# Combining optical genome mapping and RNA-seq for structural variants detection and interpretation in unsolved neurodevelopmental disorders

Bing Xiao[1†], Xiaomei Luo[1†], Yi Liu[1], Hui Ye[1], Huili Liu[1], Yanjie Fan[1*] and Yongguo Yu[1*]

## Abstract

**Background** Structural variations (SVs) are key genetic contributors to neurodevelopmental disorders (NDDs). Exome sequencing (ES), the current first-line tool for genetic testing of NDDs, falls short in SVs detection. This diagnostic gap is being actively addressed by new methods such as optical genome mapping (OGM).

**Methods** This study evaluated the utility of combining OGM and RNA-seq in the detection and interpretation of SVs in ES-negative NDDs. OGM was performed in 43 patients with NDDs with inconclusive ES results. Candidate SVs were selected based on disease association and pathogenicity evaluation, and further validated or reconstructed by alternative methods, including long-read sequencing for a complex rearrangement event. RNA-Seq was performed on blood samples from patients with candidate SVs to facilitate interpretation of pathogenicity.

**Results** OGM detected four candidate SVs, and RNA-seq confirmed the pathogenicity of three SVs in the patient cohort. This combined approach solved three cases—two cases with de novo SVs in genes associated with autosomal dominant NDDs, including a deletion encompassing the promoter and 5'UTR of *MBD5* and an intragenic duplication of *PAFAH1B1*, and a third case possessing an intragenic duplication *in trans* with a pathogenic single-nucleotide variant of *PLA2G6*, associated with autosomal recessive NDDs. The expression alteration of the affected genes and the tandem positioning of two intragenic duplications were confirmed by RNA-seq. In the fourth case, OGM detected a complex rearrangement involving chromosomes 2 and 6, much more complex than the de novo t(2:6)(q13;q15) indicated by conventional cytogenetic analysis. Reconstruction showed that 17 segments of 6q15 spanning 9.3 Mb were disarranged and joined 2q11.2, with four breakpoints detected in the 5' and 3' non-coding region of the NDD-associated gene *SYNCRIP*. RNA-seq revealed largely preserved *SYNCRIP* expression, leaving the pathogenicity of this complex rearrangement event uncertain.

**Conclusions** SVs in ES-negative NDDs can be identified by OGM, which is particularly useful for SVs in non-coding regions not covered by ES. OGM helps to construct complex SVs and provides information on the location and orientation of duplications, which is crucial for pathogenicity interpretation. The integration of RNA-seq facilitates

---

†Bing Xiao and Xiaomei Luo contributed equally to this work.

*Correspondence:
Yanjie Fan
fanyanjie@shsmu.edu.cn
Yongguo Yu
yuyongguo@shsmu.edu.cn
Full list of author information is available at the end of the article

the interpretation of the functional consequences of SVs at the transcriptional level. These findings demonstrate the utility and feasibility of combining OGM and RNA-seq in ES-negative cases with NDDs.

**Keywords**  Optical genome mapping, RNA-seq, Structural variants, Neurodevelopmental disorders

## Background

Optical Genomic Mapping (OGM) has emerged as a revolutionary technology in genomics by comprehensively mapping all types of structural variants (SVs) at high resolution in a single assay. OGM not only maps all classes of SVs with excellent concordance with established diagnostic standards but also unveils SVs that are beyond the reach of current technologies [1–3], which holds tremendous promise for clinical applications. The application scenarios of OGM are diverse, with many translational research studies highlighting its efficacy in detecting clinically relevant SVs in hematological malignancies [3–6]. Recent studies have also expanded the use of OGM to constitutional abnormalities beyond malignancies [2, 7, 8]. Noteworthy applications include its use as a cytogenomic tool for prenatal diagnostics in recent literature [9]. However, compared to the accumulated evidence on the utility of OGM in cancer, the clinical utility of OGM in developmental disorders remains to be explored.

SVs play a role in neurodevelopmental disorders (NDDs) that cannot be ignored, as highlighted by Hu et al. [10]. Exome sequencing (ES) is currently the primary clinical diagnostic tool for NDDs, which effectively detects sequence variants and provides simultaneous detection of SVs, mainly on copy number variations (CNVs). Despite the dominance of ES in the diagnosis of NDDs (31–53% diagnostic rate) [11], its reliance on short reads limits its ability to comprehensively solve complex SVs. Additionally, ES misses genomic variations outside the coding region, leaving a significant gap in disease etiology and hindering diagnostic improvement. OGM may partially fill this gap by detecting SVs. Genomic sequencing, despite its comprehensive coverage, analytical tools and substantial computational resources, is often not fully implemented in clinical settings [12–14]. Meanwhile, the interpretation of genomic SVs represents another major challenge beyond detection. The major impact of genomic SVs on disease is likely to result from alterations in gene dosage and transcript expression levels. SVs can directly alter protein products though copy number changes, coding sequence disruptions, gene silencing through promoter deletions or translocations, or create novel fusion transcripts through rearrangements. Transcriptome analysis allows the identification of variants that result in aberrant transcription and splice junctions [15]. RNA-seq, a high-throughput technology that directly measures transcriptome changes to reveal the functional impact of variants, is a complementary method in the genomic diagnosis of monogenetic disorders [16]. Unlike DNA-based analyses, RNA-seq provides a nuanced view of the transcriptome, enabling researchers to distinguish functionally relevant SVs from those with neutral effects and to pinpoint their precise effects on gene expression and splice junctions, facilitating accurate interpretation of detected SVs. A recent study showed that RNA-Seq contributed to the interpretation of three SVs at the transcriptional or regulatory level [17].

Therefore, to improve the detection and interpretation of SVs in patients with ES-negative NDDs, we investigated the clinical utility of combining OGM with RNA-seq, ultimately facilitating its translation into effective clinical practice.

## Methods

### Cohort design

This prospective observational study was approved by the Ethnic Committee of Xinhua Hospital. Forty-three unrelated Chinese Han probands with ES-negative NDDs in Xinhua Hospital from August 2022 to June 2023 were recruited, of whom 12 (28%) were female. The median age within the cohort was 12 years, ranging from 2 to 31 years. All 43 probands in this study had negative ES results prior to recruitment, including trio-ES in 30 cases and solo-ES in 13 cases. Solo-ES or trio-ES were performed using the capture kit of xGen Exome Research Panel v1.0 (Integrated DNA Technologies, Coralville, IA, USA, $n=41$) or Agilent SureSelect Exome V5 (Agilent, Santa Clara, California, USA, $n=2$). Both SNV and CNV analysis were performed following routine clinical genetic procedures. For 20 of the 43 patients, chromosomal microarray analysis was performed using Affymetrix CytoScan™ 750 K (Santa Clara, California, USA, $n=4$), CytoScan™ HD (Santa Clara, California, USA, $n=4$), or CNV-seq ($n=12$) prior to ES (Additional file 1: Table S1).

Phenotypic data were collected as Human Phenotype Ontology terms. All probands presented with syndromic or non-syndromic moderate to severe intellectual disability (HP: 0001249) or global developmental delay (HP: 0001263). Individuals were excluded if a non-genetic etiology was suspected. The detailed clinical phenotypes are listed (Additional file 1: Table S1).

This study was conducted in accordance with the principles of the Declaration of Helsinki. Approval

was obtained from the Ethic Committee of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (XHCC-C-2022–077-1). This study (ChiCTR2200062714) has been registered in Chinese Clinical Trail Registry (https://www.chictr.org.cn/). Written informed consent for genetic analysis was obtained from the patients or their legal guardians.

**Optical genome mapping (OGM)**

Ultra-high molecular weight DNA from peripheral blood of patients and available parents was extracted, quantified (at least 4 ng/ul), labeled, and processed using the Prep SP Blood and Cell DNA Isolation Kit, Qubit Fluorometers, and Prep DLS Labeling kits (Bionano Genomics Inc., San Diego, CA, USA) according to the manufacturer's instructions. Labeled DNA was loaded onto a Saphyr chip for linearization and imaging on the Saphyr instrument. After checking QC metrics, the de novo assembly pipeline was run using the Bionano Solve V3.7 software. SVs (based on the assembled genome maps) were called against the human reference genome (GRCh38/hg38) and visualized using Bionano Access software V1.7 and compared with an OGM dataset of 180 control samples from apparently healthy individuals (provided by Bionano Genomics).

For data analysis, SVs were filtered according to the following criteria (Fig. 1): (a) filter out SVs with ≥ 1% frequency in the OGM control sample SVs database, (b) filter out SVs not involving disease-associated genes, and (c) filter out SVs not involving NDD-associated genes.
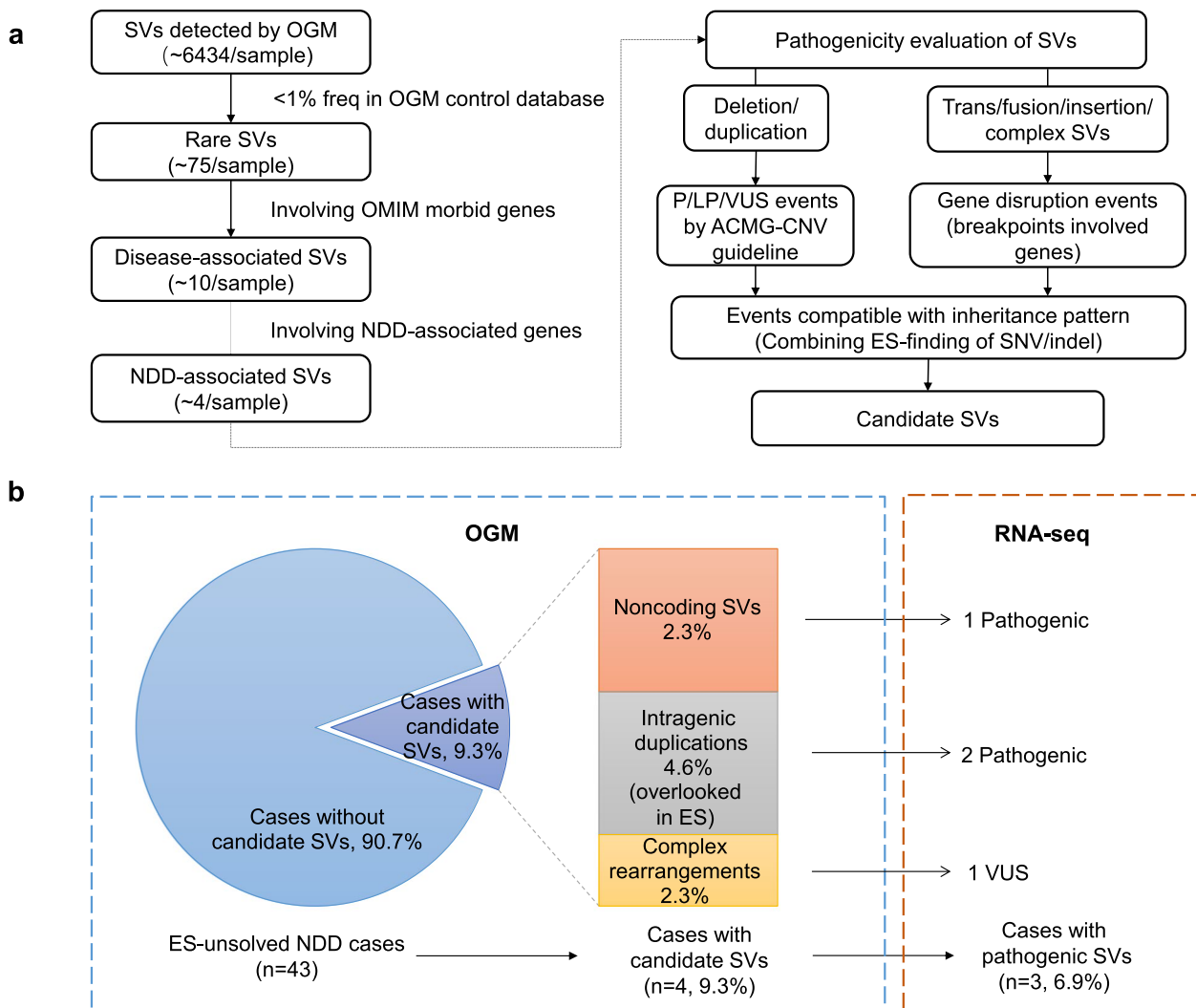


**Fig. 1** Overview of SVs filter combining OGM and RNA-seq. **a** Workflow of OGM analysis. **b** The diagnostic rate of combining OGM and RNA-seq (3/43, 6.9%)

Xiao *et al. Genome Medicine*        (2024) 16:113

Page 4 of 12

The SVs that passed the filtering were subjected to pathogenicity evaluation: deletions/duplications were evaluated according to the variant interpretation guidelines of the American College of Medical Genetics [18–20], while the other types of gene disruption events (e.g., translocations, fusions, insertions, inversions, and complex rearrangements) were determined based on the location of breakpoints. Variants could be classified into pathogenic, likely pathogenic, uncertain significance, likely benign, and benign according to the guideline. Only SVs compatible with the disease inheritance pattern were selected as candidate SVs. Database of Genomic variants (http://dgv.tcag.ca/dgv/app/home) and DECIPHER (https://www.dechiphergenomics.org/) were referred for SVs evaluation. The OGM nomenclature was used according to the recommendations of the ISCN Standing Committee [21].

### RNA-seq
RNA sequencing was performed on whole blood RNA from cases with candidate SVs. Oligo dT-enriched mRNA was used for subsequent library preparation. The library was generated using the NEBNext Ultra II Directional RNA LibraryPrep kit (#E7420 New England Biolabs, Massachusetts, USA), followed by sequencing on Illumina NovaSeq PE150 (Illumina Inc., San Diego, CA, USA) with approximately 100 million reads per sample. Reads were aligned using STAR V.2.7.9.a [22] in two-pass mode (GRCh37 release75, Gencode34). Analysis of RNA-seq data was performed using Drop V.1.2.2 [23] under the default module of aberrant expression detection (OUTRIDER) with an in-house collection of 139 samples (unrelated individuals of Chinese Han ethnicity, aged 2–16 years, 52 (37%) are female). Aberrant expression was filtered using the following thresholds: padj $\leq 0.05$ and $|Z$ score$| > 3$. Sample rank and volcano plots were generated using the R package OUTRIDER [24].

### Complementary to complex SV reconstruction with OGM using long-read whole genome sequencing (LRWGS)
Genomic DNA was extracted from patient blood leukocytes using a kit from New England Biolabs (the Monarch® HMW DNA Extraction Kit). The library was prepared for nanopore sequencing using the Ligation Sequencing Kit V14 (SQK-LSK114) and run on a PromethION 2 SOLO sequencer using a PRO114M (R10) flow cell (all Oxford Nanopore Technologies, Oxford, UK). MinKNOW version 24.02.16 was used for base calling and FASTQ conversion procedures. Control datasets were also sequenced using PromethION. Nanopore long reads were aligned to the human reference genome (GRCh38) using minimap2 v2.27 [25]. Reads were then sorted by SAMtools v1.15 [26]. Delly v1.1.8 [27] and Sniffles2 v2.2 [28] were used to detect SVs. The Integrative

Genomics Viewer (https://igv.org/) was used to analyze and visualize the data.

## Results
For the 43 samples tested, the average mapping rate of OGM was 89.4% and the average coverage was 214.6X (Additional file 1: Table S1). An average of 6434 SVs were identified per sample, with ~75 rare SVs; after further filtering by disease and phenotype association, an average of 4 NDD-associated SVs per sample required further evaluation. In this study, OGM detected four candidate SVs involving genes potentially related to the patients' phenotypes (4 out of 43, 9.3%), and subsequent RNA-seq was performed to further interpret the pathogenicity of these SVs (Fig. 1). Successfully identified pathogenic SVs solved the genetic causes of three individuals (3/43, 6.9%), including two de novo SVs (P4, P40) associated with autosomal dominant NDDs and a third diagnostic SV (P2) *in trans* with a pathogenic SNV in a gene associated with autosomal recessive NDD. OGM identified complex rearrangement events in patient 30, and several candidate genes were found in the deletion regions and at breakpoints. Detailed information on these SVs and relevant genes/disorders is described in Table 1.

### OGM revealed non-coding SVs and RNA-seq contributed to the interpretation at transcriptional level
#### Case P40: a non-coding deletion affecting the MBD5 promoter and 5′ UTR
The patient is a developmentally delayed male aged 6 years and 4 months. He is the second child born at full term with a birth weight of 3.2 kg (50th percentile). There is no history of birth asphyxia. While gross motor development appears normal, his speech remains delayed—he was able to say "dad" and "mom" at the age of two, and he still spoke mostly simple words and showed limited understanding of instructions at 6 years old. Physical examination revealed a height of 115 cm (15th–50th percentile), a weight of 21 kg (50th percentile), and a head circumference of 51.8 cm (50th percentile). Magnetic resonance imaging (MRI) of the brain showed no apparent abnormalities. Intelligence testing using the WISC-R indicated a language score of 56, an operational score of 55, and a total IQ of 48. ES revealed no diagnostic variants. OGM identified a rare heterozygous deletion of 134 kb, termed ogm[GRCh38] 2q23.1(147,966,774_148,100,811)×1 (Fig. 2a). The parents' OGM results confirmed that the *MBD5* deletion was de novo. This deletion involved partial *ORC4* and 5′-untranslated exon 1 of *MBD*5 (NM_001378120) (Fig. 2b). *ORC4* is associated with autosomal recessive Meier-Gorlin syndrome 2. *MBD5* is associated with autosomal dominant

**Table 1** Candidate structural variants identified in this study

| Patient | Age | Sex | Clinical Indication | Gene | Variants description | Variants in ISCN style format |
|---|---|---|---|---|---|---|
| P40 | 6y4m | male | Global developmental delay | *MBD5* | Pathogenic; a de novo heterozygous 134-kb deletion involving *MBD5* | ogm[GRCh38] 2q23.1(147966774_148100811)x1 |
| P2 | 2y | male | Global developmental delay, early psychomotor and motor regression | *PLA2G6* | Pathogenic; a heterozygous 19-kb insertion within *PLA2G6* | ogm[GRCh38] ins(22;?)(q13.1;?) (38101242_38118769;?) |
| P4 | 7y | male | Global developmental delay and seizure | *PAFAH1B1* | Pathogenic; a *de novo* heterozygous 9.5-kb insertion within *PAFAH1B1* | ogm[GRCh38] ins(17;?)(p13.3;?) (2654573_2670266;?) |
| P30[a] | 3y11m | male | Developmental delay and seizure | / | VUS; multiple complex rearrangement events involving chromosomes 2 and 6 | ogm[GRCh38] (2,6) cpx |

*Abbreviations: m* Month, *P* Patient, *VUS* Variants of uncertain significance, *y* Year

[a] The detailed molecular karyotype for P30 determined by LRWGS and OGM was seq[GRCh38]der(6)(pter→q14.1::q14.3inv:: q14.3q15inv::2q12.1→qter;der(2)(pter→q 11.2::q14.3::q15::q14.1q14.2inv::q14.1::q14.1::q14.3::q14.1::q14.3inv::q14.3inv::q15inv::q14.3inv:: q14.1::q14.1::q14.1 inv:: 2q11.2q12.1inv::6q15→qter)
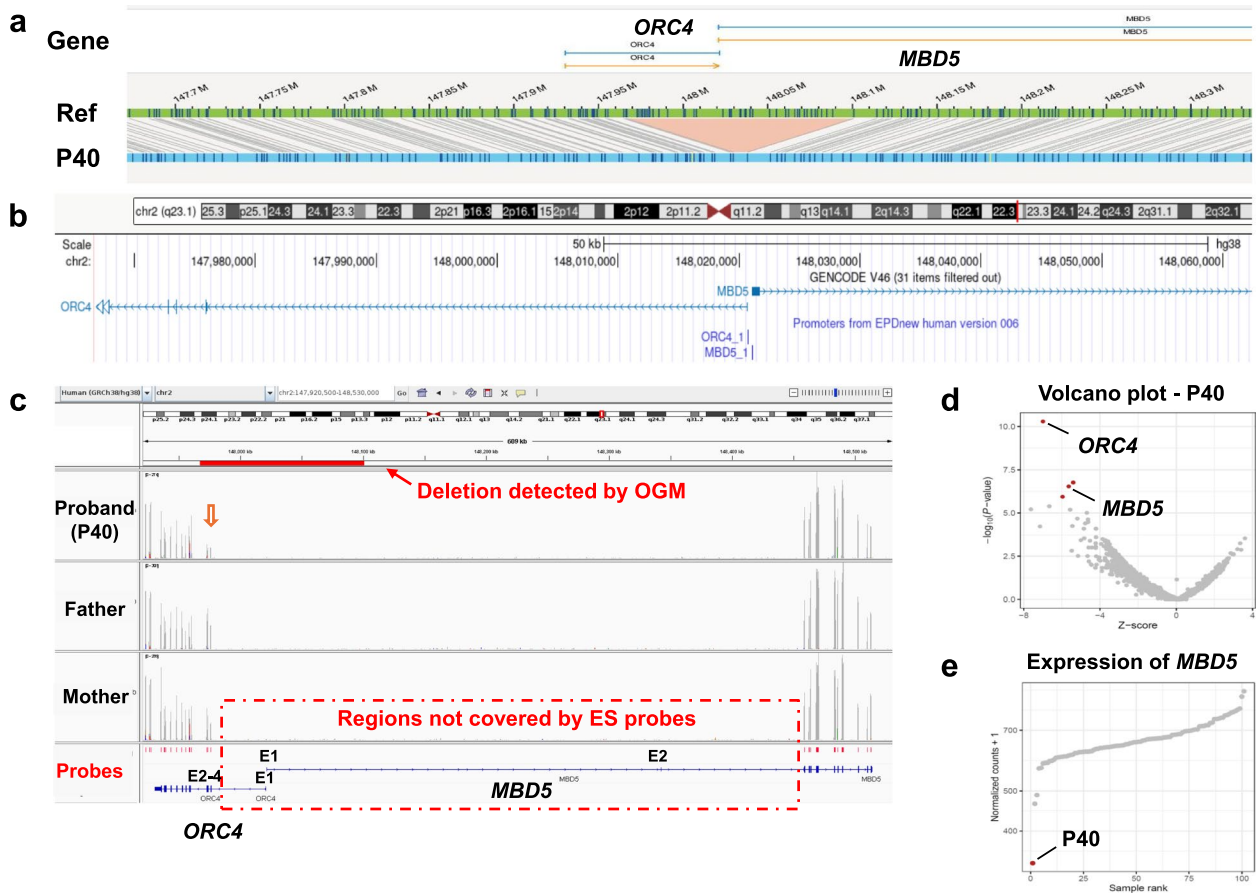


**Fig. 2** Characterization of a de novo 134-kb heterozygous deletion partially involving *MBD5* and *ORC4* in P40. **a** OGM showed a deletion of 5'-untranslated exons of *MBD5* (ogm[GRCh38] 2q23.1(147,966,774_148,100,811)×1). **b** UCSC Genome Bowser showed that the deletion included the 5'UTR and promoter of *MBD5*. **c** The deletion in MBD5 was not captured by the ES capture kit (xGen Exome Research Panel v1.0, IDT). **d** RNA-seq confirmed significantly reduced expression of *MBD5* together with the nearby gene *ORC4* in the patient's blood compared to controls. **e** The normalized read count of P40 was the lowest in the cohort tested

intellectual developmental disorder (MIM: 156200). As the disease associated with *MBD5* matched the phenotype and inheritance pattern of the patient, this 134-kb deletion spanning part of *MBD5* was considered as a candidate disease-causing SV. This deletion was validated using Affymetrix CytoScan HD array, which

revealed a 123-kb heterozygous deletion, termed arr[GRCh38]2q23.1(147,970,757−148,094,394)×1 (Additional file 2: Fig. S1). This deleted region of *MBD5* was not covered by ES (no probes in this region in xGen Exome Research Panel v1.0, IDT, see Fig. 2c). This deletion spans the promoter and 5′ untranslated exon 1 of *MBD5* (Fig. 2b), which could affect transcription, but the functional consequence remained uncertain as alternative transcription initiation cannot be excluded. The subsequent RNA-seq result showed a significantly reduced expression of *MBD5* together with the nearby *ORC4* gene in the patient's blood compared to controls (fold change = 0.51; $Z$ score = − 5.64; padj = 0.014, Fig. 2d). The expression of both *ORC4* and *MBD5* were among the most significantly downregulated genes based on RNA-seq. For *MBD5*, the normalized counts of P40 ranked the lowest among the tested cohort, with approximately half the reduction (P40: 333.5 counts, mean cohort: 660.7 counts, Fig. 2e). This approximately half reduction in *MBD5* transcription was consistent with a heterozygous deletion that abolished the original transcription initiation without alternative transcription initiation. In this case, the functional consequence provided by RNA-seq supports the pathogenic nature of this SV.

## OGM and RNA-seq assisted in confirming tandem positioning of duplicated exons

1. A 9.5-kb insertion in the *PAFAH1B1* gene in P4

The patient is a 6 years and 8 months old boy with global developmental delay and generalized hypotonia. Seizures started at 2 months of age. Currently, at 7 years of age, he has significant motor limitations, unable to lift his head, sit independently, walk, or speak. Neurological examination shows poor visual tracking and response to sound. Brain MRI shows abnormally wide cerebral gyri with thickening of the cerebral cortex.

Panel testing, trio-ES, and mitochondrial sequencing performed at local hospitals revealed no clinically significant genetic variants. OGM found a novel 9.5-kb insertion in the *PAFAH1B1* gene, designated ogm[GRCh38]ins(17;?)(p13.3;?) (2,654,573_2,670,265;?) (Fig. 3a). The lack of additional genetic markers between these points made it challenging to determine the source of the insertion. Retrospective analysis of the ES data confirmed the copy number gain of three exons (exons 3–5, Fig. 3b) of *PAFAH1B1* (NM_000430.4). RNA-seq analysis revealed a back-spliced fusion junction between exon 5 and exon 3 (Fig. 3c). These results together indicated a tandemly positioned intragenic duplication from exon 3 to exon 5 in *PAFAH1B1*. The transcriptional change was
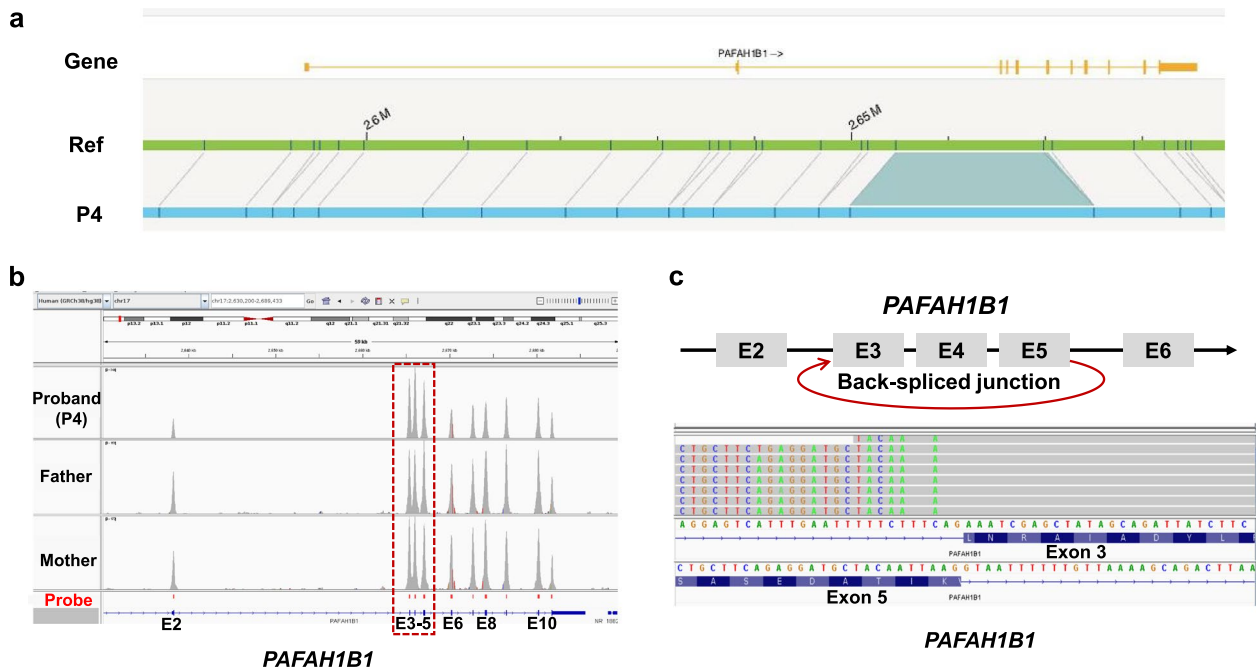


**Fig. 3** A de novo heterozygous 9.5-kb intragenic tandem duplication from exon 3 to exon 5 within the *PAFAH1B1* gene in P4. **a** OGM of the *PAFAH1B1* locus showed an intragenic insertion. **b** Integrative Genomics Viewer (IGV) of ES data showed a copy number gain of exons 3–5 in *PAFAH1B1*. **c** RNA-seq showed an abnormal fusion junction involving exon 5 to exon 3 along with abnormal splicing events

Xiao *et al. Genome Medicine*     (2024) 16:113

Page 7 of 12

r.33_399dup, which was predicted to result in an out-of-frame product p.(Val134Lysfs*14), potentially triggering nonsense-mediated decay and reducing mRNA level. RNA-seq confirmed that *PAFAH1B1* expression was reduced compared to controls (fold change $=0.74$; $Z$ score $= -6.08$; padj $=0.0002$). This duplication was verified by genome sequencing and Sanger sequencing (Additional file 2: Fig. S2-S4, Table S2). Though this duplication was identified in a Caucasian control cohort based on Affymetrix CytoScan HD data in Database of Genomic Variants (4/873, frequency: 0.45%, database record: essv9812336, essv9812335, essv9812332, essv9812333) [29], the probe coverage of CytoScan HD array was poor in this region which put the validity of these records into questions (8 probes in a region of 5.7 kb). Neither our in-house database (containing over 10,000 array data) nor gnomAD [30] recorded CNVs in this region. Based on the frequency, tandem positioning and functional consequence, this de novo intragenic duplication of *PAFAH1B1* was interpreted as pathogenic.

2. A 19-kb insertion in the *PLA2G6* gene in P2

The patient is a male aged 2 years and 1 month with global developmental delay and rapid regression. He is the third child in the family, born at term by vaginal delivery with no history of asphyxia. The patient's development was generally normal until the age of 7 months.

He then began to regress, losing independent mobility and experiencing impaired movement in all limbs. Dried blood spots and urine gas chromatography/mass spectrometry showed no abnormalities. Biochemical tests indicated elevated levels of aspartate transaminase and lactate dehydrogenase, while cardiac ultrasound showed no abnormalities. Brain MRI revealed cerebellar atrophy.

Family history revealed a deceased older sister with similar presentation: she developed normally until the age of 1 year, then lost the ability to walk, stand, and sit, followed by cognitive regression. She died at the age of 5 with cerebellar atrophy identified on MRI. Dried blood spots and urine gas chromatography/mass spectrometry showed no abnormalities, and biochemical tests indicated elevated levels of alanine transaminase, aspartate transaminase, and lactate dehydrogenase. Karyotypic analysis, trio-ES, and mitochondrial sequencing for the proband were performed at a local hospital, but no diagnostic finding was reported.

OGM analysis revealed a 19-kb insertion in the *PLA2G6* gene (ogm[GRCh38] ins(22;?)(p13.3;?) (38101242_38118769;?)) (Fig. 4a), inherited from his mother. Re-analysis of the ES data provided by the local hospital identified a heterozygous pathogenic variant in the *PLA2G6* gene (NM_003560.4:c.109C > T, p.(Arg37*)), which was inherited from the father. Retrospective analysis of the ES data also confirmed an intragenic copy number gain of seven exons (exons 6 to exon 12) that was
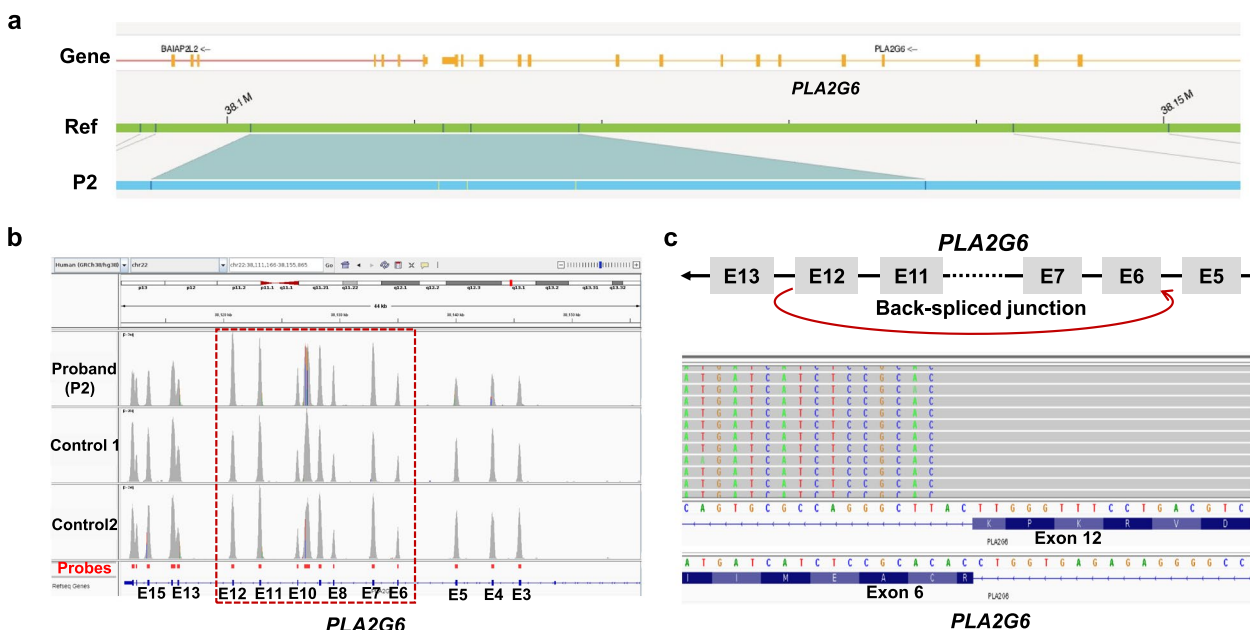


**Fig. 4** Characterization of a SV within the *PLA2G6* gene in P2. **a** OGM showing an approximately 19 kb insertion within *PLA2G6* gene. **b** IGV of the ES data showed an overlooked copy number gain of exon 6–12. **c** RNA-seq showed an abnormal back-spliced fusion junction between exon 6 and exon 12

Xiao *et al. Genome Medicine*    (2024) 16:113

Page 8 of 12

missed in the previous ES analysis (Fig. 4b). RNA-seq analysis revealed an abnormal back-spliced fusion junction between exon 6 and exon 12 (Fig. 4c). This RNA-seq finding, together with the "in situ" information from OGM, confirmed the tandem positioning of this duplication residing in *PLA2G6*. The phospholipase A2 group 6 protein encoded by *PLA2G6* gene contains 806 amino acids. The transcriptional change of *PLA2G6* in P2 was r.798_1742dup (945nt duplication), resulting in an in-frame duplication predicted to increase 39% of the total protein (315 amino acids). RNA-seq showed the expression of *PLA2G6* was increased compared to controls (fold change = 1.27; $Z$ score = 2.54; pValue = 0.008; padj = 1), though genome-wide significance was not reached. This is consistent with the presumed outcome of in-frame duplications that do not trigger nonsense-mediated RNA decay. As this duplication affects > 10% of the total amino acid sequence and is in trans with another pathogenic SNV, it is interpreted as a pathogenic duplication based on the American College of Medical Genetics guideline for single gene CNVs [18]. This duplication was verified by genome sequencing and Sanger sequencing (Additional file 2: Fig. S2-4, Table S2).

### OGM and RNA-seq assisted in determining whether candidate gene expression was altered in complex rearrangements

A male patient (P30) aged 3 years and 11 months presented with developmental delay and seizures. ES did not identify any definitive pathogenic variants. Chromosomal karyotyping analysis revealed a de novo translocation t(2:6)(q13;q15) (Additional file 2: Fig. S5a). Using OGM, we found that the chromosomal complex rearrangements (CCR) were much more complex than suggested by conventional cytogenetic analysis. In addition to the translocation between chromosomes 2q11.2 and 6q15, additional complex SVs involving 2q and 6q were identified, which including inv(6)(q14.1q14.1), fus(6;6) (q14.1;q14.3), inv(6)(q14.1q14.3) as well as multiple copy number changes around breakpoints (Additional file 2: Fig. S5b). These resulted in the following prediction: multiple double-strand breaks occurred at two chromosomes 6 and one chromosome 2, and then derivative chromosomes 2 and 6 were generated by chimeric joining of these DNA segments.

OGM identified the CCR of chromosomes 2 and 6 and provided the backbone for further refining these SVs. We used LRWGS to reconstruct these SVs on the backbone provided by OGM and refine the breakpoints to single nucleotide resolution. LRWGS combined with OGM analysis allowed us to accurately determine the full structure of the CCR, which showed chromoplexy events between the long arm

of chromosomes 2 and 6. In brief, chromosomes 2q12.1qter (chr2:104.131,686–242,193,529) and 6q15qter (chr6:89,275,859–170,805,979) showed reciprocal translocations. 6q14.1q15 (chr6:80,033,502–89,275,859) was split into 17 segments (seg 1–17), with seg 11 and 14 inversely rejoining with 6q14.1 (chr6:80,033,502) on one side and rejoining with the translocated 2q12.1qter (chr2:104.131,686–242,193,529) on the other side and the other 15 segments (seg 1–10, 12–13, 15–17) randomly rejoining with 2q11.2 (chr2:100,270,965) and at the terminal end. This resulted in a derivative chromosome 2 and a derivative chromosome 6 (Additional file 2: Fig. S5c). The molecular karyotype for P30 eventually modified by LRWGS and OGM to seq[GRCh38]der(6)(pter → q14.1::q14.3inv::q14.3q15inv::2q12.1 → qter;der(2) (pter → q11.2::q14.3::q15::q14.1q14.2inv::q14.1::q14.1::q14.3::q14.1::q14.3inv::q14.3inv::q15inv::q14.3inv::q14.1::q14.1::q14.1inv::2q11.2q12.1inv::6q15 → qter).

In addition, the combined analysis revealed 19 de novo copy number losses in chromosomes 6 and 2 and pinpointed all the breakpoints. The exact coordinates of all the breakpoints have been listed in a schematic diagram of the construction of SVs in Fig. 5, and the nearby sequence of all the breakpoints has been shown (Additional file 2: Fig. S6) as well as in depositories [31]. The deleted regions and six genes involved were listed (Additional file 2: Table S3). Notably, the *SYNCRIP* gene (MIM: 616686) was disrupted by four breakpoints (Fig. 5b). Although *SYNCRIP* is not yet included in the OMIM morbid gene list, de novo variants of *SYNCRIP* have been reported in several NDD patients, possibly through a haploinsufficiency mechanism [32, 33]. In our case, the four breakpoints were all located in the 5′ or 3′ UTR region of the gene (Fig. 5b), and the coding part was undisrupted, making it difficult to determine the functional consequence. RNA-seq results for this gene showed no statistically significant alteration of *SYNCRIP* expression (fold change = 0.91; pValue = 0.038, padj = 1; $Z$ score = − 2.11). Another gene involved is *SNHG5*, which is located in the deletion region; this gene is not yet an OMIM morbid gene, and it has not been reported to be associated with human diseases. *SNHG5* produces spliced non-coding RNAs and is host to the small nucleolar RNAs (snoRNAs) U50 (*SNORD50A*, MIM:613117) and U50-prime (*SNORD50B*, MIM:613264) [34]. Querying RNA-seq results showed the expression of *SNHG5* was reduced compared to the controls (fold change = 0.36; padj = 0.028; $Z$ score = − 2.47). For the other four relevant genes (*LONRF2*, *TTK*, *LINC02542*, and *GABRR2*), RNA-seq could not provide sufficiently confident expression outlier calling (Additional file 2: Table S3). *LONRF2* and *LINC02542* were not OMIM genes. *TTK* is required for centrosome duplication and
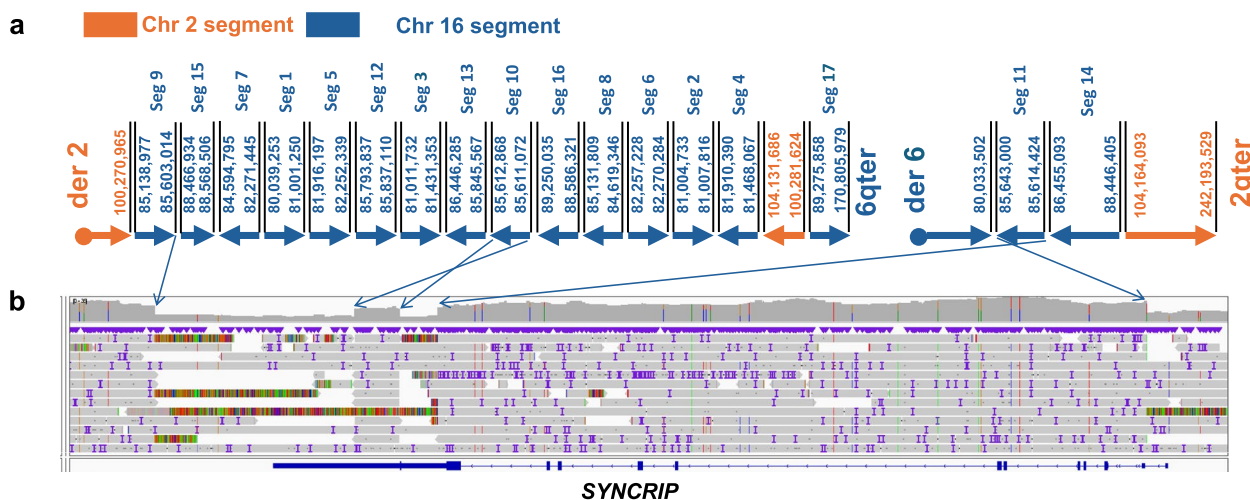
**Fig. 5** Schematic diagram of der(2) and der(6), and disruption of the NDD-associated gene *SYNCRIP*. **a** A schematic diagram illustrating the pattern of recombination after the DNA double strands break in chromosomes 2 and 6 by combining OGM and LRWGS analysis. **b** Integrative Genomics Viewer showed that the *SYNCRIP* gene was disrupted by four breakpoints

normal mitosis progression [35]; *GABRR2* is a member of a family of ligand-gated chloride channels that are the major inhibitory neurotransmitter receptors in the central nervous system [36]. *TTK* and *GABRR2* have not been reported to be associated with human disease. Taken together, the blood RNA-seq did not support that the expression of the NDD-related gene *SYNCRIP* was significantly altered but could not rule out the possibility of altered mRNA stability or protein translation due to the disrupted 5′ and 3′ UTR, leaving the pathogenicity of this SV uncertain.

## Discussion

ES is a first-tier molecular diagnostic test for individuals with NDDs [11, 37, 38]. It is adept at detecting sequence variants within the coding regions of genes and, when combined with specific CNV pipelines, can identify some SVs associated with copy number changes [25, 26]. However, when ES fails to provide a genetic diagnosis, there is often a lack of clear guidance on the next steps to be taken. Further genetic investigations should focus on variants not captured or inadequately covered by ES, including non-coding region variations, missed CNVs, and undetected SVs. Recently, a study using OGM in short-read genome sequencing negative cases of retinal diseases revealed that 25% of newly identified SVs disrupting disease-associated genes were previously overlooked [39]. Shieh et al. identified pathogenic or likely pathogenic SVs in 12% of 50 undiagnosed cases of rare monogenic disorders [40]. Iqbal et al. reported a yield of 4.5% (1/22) for pathogenic SVs in one cohort and 22.7% (5/22) for candidate SVs in another cohort [1]. Schrauwenet

al. found pathogenic or likely pathogenic SVs missed by ES in 10.6% of 47 unresolved NDD patients [41]. Therefore, we speculate that the undetected or overlooked SVs are important genetic contributors in ES-negative NDD cases.

In this study, 43 unsolved cases with NDDs were investigated by OGM, and RNA-seq was subsequently performed to further determine the functional impacts of candidate SVs. The clinical utility of combining OGM and RNA-seq to improve SV detection and interpretation was assessed in the ES-negative NDDs cohort. OGM identified three pathogenic SVs in our ES-negative NDDs cohort of 43 patients, achieving a diagnostic rate of 6.9%. If the two duplications that could be identified but were missed by previous clinical ES were excluded, the diagnostic rate would be reduced to 2.3%. The relatively low detection rate of pathogenic SVs in this study indicates that the proportion of pathogenic SVs in ES-negative cases with NDDs is not particularly high. SVs may not be a predominant etiological factor, especially when CNVs have been carefully analyzed together with routine exome analysis. Nevertheless, the inclusion of SVs analysis still contributes to solving ES-negative cases and improving the diagnostic outcome.

In our cases, SVs located in non-coding regions outside the coverage of ES were identified in one patient (P40). This highlights the potential importance of non-coding SVs in unsolved genetic cases. A heterozygous deletion of 134 kb affecting the promoter and 5′-untranslated exons of *MBD5*, outside the captured coverage of ES, were detected by OGM in this study. It is worth noting that although OGM can effectively detect SVs

Xiao *et al. Genome Medicine*    (2024) 16:113

Page 10 of 12

located in non-coding regions, the functional impact of SVs remains to be determined. In our study, RNA-seq confirmed approximately half of the reduction in *MBD5* expression, which is crucial for the accurate interpretation of the detected SV. This finding is consistent with previous reports documenting four cases with partial deletions of non-coding 5′-untranslated exons of *MBD5* [42, 43]. The region of deletion included the promoter and exon 1 of *MBD5*, which are probably essential for transcription initiation. As all the cases in the literature and in this study are associated with reduced mRNA expression of MBD5, the pathogenic mechanism is likely to be haploinsufficiency.

Another advantage of OGM is to provide "in situ" information, offering detailed insights into the location and orientation of duplication events that could not be achieved by ES. OGM has been reported to characterize the tandem location and orientation of a de novo heterozygous 13-kb duplication in *PUM1* and a 32-kb heterozygous de novo intronic duplication within the *NHEJ1* gene. These location and orientation information were readily available in the OGM results but would not be possible to obtain by ES-CNV analysis or traditional chromosomal microarray analysis [17, 40]. In another study by Jean et al., OGM pinpointed the specific number of extra copies in a multiplied region within *PAX5* and confirmed their presence as intragenic tandem multiplication [44]. Although the affected exon(s) may be readily identified using short read exome or genome sequencing, the location of the duplicated fragments(i.e., in tandem or in different genomic regions) is often unknown, which is critical for interpreting the pathogenicity of the duplication [18]. According to the American College of Medical Genetics guideline, only intragenic exonic duplications with deleterious coding consequences could be interpreted as likely pathogenic or pathogenic. Therefore, it is considered necessary to clarify the coding consequence of the duplication. In one study, several duplicated exons in the *DMD* gene were found to be non-contiguous and the *DMD* gene was intact based on OGM, which were reclassified as benign [45].

In our study, two small-sized SVs within genes were overlooked in previous ES-CNV analyses. Although the small duplications could be analyzed by ES, the pathogenicity was uncertain. OGM provided in situ information that these duplications were intragenic and located in tandem. RNA-seq analysis further confirmed the positioning of the duplicated exons by showing splice junctions consistent with the tandem duplication. Moreover, the presumed effect on RNA expression: out-of-frame duplication triggering nonsense-mediated decay (*PAFAH1B1* in P40) or in-frame duplication

preserving mRNA expression (*PLA2G6* in P2) was confirmed respectively. This combination of OGM and RNA-seq is an optimal approach to characterize structural duplications, providing valuable additional information on location, orientation and mRNA expression, thus improving the interpretation of SVs and clinical genetic counseling.

CCRs are thought to be underestimated causes of rare diseases that are often missed by routine genetic screening [46]. OGM can identify and delineate CCRs [47]. Fine characterization of CCRs is crucial for pinpointing breakpoints and identifying gene disruption events. Previous studies have shown that OGM refines breakpoints to reveal disease-causing genes [48, 49]. In P30 of this study, OGM effectively identified the CCR between chromosomes 2 and 6. LRWGS analysis based on OGM allowed us to identify breakpoints at single nucleotide resolution and to accurately construct the complete map of complex SVs. In a recent study [50], LRWGS initially found the CCR, but there were two possible ways to arrange the broken segments, and the final map of SVs was not fully elucidated. LRWGS with single-nucleotide resolution and OGM with longer reads (up to 2 M) were complementary for complex SVs analysis using long-molecule DNA [50]. Thus, combining LRWGS with OGM allowed us to accurately determine the complete structure of the CCR at the single nucleotide level. Moreover, RNA-seq analysis aided in determining the gene expression involved in candidate SVs. Although no known disease-causing genes associated with the patient's phenotype were identified in this study, this strategy may help to uncover novel NDD genes disrupted in the CCR that may be missed by routine genetic testing, such as *SNHG5*, which was found to have reduced RNA expression in this study and warrants further investigation.

Although OGM could reliably detect SVs, there are some limitations. OGM is a label-dependent technique that visualizes and analyzes genomic structure. While it offers high-resolution karyotyping, it is limited in resolving SVs lacking the specific DNA sequences required for labeling. In P2 and P4, where insertions have been detected with insufficient coverage of the corresponding region, it should be used in conjunction with other techniques to determine the inserted materials and their positions. In addition, OGM does not detect SNVs, and if an SV is found in a disease-causing gene that causes a recessive disorder, other sequencing technologies are needed to fully resolve the cases. Another limitation of the study is that the gene expression analysis based on peripheral blood may not fully reflect the affected tissues of patients with NDD.

## Conclusions

This study demonstrates the clinical utility of combining OGM and RNA-seq in ES-negative NDDs. The combination of OGM and RNA-seq allows a more comprehensive understanding of the impact of SVs at the transcript level. This synergy is particularly vital in unraveling the functional consequences of non-coding SVs, intragenic duplications and complex rearrangements. OGM is an ideal complement to ES and, when integrated with RNA-seq, provides a holistic and comprehensive approach to NDD diagnosis and genomic medicine.

## Abbreviations

| | |
|---|---|
| CCR | Chromosomal complex rearrangements |
| CNV | Copy number variation |
| ES | Exome sequencing |
| LRWGS | Long-read whole genome sequencing |
| MRI | Magnetic resonance imaging |
| NDD | Neurodevelopmental disorder |
| OGM | Optical genome mapping |
| SV | Structural variation |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-024-01382-9.

Additional file 1: Table S1. The detailed clinical data of 43 patients in our study cohort.

Additional file 2: Fig. S1. CMA analysis for P40 indicated the deletion involving partial *ORC4* and *MBD5*. Fig. S2. IGV visualization of WGS bam file in P2 and P4 suggests exons duplication. Fig. S3. Agarose gel electrophoresis for PCR products of P2 and P4 using primers in Table S2. Fig. S4. Sanger sequencing for P2 and P4 verified the breakpoints. Fig. S5. Genetic testing for patient 30. (a) G-banded karyotyping revealed a reciprocal translocation t(2:6)(q13;q15). (b) OGM results revealed multiple complex rearrangement events involving chromosomes 2 and 6. (c) A schematic diagram illustrating the pattern of recombination after the DNA double strands break of chromosomes 2 and 6. Fig. S6. The exact coordinates and nearby sequence of all the breakpoints in P30. Table S2. PCR primers for verification of the breakpoints in P2 and P4. Table S3. List of genes disrupted or deleted in the complex chromosomal rearrangements.

## Authors' contributions

The study was designed by YY, YF, and BX. Samples were collected by XL and YL. Optical genome mapping was analyzed by BX and XL. RNA-seq was analyzed by YF. Karyotyping was analyzed by HY. Exome sequencing data was analyzed by HL. The first draft of the manuscript was written by BX, and it was reviewed by YY, YF, and XL. All authors read and approved the final manuscript.

## Availability of data and materials

The information of four patients (P2, P4, P30, P40) described in Table 1 was uploaded to LOVD (https://databases.lovd.nl/shared/screenings). They were recorded in the database as individuals #00453475, #00453477, #00453478, and #00453480. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request. IGV visualization of LRWGS BAM file for all the 19 breakpoints involved in the complex rearrangements was uploaded and shown in depositories [31].

## Declarations

### Ethics approval and consent to participate

This study was conducted in accordance with the principles of the Declaration of Helsinki. This study was approved by the ethic board of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (XHCC-C-2022–077-1). Written informed consent for genetic analysis was obtained from the patients or their legal guardians.

### Consent for publication

Consent was obtained from the patients and their legal guardians for patient-related information that could be published.

### Competing interests

The authors declare no competing interests.

### Author details

[1]Department of Pediatric Endocrinology and Genetic Metabolism, Science and Education Building, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai Institute for Pediatric Research, Room 801, No.1665, Kong Jiang Road, Shanghai 200092, China.

## References

1. Iqbal MA, Broeckel U, Levy B, Skinner S, Sahajpal NS, Rodriguez V, et al. Multisite assessment of optical genome mapping for analysis of structural variants in constitutional postnatal cases. J Mol Diagn. 2023;25(3):175–88.
2. Mantere T, Neveling K, Pebrel-Richard C, Benoist M, van der Zande G, Kater-Baats E, et al. Optical genome mapping enables constitutional chromosomal aberration detection. Am J Hum Genet. 2021;108(8):1409–22.
3. Neveling K, Mantere T, Vermeulen S, Oorsprong M, van Beek R, Kater-Baats E, et al. Next-generation cytogenetics: comprehensive assessment of 52 hematological malignancy genomes by optical genome mapping. Am J Hum Genet. 2021;108(8):1423–35.
4. Levy B, Baughn LB, Akkari Y, Chartrand S, LaBarge B, Claxton D, et al. Optical genome mapping in acute myeloid leukemia: a multicenter evaluation. Blood Adv. 2023;7(7):1297–307.
5. Smith AC, Neveling K, Kanagal-Shamanna R. Optical genome mapping for structural variation analysis in hematologic malignancies. Am J Hematol. 2022;97(7):975–82.
6. Yang H, Garcia-Manero G, Sasaki K, Montalban-Bravo G, Tang Z, Wei Y, et al. High-resolution structural variant profiling of myelodysplastic syndromes by optical genome mapping uncovers cryptic aberrations of prognostic and therapeutic significance. Leukemia. 2022;36(9):2306–16.
7. Dai Y, Li P, Wang Z, Liang F, Yang F, Fang L, et al. Single-molecule optical mapping enables quantitative measurement of D4Z4 repeats in facioscapulohumeral muscular dystrophy (FSHD). J Med Genet. 2020;57(2):109–20.
8. Zhang S, Pei Z, Lei C, Zhu S, Deng K, Zhou J, et al. Detection of cryptic balanced chromosomal rearrangements using high-resolution optical genome mapping. J Med Genet. 2023;60(3):274–84.
9. Sahajpal NS, Mondal AK, Fee T, Hilton B, Layman L, Hastie AR, et al. Clinical validation and diagnostic utility of optical genome mapping in prenatal diagnostic testing. J Mol Diagn. 2023;25(4):234–46.

Xiao *et al. Genome Medicine*     (2024) 16:113

Page 12 of 12

10. Hu WF, Chahrour MH, Walsh CA. The diverse genetic landscape of neurodevelopmental disorders. Annu Rev Genomics Hum Genet. 2014;15:195–213.

11. Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. Genet Med. 2019;21(11):2413–21.

12. Gross AM, Ajay SS, Rajan V, Brown C, Bluske K, Burns NJ, et al. Copy-number variants in clinical genome sequencing: deployment and interpretation for rare and undiagnosed disease. Genet Med. 2019;21(5):1121–30.

13. Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol. 2019;20(1):117.

14. Sahajpal NS, Barseghyan H, Kolhe R, Hastie A, Chaubey A. Optical genome mapping as a next-generation cytogenomic tool for detection of structural and copy number variations for prenatal genomic analyses. Genes (Basel). 2021;12(3):398.

15. Saeidian AH, Youssefian L, Vahidnezhad H, Uitto J. Research techniques made simple: whole-transcriptome sequencing by RNA-Seq for diagnosis of monogenic disorders. J Invest Dermatol. 2020;140(6):1117-1126.e1.

16. Dekker J, Schot R, Bongaerts M, de Valk WG, van Veghel-Plandsoen MM, Monfils K, et al. Web-accessible application for identifying pathogenic transcripts with RNA-seq: Increased sensitivity in diagnosis of neurodevelopmental disorders. Am J Hum Genet. 2023;110(2):251–72.

17. Riquin K, Isidor B, Mercier S, Nizon M, Colin E, Bonneau D, et al. Integrating RNA-Seq into genome sequencing workflow enhances the analysis of structural variants causing neurodevelopmental disorders. J Med Genet. 2023;61(1):47–56.

18. Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med. 2020;22(2):245–57.

19. Brandt T, Sack LM, Arjona D, Tan D, Mei H, Cui H, et al. Adapting ACMG/AMP sequence variant classification guidelines for single-gene copy number variants. Genet Med. 2020;22(2):336–44.

20. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–24.

21. Moore S, McGowan-Jordan J, Smith AC, Rack K, Koehler U, Stevens-Kroef M, et al. Genome Mapping Nomenclature. Cytogenet Genome Res. 2023;163(5–6):236–46.

22. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

23. Yépez VA, Mertes C, Müller MF, Klaproth-Andrade D, Wachutka L, Frésard L, et al. Detection of aberrant gene expression events in RNA sequencing data. Nat Protoc. 2021;16(2):1276–96.

24. Brechtmann F, Mertes C, Matusevičiūtė A, Yépez VA, Avsec Ž, Herzog M, et al. OUTRIDER: a statistical method for detecting aberrantly expressed genes in RNA sequencing data. Am J Hum Genet. 2018;103(6):907–17.

25. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100.

26. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2):giab008.

27. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012;28(18):i333–9.

28. Smolka M, Paulin LF, Grochowski CM, Horner DW, Mahmoud M, Behera S, et al. Detection of mosaic and population-level structural variants with Sniffles2. Nat Biotechnol. 2024. https://doi.org/10.1038/s41587-023-02024-y. Epub ahead of print.

29. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. 2014;42(Database issue):D986-92.

30. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, et al. A genomic mutational constraint map using variation in 76,156 human genomes. Nature. 2024;625(7993):92–100.

31. Luo X. IGV visualization of LRWGS bam file for all 19 breakpoints involved in complex rearrangements in P30. Figshare. 2024. https://doi.org/10.6084/m9.figshare.26508550.v1.

32. Gillentine MA, Wang T, Hoekzema K, Rosenfeld J, Liu P, Guo H, et al. Rare deleterious mutations of HNRNP genes result in shared neurodevelopmental disorders. Genome Med. 2021;13(1):63.

33. Semino F, Schröter J, Willemsen MH, Bast T, Biskup S, Beck-Woedl S, et al. Further evidence for de novo variants in SYNCRIP as the cause of a neurodevelopmental disorder. Hum Mutat. 2021;42(9):1094–100.

34. Tanaka R, Satoh H, Moriyama M, Satoh K, Morishita Y, Yoshida S, et al. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. Genes Cells. 2000;5(4):277–87.

35. Fisk HA, Mattison CP, Winey M. Human Mps1 protein kinase is required for centrosome duplication and normal mitotic progression. Proc Natl Acad Sci U S A. 2003;100(25):14875–80.

36. Cutting GR, Curristin S, Zoghbi H, O'Hara B, Seldin MF, Uhl GR. Identification of a putative gamma-aminobutyric acid (GABA) receptor subunit rho2 cDNA and colocalization of the genes encoding rho2 (GABRR2) and rho1 (GABRR1) to human chromosome 6q14-q21 and mouse chromosome 4. Genomics. 1992;12(4):801–6.

37. Vetri L, Calì F, Saccone S, Vinci M, Chiavetta NV, Carotenuto M, et al. Whole exome sequencing as a first-line molecular genetic test in developmental and epileptic encephalopathies. Int J Mol Sci. 2024;25(2):1146.

38. Wayhelova M, Vallova V, Broz P, Mikulasova A, Smetana J, Dynkova Filkova H, et al. Exome sequencing improves the molecular diagnostics of paediatric unexplained neurodevelopmental disorders. Orphanet J Rare Dis. 2024;19(1):41.

39. de Bruijn SE, Rodenburg K, Corominas J, Ben-Yosef T, Reurink J, Kremer H, et al. Optical genome mapping and revisiting short-read genome sequencing data reveal previously overlooked structural variants disrupting retinal disease-associated genes. Genet Med. 2023;25(3):100345.

40. Shieh JT, Penon-Portmann M, Wong KHY, Levy-Sakin M, Verghese M, Slavotinek A, et al. Application of full-genome analysis to diagnose rare monogenic disorders. NPJ Genom Med. 2021;6(1):77.

41. Schrauwen I, Rajendran Y, Acharya A, Öhman S, Arvio M, Paetau R, et al. Optical genome mapping unveils hidden structural variants in neurodevelopmental disorders. Sci Rep. 2024;14(1):11239.

42. Ohori S, Tsuburaya RS, Kinoshita M, Miyagi E, Mizuguchi T, Mitsuhashi S, et al. Long-read whole-genome sequencing identified a partial MBD5 deletion in an exome-negative patient with neurodevelopmental disorder. J Hum Genet. 2021;66(7):697–705.

43. Mullegama SV, Rosenfeld JA, Orellana C, van Bon BW, Halbach S, Repnikova EA, et al. Reciprocal deletion and duplication at 2q23.1 indicates a role for MBD5 in autism spectrum disorder. Eur J Hum Genet. 2014;22(1):57–63.

44. Jean J, Kovach AE, Doan A, Oberley M, Ji J, Schmidt RJ, et al. Characterization of PAX5 intragenic tandem multiplication in pediatric B-lymphoblastic leukemia by optical genome mapping. Blood Adv. 2022;6(11):3343–6.

45. He W, Meng G, Hu X, Dai J, Liu J, Li X, et al. Reclassification of DMD duplications as benign: recommendations for cautious interpretation of variants identified in prenatal screening. Genes (Basel). 2022;13(11):1972.

46. Schuy J, Grochowski CM, Carvalho CMB, Lindstrand A. Complex genomic rearrangements: an underestimated cause of rare diseases. Trends Genet. 2022;38(11):1134–46.

47. Qu J, Li S, Yu D. Detection of complex chromosome rearrangements using optical genome mapping. Gene. 2023;884:147688.

48. Schnause AC, Komlosi K, Herr B, Neesen J, Dremsek P, Schwarz T, et al. Marfan syndrome caused by disruption of the FBN1 gene due to a reciprocal chromosome translocation. Genes (Basel). 2021;12(11):1836.

49. Orlando V, Di Tommaso S, Alesi V, Loddo S, Genovese S, Catino G, et al. A complex genomic rearrangement resulting in loss of function of SCN1A and SCN2A in a patient with severe developmental and epileptic encephalopathy. Int J Mol Sci. 2022;23(21):12900.

50. Ohori S, Numabe H, Mitsuhashi S, Tsuchida N, Uchiyama Y, Koshimizu E, et al. Complex chromosomal 6q rearrangements revealed by combined long-molecule genomics technologies. Genomics. 2024;116(5):110894.

## Publisher's Note