size ～7.7kb) gene might be associated with NDDs without causing neurosenosynrome syndrome.

In this study, we have found a 20 year old girl (#106) born in a consanguineous marriage with healthy parents and was delivered at preterm after an eventful pregnancy (IUGR). She had the history of delayed development. She had intellectual disability with dysmorphic features of elongated face, long fingers (Table S1). Her mental and physical conditions were progressively worsening from the age of 15 years. At the age of 18 years, she developed dyskinesia and swallowing difficulty. From this age, she was unable to walk independently or talk clearly. Her MRI finding was normal at this age. Analyzing the overlapping region of fourteen previously reported deletions (length 0.19 to 4.91 Mb) [55, 90–94] and three duplications (length 3.31 to 12.63 Mb) [90], our patient (#106) contains a shorter 2.01Mb duplication with a 38.4kb (chr19:35,700,296-35,738,700) common overlapping region among the CNVs disrupting 2 genes, ZBTB32 and KMT2B (Figure 7 and Table S8). Further, CEG and GenomeArc analysis also identified UBA2, USF2, SCN1B, KMT2B, COX6B1, LG4 and ZNF599. SCN1B is the most interesting gene known to be associated with Atrial fibrillation, familial, 13 (OMIM#615377), Brugada syndrome 5 (OMIM#612838), developmental and epileptic encephalopathy 52 (OMIM#617350) and epilepsy with febrile seizures plus, type 1 (OMIM#604233) when get abnormal due to pathogenic mutations. Our patients had no history of seizures or epilepsy or any cardiac problems that indicate that SCN1B duplication might not be associated with our patient's clinical condition. KMT2B is another interesting gene within this breakpoint and pathogenic mutations are associated with Dystonia 28, childhood-onset (OMIM#617284) which core phenotypes are described as limb-onset childhood dystonia that tends to spread progressively, resulting in generalized dystonia with cranioocular involvement. Co-occurring signs such as distinct facial dysmorphism and intellectual disability are most common [54, 55]. There is a distinct group of KMT2B patients presenting with a neurodevelopmental disorder in the absence of dystonia or related movement disorder [52, 54, 56]. We have also found some Decipher [90] patients carrying duplication CNV containing 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, swallowing problem, walking and talking problem match with the KMT2B related disorder. That's why we hypothesize that KMT2B gene duplication 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 are the primary cause of disease mechanism [54–56] but 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 [54, 56]. A report from China showed that DYT1-KMT2B and KMT2B-related neurodevelopmental disease without dystonia can occur even within the same family [95]. From the current knowledge of KMT2B related studies we hypothesize that patients with KMT2B duplication variants had reduced penetrance for KMT2B related disorders than those with chromosomal deletions and loss of function mutations [54, 56]. 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 patient we have also found another 891Kb VOUS which disrupt the most important gene FTL. Pathogenic mutations in this gene are associated with neurodegeneration with brain iron accumulation 3; NBIA3 (OMIM#606159) or Hyperrferitinemina with or without cataract; hrftc (OMIM#600886). NBIA3 is characterized by progressive iron accumulation in the basal ganglia and other regions of the brain, resulting in extrapyramidal movements, such as Parkinsonism, dystonia and dyskinesia. Age at onset is variable 13 to 63 years [96, 97]. No cavitation of the basal ganglia or any evidence of iron deposition was found in the MRI report of our patient at the age of 18 years (report normal finding). Although our patient has no confirmed pathology findings of NBIA3 or cataract, it is not possible to exclude FTL duplication association with the clinical conditions of our patient due to variable age of onset of the disease.

In our cohort, a series of overlapping rare clinically relevant variants have been identified in multiple patients. For example, a 96.8 Kb deletion was found in two unrelated ASD patients (#8 and #127) at 6p21.33 that disrupts major histocompatibility complex (MHC) class I gene MICA that was not previously associated with broader neurodevelopmental disorder patients. Another three unrelated patients (#13, #15 and #207) carrying around 48Kb duplication at 15q13.3 that disrupt the previously reported [98–100] broader NDD genes CHRNA7 and OTUD7A. We have also found a 381Kb duplication in two unrelated ASD patients (#34 and #104) at 8q21.2 that disrupts Carbonic anhydrases II, CA2 gene associated with the disease osteopetrosis, autosomal recessive 3; OPTB3 (OMIM#259730). We have found two siblings (#42, #43) affected with variable NDD phenotypes and one unrelated patient (#152) carrying 233 kb to 340 kb duplication at 3p26.3 disrupting previously reported [101, 102] broader NDD gene CNTN6 gene. We have found another two unrelated ASD patients carrying around 130Kb duplication CNV disrupting previously reported [103] NDD gene NPHPT1. Three unrelated patients (#57, #159 and #200) carrying 44Kb to 691Kb duplication CNVs with a 236Kb overlapping region disrupting two constraint genes NPY4R and GPRIN2. In addition, we have found two unrelated ASD patients (#65 and #99) carrying 51Kb and 59Kb duplication at 4p16.3 disrupting Alpha-L-iduronidase, IDUA gene.

**Conclusion**

In this paper, we have shown the utility of whole genome microarray as a first tier diagnostic technology for neurodevelopmental disorders patients. Without a proper genetic test, the clinical complexity alone may not enough to identify the cause and often lead to a diagnostic odyssey. The price of microarray is getting cheaper and in near future developing countries will be able to implement such technology within their healthcare setting. To resolve diagnosis of NDD cases, we highly recommend to use whole genome microarray test in developing countries that eventually will lead to precision diagnosis for 10-20% of NDD cases and will enable the detection of novel variants and genes from underrepresented populations.

**Abbreviations**

Neurodevelopmental Disorders (NDD); Chromosomal Microarray Analysis (CMA); Copy number variation (CNV); Autism Diagnostic Observation Schedule-Second Edition (ADOS-2); Variant of Uncertain Significance (VOUS); Critical Exon Gene (CEG); Autism Spectrum Disorder (ASD); Intellectual Disability (ID); Attention Deficit Hyperactive Disorders (ADHD); Developmental Delay (DD); Droplet Digital PCR (ddPCR); Gene set enrichment analysis (GSEA)

**Declarations**

Ethics approval and consent to participate
The study was conducted according to the Declaration of Helsinki and was approved by the Institutional Ethical Review Committee (IERC) of Holy Family Red Crescent Medical College and Hospital, and all samples were collected with written informed consent.

Consent for publication

Written informed consent for publication was obtained from parents or legal guardians for all individuals involved in the study.

Availability of data and materials

Patient's phenotypic data is contained within the supplementary material. Genomic data can be shared for any collaborative research that involves NeuroGen Healthcare. Please request this via the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported, in whole or in part, by internal grant awards from Startup Bangladesh–IDEA project, Genetics and Genomic Medicine Centre, NeuroGen Healthcare; Mohammed Bin Rashid University of Medicine and Health Sciences (MBRU) – College of Medicine; Sandooz Al Watan Research & Development Grant (SWARD-F2018-002); AlMahmeed Collaborative Research Awards (ALM1801, ALM20-0074); Dr. Nasna Nassir was supported by MBRU Post-Doctoral Fellow Award (MBRU-PD-2020-02).

Authors' Contributions

Conceptualization, MU, HA, MMR; Preparation of draft manuscript by MU, HA, MMR, ShaS, SS, MB, MI, AR, MMH, ML, NJ, AM, MOF, TBE, NS, GK, SIK, RKK, MF, MH, MAHM, NK, ZBA, NN, AHMNN, KMFU; The figures were prepared by MU, HA, SH, MAR; Microarray was done by HA, AR, NJD; Analysis, MU, GB, AI, HA, NN, MIH, AR, MAR, NJD; Interpretation of variants, MU, HA, AI; Comprehensive literature search was conducted by MU, HA, NJD, AR, NN, ZBA, SH, TBE, SS. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

We thank all the patients and participants of this study. We graciously thank Mohammed Salah Uddin Danis, Managing Director, NeuroGen Healthcare, Dhaka, Bangladesh for helping us at NeuroGen conducting these experiments.

Authors' Information

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 Children's 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. International Horticulture, Leibniz University Hannover, Section Plant Biotechnology 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 Biochemistry and Biotechnology, University of Houston 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 & Gynaecology, Bangabandhu Sheikh Mujib Medical University, Dhaka, Bangladesh 15. Department of Pharmaceutical Sciences, North South University, Dhaka, Bangladesh 16. Department of Biomedical Engineering, Military Institute of Science and Technology, Dhaka, Bangladesh 17. Department of Paediatrics, BIRDEM General Hospital 18. College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, UAE 19. Department of Biochemistry, Holy Family Red Crescent Medical College, Dhaka, Bangladesh 20. Cellular Intelligence (Ci) Lab, GenomeArc Inc., Toronto, ON, Canada

References

1. Bale TL, Baram TZ, Brown AS, Goldstein JM, Insel TR, McCarthy MM, et al. Early life programming and neurodevelopmental disorders [Internet]. Biol Psychiatry. Biol Psychiatry; 2010 [cited 2021 Jun 13]. p. 314–9. Available from: https://pubmed.ncbi.nlm.nih.gov/20674602/.
2. Zoghbi HY. Postnatal Neurodevelopmental Disorders: Meeting at the Synapse? [Internet]. Science (80-.). American Association for the Advancement of Science; 2003 [cited 2021 Jun 13]. p. 826–30. Available from: https://science.sciencemag.org/content/302/5646/826.
3. Krakowiak P, Walker CK, Bremer AA, Baker AS, Ozonoff S, Hansen RL, et al. Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. Pediatrics [Internet]. American Academy of Pediatrics; 2012 [cited 2021 Jun 13];129:e1121. Available from: /pmc/articles/PMC3340592/.
4. Dennis M, Francis DJ, Cirino PT, Schachar R, Barnes MA, Fletcher JM JM. Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders [Internet]. J. Int. Neuropsychol. Soc. Cambridge University Press; 2009 [cited 2021 Jun 13]. p. 331–43. Available from: https://pubmed.ncbi.nlm.nih.gov/19402919/.
5. Levitt P, Eagleson KL, Powell EM. Regulation of neocortical interneuron development and the implications for neurodevelopmental disorders [Internet]. Trends Neurosci. Trends Neurosci; 2004 [cited 2021 Jun 13]. p. 400–6. Available from: https://pubmed.ncbi.nlm.nih.gov/15219739/.
6. Abou Jamra R, Wohlfart S, Zweier M, Uebe S, Priebe L, Ekici A, et al. Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity. Eur J Hum Genet [Internet]. Nature Publishing Group; 2011 [cited 2021 Jun 13];19:1161–6. Available from: http://projects.tcag.ca/variation/.

7. Shevell M, Ashwal S, Donley D, Flint J, Jingold M, Hirtz D, et al. 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 [Internet]. Lippincott Williams and Wilkins; 2003 [cited 2021 Jun 13];60:367–80. Available from: https://pubmed.ncbi.nlm.nih.gov/12578916/.

8. CDC estimate on autism. Prevalence increases by nearly 10 percent, to 1 in 54 children in the U.S. | Autism Speaks [Internet]. [cited 2021 Jun 13]. Available from: https://www.autismspeaks.org/press-release/cdc-estimate-autism-prevalence-increases-nearly-10-percent-1-54-children-us.

9. Hampson DR, Adusei DC, Pacey LKK. The neurochemical basis for the treatment of autism spectrum disorders and Fragile X Syndrome [Internet]. Biochem. Pharmacol. Elsevier Inc.; 2011 [cited 2021 Jun 13]. p. 1078–86. Available from: https://www.europepmc.org/article/med/21336364.

10. Matson JL, Shoemaker M. Intellectual disability and its relationship to autism spectrum disorders [Internet]. Res. Dev. Disabil. Res Dev Disabil; 2009 [cited 2021 Jun 13];30:1107–14. Available from: https://pubmed.ncbi.nlm.nih.gov/19604668/.

11. Zhang Y, Haraksingh R, Grubert F, Abyzov A, Gerstein M, Weissman S, et al. Child Development and Structural Variation in the Human Genome. Child Dev [Internet]. Child Dev; 2013 [cited 2021 Jun 13];84:34–48. Available from: https://pubmed.ncbi.nlm.nih.gov/23317672/.

12. Gruber I, McVicar T, Yen S, Buxbaum J, Rorie K, et al. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. Sci Transl Med [Internet]. Sci Transl Med; 2011 [cited 2021 Jun 13];3. Available from: https://pubmed.ncbi.nlm.nih.gov/21832240/.

13. Ahn K, Gotay N, Andersen TM, Anvari AA, Gochman R, Lee Y, et al. High rate of disease-related copy number variations in childhood onset schizophrenia. Mol Psychiatry [Internet]. Nature Publishing Group; 2014 [cited 2021 Jun 13];19:568–72. Available from: https://pubmed.ncbi.nlm.nih.gov/23689535/.

14. Oskoui M, Gazzellone MJ, Thiruvahindrapuram B, Zarrei M, Andersen J, Wei J, et al. Clinically relevant copy number variations detected in cerebral palsy. Nat Commun [Internet]. Nature Publishing Group; 2015 [cited 2021 Jun 13];6. Available from: https://pubmed.ncbi.nlm.nih.gov/26236009/.

15. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, et al. A copy number variation morbidity map of developmental delay. Nat Genet [Internet]. Nat Genet; 2011 [cited 2021 Jun 13];43:838–46. Available from: https://pubmed.ncbi.nlm.nih.gov/21841781/.

16. Pinto D, Delaby E, Merico D, Barbosa M, Merikangas K, Klei L, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am J Hum Genet [Internet]. Cell Press; 2014 [cited 2021 Jun 13];94:677–94. Available from: https://pubmed.ncbi.nlm.nih.gov/24768552/.

17. Chien WH, Gau SSF, Liao HM, Chiu YN, Wu YY, Huang YS, et al. Deep exon resequencing of DLGAP2 as a candidate gene of autism spectrum disorders. Mol Autism [Internet]. Mol Autism; 2013 [cited 2021 Jun 13]. Available from: https://pubmed.ncbi.nlm.nih.gov/23915500/.

18. Hamdan FF, Daoud H, Piton A, Gauthier J, Dobrzeniecka S, Krebs MO, et al. De novo synap1 mutations in nonsyndromic intellectual disability and autism. Biol Psychiatry [Internet]. Biol Psychiatry; 2011 [cited 2021 Jun 13];69:898–901. Available from: https://pubmed.ncbi.nlm.nih.gov/21237447/.

19. Moessner R, Marshall CR, Sutcliffe JS, Skau J, Pinto D, Vincent J, et al. Contribution of SHANK3 mutations to autism spectrum disorder. Am J Hum Genet [Internet]. University of Chicago Press; 2007 [cited 2021 Jun 13];81:1289–97. Available from: https://www.nature.com/articles/17999366/.

20. Liu Y, Zhang Y, Zhao D, Dong R, Yang X, Tammimies K, et al. 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 [Internet]. John Wiley & Sons, Ltd; 2015 [cited 2021 Sep 22];168:258–64. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ajmg.b.32306.

21. Woodbury-Smith M, Nicolson R, Zarrei M, Yuen RKC, Walker S, Howe J, et al. Variable phenotype expression in a family segregating microdeletions of the NRXN1 and MBD5 autism spectrum disorder susceptibility genes. npj Genomic Med 2017 21 [Internet]. Nature Publishing Group; 2017 [cited 2021 Sep 22];2:1–8. Available from: https://www.nature.com/articles/s41525-017-0020-9.

22. Cristino AS, Williams SM, Hawi Z, An JY, Bellgrove MA, Schwartz CE, et al. Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. Mol Psychiatry [Internet]. Nature Publishing Group. 2014;19:294–301. [cited 2021 Jun 13];. Available from: www.nature.com/mp.

23. Chen ES, Giegek CO, Rosenfeld JA, Diallo AB, Maussion G, Chen GG, et al. Molecular convergence of neurodevelopmental disorders. Am J Hum Genet [Internet]. Cell Press; 2014 [cited 2021 Jun 13];95:490–508. Available from: https://pubmed.ncbi.nlm.nih.gov/25307298/.

24. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. Am J Hum Genet [Internet]. Am J Hum Genet; 2010 [cited 2021 Apr 17];86:749–64. Available from: https://pubmed.ncbi.nlm.nih.gov/20466091/.

25. (11) (PDF) Prevalence of mental disorders, mental retardation, epilepsy and substance abuse in children [Internet]. [cited 2021 Jun 12]. Available from: https://www.researchgate.net/publication/27406488_Prevalence_of_mental_disorders_mental_retardation_epilepsy_and_substance_abuse_in_children.

26. Mullick M, Goodman R. 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 [Internet]. Soc Psychiatry Psychiatr Epidemiol; 2005 [cited 2021 Jun 12];40:663–71. Available from: https://pubmed.ncbi.nlm.nih.gov/16091858/.

27. Lord C, Rutter M, Dilavore PC, Risi S, Gotham K, Bishop SL, et al. ADOS-Autisme diagnostisch observatieschema Handleiding. 2012.

28. Rahaman MA, Lopa M, Uddin KMF, Baqui MA, Keya SP, Faruk MO, et al. An Exploration of Physical and Phenotypic Characteristics of Bangladeshi Children with Autism Spectrum Disorder. J Autism Dev Disord [Internet]. Springer; 2020 [cited 2021 Jun 13];51. Available from: https://pubmed.ncbi.nlm.nih.gov/32975665/.

29. Akhter S, Hussain AHME, Shefa J, Kundu GK, Rahman F, Biswas A. Prevalence of Autism Spectrum Disorder (ASD) among the children aged 18-36 months in a rural community of Bangladesh: A cross sectional study [version 1; referees: 1 approved, 2 approved with reservations]. F1000Research
98. Gillentine MA, Yin J, Bajic A, Zhang P, Cummock S, Kim JJ, et al. Functional Consequences of CHRNA7 Copy-Number Alterations in Induced Pluripotent Stem Cells and Neural Progenitor Cells. Am J Hum Genet [Internet]. Elsevier Company; 2017;101:874–87. Available from: https://doi.org/10.1016/j.ajhg.2017.09.024.

99. Gillentine MA, Schaaf CP. The human clinical phenotypes of altered CHRNA7 copy number. Biochem Pharmacol [Internet]. Elsevier Inc.; 2015;97:352–62. Available from: http://dx.doi.org/10.1016/j.bcp.2015.06.012.

100. Uddin M, Unda BK, Kwan V, Holzapfel NT, White SH, Chalil L, et al. OTUD7A Regulates Neurodevelopmental Phenotypes in the 15q13.3 Microdeletion Syndrome. Am J Hum Genet [Internet]. Cell Press; 2018 [cited 2021 May 29];102:278–95. Available from: /pmc/articles/PMC5985537/.

101. SA JHULMSSKJS Y, et al. CNTN6 copy number variations in 14 patients: a possible candidate gene for neurodevelopmental and neuropsychiatric disorders. J Neurodev Disord [Internet]. J Neurodev Disord; 2015 [cited 2021 Jul 7];7. Available from: https://pubmed.ncbi.nlm.nih.gov/26257835/.

102. Repnikova EA, Lyalin DA, McDonald K, Astbury C, Hansen-Kiss E, Cooley LD, et al. CNTN6 copy number variations: Uncertain clinical significance in individuals with neurodevelopmental disorders. Eur J Med Genet Elsevier Masson. 2020;63:103636.

103. Y Y, R H, R F, N O, Y H, H Y, et al. Duplication of the NPHP1 gene in patients with autism spectrum disorder and normal intellectual ability: a case series. Ann Gen Psychiatry [Internet]. Ann Gen Psychiatry; 2014 [cited 2021 Jul 7];13. Available from: https://pubmed.ncbi.nlm.nih.gov/25126106/.

104. Karczewski K. The Genome Aggregation Database (gnomAD). 2019.

**Figures**

**Figure 1**

Physical characteristics difference based on sex and variant types (Full cohort). A), B) C) and D) comparison of age (years), head circumference (cm), weight (kg) and BMI (kg/m2) of male and female in the full cohort. E) and F) comparison of head circumference (cm) and weight (kg) between male-female CNV positive groups. Here, CNV=copy number variation, CNV positive=carrying deletion or duplication or both deletion and duplication, CNV negative=those have no or benign CNVs. All p values were derived from independent-samples t-test. Bold p values are Bonferroni corrected.
Figure 2

Physical and phenotypic characteristics difference based on sex, autism diagnosis and variant types (ADOS-2 sub cohort). A) and B) comparison of head circumference of male and female in the combined (ASD positive/negative in ADOS-2 assessment) and ASD positive groups. C) The percentage of ASD positive and non-spectrum groups in the ADOS-2 sub cohort impacted by copy number variation. P value was calculated by Fisher’s exact test. D) comparison of social affect score in deletion, duplication and CNV negative groups. E) comparison of total ADOS-2 score in deletion, duplication and CNV negative groups. ASD=Autism Spectrum Disorder, positive=ASD diagnosed by ADOS-2 test, CNV=copy number variation.

Figure 3

Physical and phenotypic characteristics difference based on sex, autism diagnosis and variant types (ADOS-2 sub cohort). A) and B) comparison of head circumference of male and female in the combined (ASD positive/negative in ADOS-2 assessment) and ASD positive groups. C) The percentage of ASD positive and non-spectrum groups in the ADOS-2 sub cohort impacted by copy number variation. P value was calculated by Fisher’s exact test. D) comparison of social affect score in deletion, duplication and CNV negative groups. E) comparison of total ADOS-2 score in deletion, duplication and CNV negative groups. ASD=Autism Spectrum Disorder, positive=ASD diagnosed by ADOS-2 test, CNV=copy number variation.
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 1053 CNVs, 3% pathogenic, 17% VOUS and rest 80% 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) The 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 5365.26kb and 17451.72kb respectively. Whereas, the average length of VOUS deletion and duplication is 129.17kb and 476.76kb.

Figure 4
Circos[50]plot illustrating neurodevelopmental disorder map showing frequency of rare and common CNV. (A) Frequency distribution of all rare CNVs throughout the chromosome 1 to 22. Dots in the red, green and blue circles indicate frequency of >30kb deletion, >50kb 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 the chromosome 1 to 22. Dots in the green, purple and blue indicate frequency of >30kb deletion, >50kb duplication and LOH CNVs, respectively. Each colored circle contains values less than 0.01 and more than 0.30.

Figure 5
KEGG and Gene ontology pathway enrichment analysis of pathogenic CNVs identified in the study cohort. (A) The impacted genes within the CNV breakpoints of all pathogenic deletions identified “GO:0019787”, “GO:0060548” and “GO:0099003” pathways to be highly significant. (B) The 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 was 0.01 and 0.001, respectively. P-values are denoted using color gradient (low p values with darker colors).
Figure 6

Detail information of PSMC3 gene from various databases. A) and B) list of variants overlapping with PSMC3 gene from DECIPHER[90] and ClinVar[89] databases, respectively. C) summary of 529 rare SNVs in PSMC3 gene from gnomAD[104] database. D) A 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 7.7kb (chr11:47,418,769-47,426,473) disrupting PSMC3, in common. Red and blue rectangles symbolize PSMC3-involving gross deletion and duplication, respectively. And orange single bar symbolizes PSMC3-involving SNP. Here, CEG=critical exon gene, SNV=short nucleotide variation, SNP=single nucleotide polymorphism, gnomAD= The Genome Aggregation Database.

Figure 7

A 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 38.4kb (chr19:35,700,296-35,738,700) disrupting 2 genes, ZBTB32 and KMT2B, in common. Red and blue rectangles symbolize KMT2B-involving gross deletion and duplication, respectively.