Computational and Structural Biotechnology Journal 27 (2025) 2233–2242

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

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

Research Article

Optical genome mapping reveals novel structural variations in an autism spectrum disorder cohort

Yunjia Zhang [a,b,1], Wai-Tong Chien [c,d,e,1], Judy Yuet-Wa Chan [c], Tsz-Yan Cheung [a,b], Zhiqian He [a], Kai-Hang Yip [a], Dorothy Fung-Ying Chan [f], Josephine Shuk-Ching Chong [f,g], Mary Miu-Yee Waye [c,d,e], Sek-Ying Chair [c,d,e,*], Ting-Fung Chan [a,b,**]

[a] School of Life Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, China
[b] The State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong, China
[c] The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China
[d] Asia-Pacific Genomic and Genetic Nursing Centre, The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China
[e] The Croucher Laboratory for Human Genomics, The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China
[f] Department of Paediatrics, The Chinese University of Hong Kong, Hong Kong, China
[g] Joint Baylor-CUHK Center of Medical Genetics, The Chinese University of Hong Kong, Hong Kong, China

ARTICLE INFO

Keywords:
Optical genome mapping
Autism spectrum disorder
Structural variants
Genetic architecture

ABSTRACT

Structural variants (SVs) contribute to the genetic architecture of autism spectrum disorder (ASD), but their comprehensive characterization is limited by technological challenges in their detection. Optical genome mapping (OGM) offers a promising alternative, enabling the identification of large-scale SVs that might be overlooked by traditional sequencing methods. This study aimed to use OGM to identify SVs associated with ASD. We generated in-house OGM data from 26 participants diagnosed with ASD, leading to the discovery of 1593 novel SVs. Among them, 114 novel SVs were identified in at least two non-sibling participants, with 57 of them putatively overlapping known gene regions. To validate our findings, two novel SVs were confirmed by Sanger sequencing. The dataset generated in this study can serve as a novel and valuable resource for future research and facilitate the exploration of SVs related to ASD. Our work also underscores the importance of large-scale genomic rearrangements in neurodevelopmental disorders and provides insights into SVs as potential molecular diagnostic and therapeutic targets for ASD.

1. Introduction

Autism spectrum disorder (ASD), also known as autism, is a heterogeneous neurodevelopmental condition characterized by deficits in social interaction and communication and the presence of repetitive behaviors [1]. Several comorbid medical conditions, psychiatric disorders, and behavioral and motor dyscontrol symptoms, such as attention deficit hyperactivity disorder (ADHD), intellectual disability, obsessive-compulsive disorder (OCD), and dyslexia, are associated with ASD [2,3]. Its prevalence has increased over time and has varied greatly within and across sociodemographic groups. Approximately 1/100 children are diagnosed with ASD worldwide [4], with a male-to-female ratio of approximately 4:1 [5]. The increase in the prevalence of ASD is likely associated with changes in the diagnostic criteria, improved performance of screening and diagnostic tools, and increased public awareness [6]. In Hong Kong, the prevalence was 16.1 per 10,000 among children aged less than 15 years, which is much lower than the worldwide statistical data [7]. However, the number of students diagnosed with ASD has been increasing rapidly [8]. There are currently no reliable biomarkers for ASD [9]. Therefore, diagnosis is conducted mainly using observational screening tools such as the Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition (DSM-5) and Autism Diagnostic Observation Schedule (ADOS) [10]. ASD treatment mainly focuses on behavioral training aimed at reconditioning target behaviors and developing vocational, social, cognitive, and living skills [11].

* Corresponding author at: The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China.
** Corresponding author at: School of Life Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, China.
E-mail addresses: sychair@cuhk.edu.hk (S.-Y. Chair), tf.chan@cuhk.edu.hk (T.-F. Chan).
[1] Co-first authors

https://doi.org/10.1016/j.csbj.2025.05.043
Received 23 February 2025; Received in revised form 26 May 2025; Accepted 26 May 2025
Available online 27 May 2025
2001-0370/© 2025 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

The etiology of ASD has been reported to be influenced by genetic factors. Infants born into families with one or more affected older siblings have an estimated 18.7 % elevated risk of developing ASD [12]. Furthermore, it is estimated that 74–93 % of the associated ASD risk is attributable to heritable genetic factors [13], with up to 1000 genes considered to be potentially implicated in the disorder [14]. To date, more than 100 ASD risk genes and genomic regions have been identified and confirmed, underscoring the substantial heritability of this disorder [1].

Studies have consistently recognized the intricacy and diversity of genetic variations in ASD risk architecture and reported valuable findings, particularly in the realm of ASD-associated single nucleotide polymorphisms (SNPs). For instance, an umbrella review of meta-analyses of most ASD genetic studies identified that at least 12 SNPs within various ASD risk genes have robust significance in ASD [15]. Inherited protein-truncating SNPs are also observed to be enriched in ASD probands compared with their unaffected siblings [16].

In addition to SNPs, studies have acknowledged that structural variations are associated with ASD in a more diverse and complex fashion than previously known [17]. In contrast to millions of SNPs, a human genome typically comprises an estimated 2500 structural variants (SVs) that affect around 20 million nucleotides, exhibiting a remarkably greater impact on the aberrant expression of genes and phenotypic changes than SNPs, despite their lower prevalence [18–20]. Theoretically, SVs could lead to the loss of function or gene dosage effect of protein-coding genes through the deletion of nucleotides or alteration of open-reading frames and gain of gene copies by duplication, respectively [21,22]. Rare copy number variants (CNVs) have been found to be present in 5–10 % of all ASD cases [23], with a significantly increased global burden of rare CNVs in ASD-affected individuals compared with control individuals [24]. ASD probands are also observed to exhibit an enrichment of both inherited rare CNVs and *de novo* CNVs compared with their unaffected siblings, and this enrichment is thought to contribute to an elevated risk of ASD [16,25]. Among Chinese, a systematic review including 9 case-control studies reported SVs in genes associated with ASD, with all CNVs involving deletions and duplications [26]. Apart from protein-coding genes, rare inherited noncoding SVs of smaller size also contribute to neurodevelopmental disorder pathogenesis by altering regulatory element positions and chromatin architecture, thereby influencing gene regulation during neurodevelopment and predisposing individuals to ASD. [22,27–29].

Despite the evident link between ASD and SVs, the majority of genetic research endeavors in ASD have focused on SNPs [15], with large genomic SVs being comparatively neglected and understudied [23]. However, considering the polygenic nature of ASD etiology, where individual genetic variants substantially contribute to ASD risk via additive genetic effects [30], studies centered on single ASD mutations may provide limited insights into the broader ASD landscape. Conversely, large genetic variants like SVs, which have greater potential to disrupt or rearrange multiple functional elements in the genome [29], are more likely to offer a holistic perspective on how genetics contribute to ASD risk. Therefore, a greater emphasis on investigating SVs in ASD research is pressing, as a more comprehensive understanding of the disorder's genetic basis might be achieved.

To identify potential ASD-associated SVs, two types of strategies, namely hybridization-based and sequencing-based approaches, have been commonly used for SV detection [31,32]. Hybridization-based approaches, such as array comparative genomic hybridization, allow high-throughput analysis of several types of SVs for routine clinical diagnosis but cannot detect balanced SVs, including translocation and inversion [33]. With the resolution of a few kilobases, this type of detection method has limited sensitivity to smaller-sized SVs and is not able to precisely define the breakpoints and positions of the variants [31,34]. With advances in sequencing technologies, SVs could be detected by sequencing-based approaches with a higher resolution, i.e., single-base resolution [35]. In principle, whole-genome sequencing could detect all types of SVs, but this remains a challenge due to the limited length of sequencing reads and the repetitive regions across the human genome [36,37]. Second-generation sequencing generates high-throughput data, but the short read length (300 bp) restricts the analysis of large-scale SVs in the complex human genome, which contains plenty of repetitive regions [37]. In contrast, third-generation sequencing, namely Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT), generates longer reads (15 kb) but lower throughput than other sequencing platforms, limiting the discovery of the entire spectrum of SVs [38].

Optical genome mapping (OGM), a light microscope-based imaging technology, fluorescently labels specific sequence motifs of DNA molecules, captures images of the labelling patterns, and generates single-molecule optical maps [37,39]. Although OGM does not provide nucleotide-level information, it can overcome the problems associated with sequencing-based detection methods by generating high-throughput data with an average molecule length of 225 kb, allowing it to easily span repetitive and complex genomic regions and facilitating the detection of large-scale SVs that are missed by sequencing [36,38–40]. Recent studies have demonstrated that OGM outcompetes both short-read and long-read sequencing in capturing SVs. For instance, OGM was able to detect duplications and inversions that were missed by whole-genome sequencing, including both current long-read platforms from PacBio and ONT [41]; optical maps resolved more than 1000 nonredundant SV clusters, which were not detected in the PacBio phased assemblies [42]; and a more complex SV pattern showing an inversion with inverted segmental duplication was identified by optical maps compared with the single deletion detected by PacBio assemblies [42]. Therefore, in this study, we investigated the genome-wide SVs of 26 participants with ASD using OGM to achieve a comprehensive SV profile of the human genome and further unravel the etiology of genetic disorders related to ASD. A schematic overview of the workflow of this study is illustrated in Fig. 1.

2. Materials and methods

2.1. Participants selection and blood collection

This study was approved by The Joint Chinese University of Hong Kong - New Territories East Cluster Clinical Research Ethics Committee (Reference number: 2022.425) and registered at International Standard Randomised Controlled Trial Number (ISRCTN79375633). Autistic participants diagnosed by medical doctors in the Department of Paediatrics at Prince of Wales Hospital were recruited. The inclusion criteria were participants who had received a valid assessment and met cutoffs for autism spectrum or autism (e.g. the newest ADOS algorithms to be used for Modules 1–3 and the original cutoff algorithms to be used for Module 4 or a clinical "Best Estimate Diagnosis" of Autistic Disorder, Asperger's Disorder, or PDD-NOS according to the DSM-IV-TR) and were of Chinese ethnicity regardless of age and gender. Individuals with a major injury, abnormality, or disorders that had effects on the brain; extensive complications during birth or pregnancy; considerable nutritional and psychological deprivation; a known genetic disorder; or a psychiatric disorder requiring medication were excluded. During recruitment, 47 subjects who met the criteria were approached. Nineteen subjects were refused and twenty-eight were recruited. All the recruited participants were of Chinese ethnicity in Hong Kong.

For those subjects who agreed to participate in the study, their parents were asked to provide signed informed consent and fill out a questionnaire about the participants' family history of ASD, environmental exposure, medication, ASD-associated symptoms, and other demographic data. Whole blood samples (5–10 mL) were collected from the participants and frozen at −80°C until the genomic DNA extraction.

From December 2022 to March 2024, 28 participants with autism were recruited, including two pairs of twin brothers (C16 and C22 pairs). OGM analysis and data statistics were conducted exclusively on

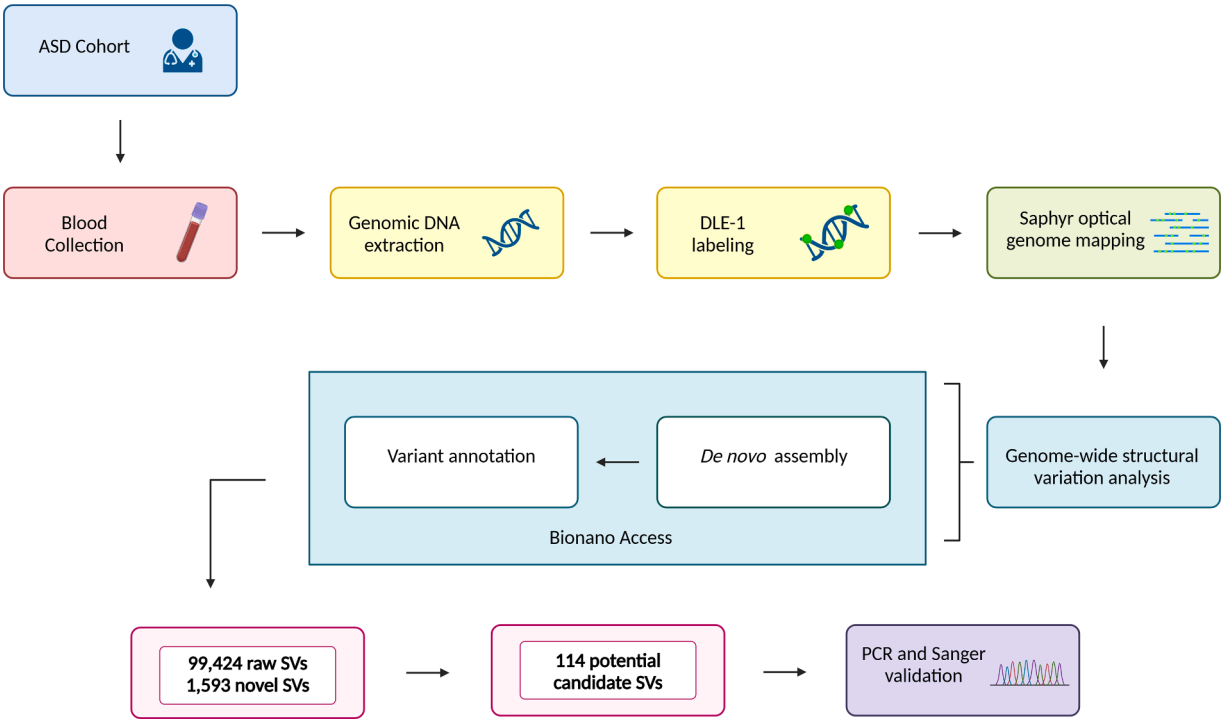

Fig. 1. Workflow of OGM for SV detection in the ASD cohort.

one individual from each twin pair. Thus, all statistical results reported in this paper are based on a sample size of 26. The demographic characteristics of the participants are shown in Table 1. The male-to-female ratio was 4.2:1, which is close to the 2020 global male-female prevalence estimate of 4:1 [5]. The participants' ages ranged from 3 to 27 years with the average age being 6.6 ± 5.0 years, and the mean age at diagnosis was 2.9 ± 1.3 years. The participants with autism were also reported to have ADHD (34.6 %), delayed development (11.5 %) and dyslexia (3.8 %). According to the Centers for Disease Control and Prevention, people with ASD often have problems with social communication and interaction, in addition to restricted or repetitive behaviors or interests [43]. Most of our participants exhibited common signs and symptoms of autism, such as no eye contact (76.9 %), poor conversation (69.2 %), poor social skills (69.2 %), not sharing interests (69.2 %), and repetitive behavior (65.4 %) and narrow interests (65.4 %). Overall, 34.6 % of the participants had at least one sibling diagnosed with autism or another mental/developmental disorder.

2.2. Optical genome mapping

Ultra-high molecular weight (UHMW) DNA was extracted from the frozen whole blood samples using the SP Blood and Cell Culture DNA Isolation Kit (Bionano Genomics, Cat No. 80042) or the SP-G2 Blood & Cell Culture DNA Isolation Kit (Bionano Genomics, Cat No. 80060) according to the manufacturer's protocol. The number of white blood cells (WBCs) was counted by HemoCue WBC Analyzer (Fisher Scientific, Cat No. 22–601–017). For each sample, 1.5 million WBCs were used for the extraction of UHMW DNA. Briefly, the WBCs were lysed and digested, followed by the binding of genomic DNA to the Nanobind disk. The extracted UHMW DNA was quantified by Qubit Fluorometer (Thermo Fisher Scientific), and 750 ng of UHMW DNA was then fluorescently labelled using the Direct Label and Stain (DLS) Kit (Bionano Genomics, Cat No. 80005) following the Bionano Prep DLS Protocol (Bionano Document No. 30206, Rev G). Briefly, the UHMW DNA molecules were labeled with DL-green fluorophore using the Direct Label Enzyme (DLE-1) to generate sequence-specific patterns, followed by overnight fluorescent staining of DNA backbones to visualize the full length of the DNA molecules. The labelled UHMW DNA was then loaded onto Saphyr Chip G1.2 (Bionano Genomics, Cat No. 20319) and run on the Bionano Saphyr system (Bionano Genomics) for 36 h.

De novo assembly and variant annotation were performed using Bionano Solve software V3.8.1 (Bionano Genomics, Inc.). The analysis used the hg38 reference genome with the Y chromosome pseudoautosomal regions (Y-PAR) masked. Filtering was conducted with the minimum confidence value: insertion/deletion/inversion = 0.8, duplications = -1, intra-fusion/inter-translocation = 0.05, CNV = 0.8, and aneuploidy = 0.8. Filtered variants were cross-compared with the Bionano control human samples database (Bionano Genomics, Inc.). Duplicate SVs between individuals were filtered, and novel SVs absent from the database of 285 phenotypically normal individuals were retained for further analysis (Bionano Solve Theory of Operation; CG-30190). These SVs were then compared against participant information and the annotated genes in hg38. Novel SVs that overlapped with those of non-sibling participants were ultimately selected.

2.3. Breakpoint polymerase chain reaction (PCR) and Sanger sequencing

The *MAU2* and *ZFHX3* genes were selected for breakpoint PCR and Sanger sequencing confirmation. The following unique primers were used for PCR amplification of the entire SV regions detected by OGM: *MAU2*_SV (forward 5′-AGACTCATCCAATCCCACAATGCT-3′, reverse 5′-GGACATTTAATGGATGATTCTTTCAAAGTTAGTTGACT-3′) and *ZFHX3*_SV (forward 5′-CCCTAAAGTCATAAGGTTTTGGCACAAG-3′, reverse 5′-AAGGCTTGGGGACCCAATGA-3′). PCR reactions were conducted using Q5® Hot Start High-Fidelity 2 × Master Mix (New England BioLabs, Cat No. M0494S) with 0.8 μL of template DNA and 0.5 μL of 10 μM primers under the following conditions: 98°C for 30 s (initial denaturation), followed by 45 cycles of 98°C for 10 s, 65°C for 30 s, and 72°C for 3 min 30 s, with a final extension at 72°C for 2 min. The PCR products were separated by gel electrophoresis. For samples with heterozygous deletions, bands representing deletion (smaller in size) and wild-type bands were excised separately. All excised bands were purified using the PureLink™ Quick Gel Extraction Kit (Invitrogen™, Cat No. K210012). Subsequently, Sanger sequencing was performed starting from the beginning of the fragment

Table 1
Demographic characteristics of the participants.

Characteristics	N = 26
Sex	
Male	21 (80.8 %)
Female	5 (19.2 %)
Age (years)	
Range	3–27
Average	6.6
Father's age (years)	
Range	20–52
Average	36.0
Mother's age (years)	
Range	19–37
Average	31.7
Age at diagnosis (years)	
Range	1.5–7
Average	2.9
Other diagnosed mental/developmental diseases	
Attention deficit hyperactivity disorder	9 (34.6 %)
Delayed development	3 (11.5 %)
Dyslexia	1 (3.8 %)
Family history	
Have siblings with autism	8 (30.8 %)
Have siblings with other diagnosed mental or developmental diseases	1 (3.8 %)
Mother with diagnosed mental/developmental diseases	2 (7.7 %)
Father with diagnosed mental/developmental diseases	0
Signs and symptoms of ASD	
Poor conversation	18 (69.2 %)
Poor social skills	18 (69.2 %)
No eye contact	20 (76.9 %)
Poor nonverbal skills	14 (53.8 %)
Not sharing interests	18 (69.2 %)
Abnormal sensory sensitivities	13 (50 %)
Repetitive behavior	17 (65.4 %)
Reluctance for change	15 (57.7 %)
Narrow interests	17 (65.4 %)
Hyperactivity	15 (57.7)
Anxiety	9 (34.6 %)
Cognitive impairment	13 (50 %)
Delayed development	16 (61.5 %)
Poor coordination	9 (34.6 %)
Sleep problem	7 (26.9 %)
Seizure	2 (7.7 %)
Attack or self-harm	8 (30.8 %)

until the breakpoint was detected, and the primer sequences covering the breakpoint were *MAU2*_sanger (5′- GCGGGCTCCTTGCAAGATCTGGGCT-3′) and *ZFHX3*_sanger (5′- CCCTAAAGTCATAAGGTTTTGGCACAAG-3′). The Sanger sequencing results were mapped to the human genome hg38 using Minimap2 v2.28-r1209 [44] and converted to BAM files using SAMtools v1.15.1 [45]. BAM files were visualized using Integrative Genomics Viewer (IGV) v2.12.2 [46].

3. Results

3.1. OGM reveals novel SVs in the ASD cohort

Saphyr optical mapping molecules that were 150 kbp or longer with at least nine DLE-1 labels were used in the subsequent analysis. The molecules had an average N50 of 258.8 kbp, with a range of 200.6 kbp to 315.8 kbp across all participants (recommended: > 150 kb). On average, 92.3 % of the molecules were successfully mapped to the reference genome, with percentages ranging from 81.7 % to 95.2 %. The average amount of total DNA collected from the participants was 1253.8 Gbp, ranging from 200.6 Gbp to 1472.2 Gbp. The average effective coverage of usable data was 374.41X, with a range of 117.5X to 446.7X (recommended 80X) (Table 2).

In total, 99,424 raw SVs were identified, encompassing deletions, duplications, insertions, inversions, and translocations, with sizes ranging from 500 bp to 6 Mbp. Among the raw SVs, 1593 novel SVs did not overlap with the control database, which ranged in size from 500 bp to 6 Mbp. The size distribution of both raw SVs and novel SVs is illustrated in Fig. 2a-b. SVs larger than 80 kbp (3 %) are omitted from the figures for clarity and readability. The majority of both novel and raw SVs were individual-specific SVs. Among raw SVs, a substantial portion also represented common SVs that occurred in multiple participants. Among novel SVs, apart from individual-specific SVs, the number of identical SVs among siblings was notably higher than that among non-sibling participants (Fig. 2d-f). 51 novel SVs were exclusively shared by one pair of siblings. Detailed information on all raw and novel SVs is provided in Supplementary Tables 1 and 2, respectively.

Table 2
Statistics of raw OGM molecules and detected SVs.

Characteristics	N = 26
N50 (kbp)	258.8
Mapping rate (hg38)	92.3 %
Throughput (Gbp)	1253.8
Maximum molecule length (Mbp)	2.8
Effective coverage (X)	374.4
Raw SVs	99,424
Novel SVs (0 % overlapping with control database)	1593
Novel SVs with overlapping between non-siblings	114
Novel SVs with overlapping between non-siblings overlapping genes	57

Among the 1593 novel SVs, to eliminate the impact of familial relationships, novel SVs that overlapped between non-sibling participants were selected as potential ASD-associated candidate SVs. Altogether, 114 novel SVs were identified in at least two non-sibling participants, 57 of which putatively overlapped known gene regions (Fig. 2f). Around half of the detected SVs overlapped genes in different categories (Fig. 2c). Among the 114 identified novel SVs, there were 47 deletions, 36 insertions, 19 inversions, and 12 duplications. Fourteen of these novel SVs (12.3 %) were shared not only within siblings but also with non-sibling participants. Among all human chromosomes, chromosome 12 contained the highest number of novel SVs, amounting to 14. In contrast, no novel SVs were detected on chromosomes 13, 17, 22, and the Y chromosome. As a potential response to the gender bias in ASD, seven novel SVs were detected on the X chromosome, which had the sixth highest number of novel SVs among all chromosomes (Fig. 3). The detailed information regarding all 114 and 57 novel SVs is provided in Supplementary Table 3 and Table 3, respectively.

3.2. Validation of novel SVs in MAU2 and ZFHX3 genes

Two novel SVs, *MAU2*_SV and *ZFHX3*_SV, overlapping *MAU2* and *ZFHX3* genes, respectively, were selected for breakpoint PCR and Sanger sequencing confirmation. The selection was based on four criteria: prevalence among participants, the correlation of their overlapping genes with ASD, recurrence among sibling participants, and suitable size for reliable breakpoint PCR amplification. Five initial candidates were selected for PCR, three of which failed due to highly repetitive nearby sequences. Eventually, *MAU2*_SV and *ZFHX3*_SV were successfully validated by Sanger sequencing.

The MAU2 protein is an essential protein for the assembly of a heterodimeric cohesin loader (also known as the kollerin complex) with the NIPBL protein [47]. This cohesin loader complex enables the attachment of cohesin to DNA, thereby contributing to various chromatin-related processes such as sister chromatid cohesion, DNA repair, transcriptional regulation, and chromatin structural organization [48]. Recently, deletions in *MAU2* have been suggested to disrupt the heterodimerization of the cohesin loader subunits and cause Cornelia de Lange syndrome (CdLS) [47]. CdLS is a rare congenital developmental disorder characterized by psychomotor delay and/or intellectual disability and is usually associated with ASD [49]. SVs overlapping the *MAU2* gene have been noted in individuals with neurodevelopmental disorders (NDD). *De novo* microdeletions ranging from 0.7 to 5.2 Mb have been detected in individuals with NDD, frequently overlapping the chromosomal region

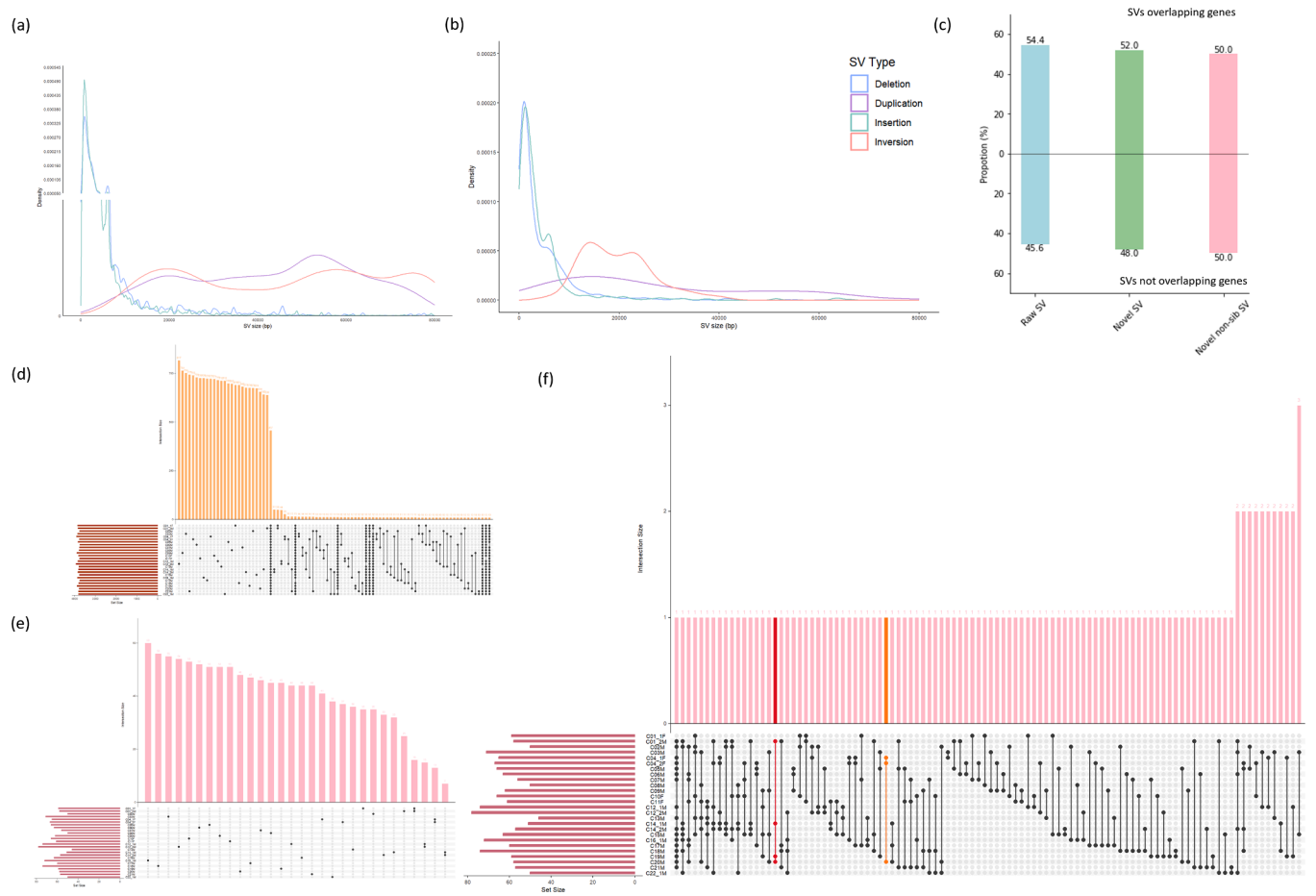

Fig. 2. Identification and characterization of SVs detected in the ASD cohort. (a)Size distribution of raw SVs detected in 26 participants with ASD (N = 99,061). SVs larger than 80 kbp (363) are omitted from the figure. (b) Size distribution of raw SVs detected in 26 participants with ASD (N = 1539). SVs larger than 80 kbp (54) are omitted from the figure. (c) Proportions of SVs relative to known genes in different categories: raw SVs (N = 99,424), novel SVs (N = 1593), and novel SVs in at least two non-sibling participants (N = 114). (d) Upset plot showing the overlap of raw SV positions across 26 samples. Sets with overlap count less than 10 are omitted from the figure. (e) Upset plot showing the overlap of novel SV positions within individuals or between siblings. (f) Upset plot showing the overlap of novel SV positions between non-sibling participants. The sets representing *MAU2* and *ZFHX3* are highlighted in red and orange, respectively.

19p13.11p12, which includes the *MAU2* gene [50]. Despite the association with neurodevelopmental processes, there have been no reports on the direct relationship between SV in *MAU2* and ASD.

Our optical mapping results revealed *MAU2*_SV as a novel heterozygous deletion, with an estimated size of 941 bp, in four male participants with ASD (15.4 % of all participants, 19.0 % of all male participants) but not in the control database. Both gel electrophoresis and Sanger sequencing results supported the optical mapping prediction, ultimately confirming a deletion of 964 bp located at chr19: 19321705–19322670. All four participants exhibited the same breakpoint, with the deletion encompassing 198 bp of the exon 1 region (5′-UTR) and 766 bp of the intron 1 region of *MAU2* (Fig. 4a).

ZFHX3, alternatively known as AT motif-binding transcription factor 1, encodes a zinc-finger homeodomain transcription factor [51], which is classified within the C2H2-type zinc finger protein family [52]. *ZFHX3* has been identified to be a key regulator of various biological processes, encompassing embryonic development, cell proliferation, tumorigenesis, neuronal differentiation, and neuronal death [51,53]. With its ubiquitous expression, notably prominent in the developing brain [53], *ZFHX3* is significantly involved in neurodevelopment by regulating genes associated with neuron development [51]. Chromatin immunoprecipitation sequencing of *ZFHX3* coupled with subsequent functional enrichment analysis and gene ontology analysis has revealed its predominant binding to the promoters of genes that engage in neurodevelopmental processes such as axonogenesis, axon development, and neurogenesis [51]. The proteins encoded by these neuron-related genes modulate neurodevelopment through diverse pathways, including the Wnt, Hippo/YAP, and mTOR signaling pathways [51,53].

Given its important role in neural development, *ZFHX3* has been implicated in the pathogenesis of multiple neurodevelopmental disorders. Loss-of-function variations in *ZFHX3* have been identified as the underlying cause of syndromic intellectual disability [51]. Furthermore, *de novo* variants of *ZFHX3* have occasionally been detected in individuals with neurodevelopmental disorders, such as developmental disorder and developmental and epileptic encephalopathy [53]. Studies have revealed a potential association between *ZFHX3* and ASD, with multiple *ZFHX3* variants being identified in individuals with ASD [51, 53–55]. Based on the neurodevelopmental disorder risk gene prediction tool mantis-mL, *ZFHX3* is among the top 5 % of genes associated with monoallelic ASD and/or developmental delay [51].

Our optical mapping data revealed a novel SV in the intron region of *ZFHX3*. *ZFHX3*_SV is a novel heterozygous deletion, with an estimated size of 2171 bp, that existed in two sisters and one male participant in our study (11.5 % of all participants, 4.8 % of all male participants, 40.0 % of all female participants) but not in the control database. Both gel electrophoresis and Sanger sequencing results supported the optical mapping prediction as well, confirming a deletion of 2207 bp (chr16: 73389202–73391408) in the intron 3 of *ZFHX3*. All four participants also exhibited the same breakpoint (Fig. 4b).

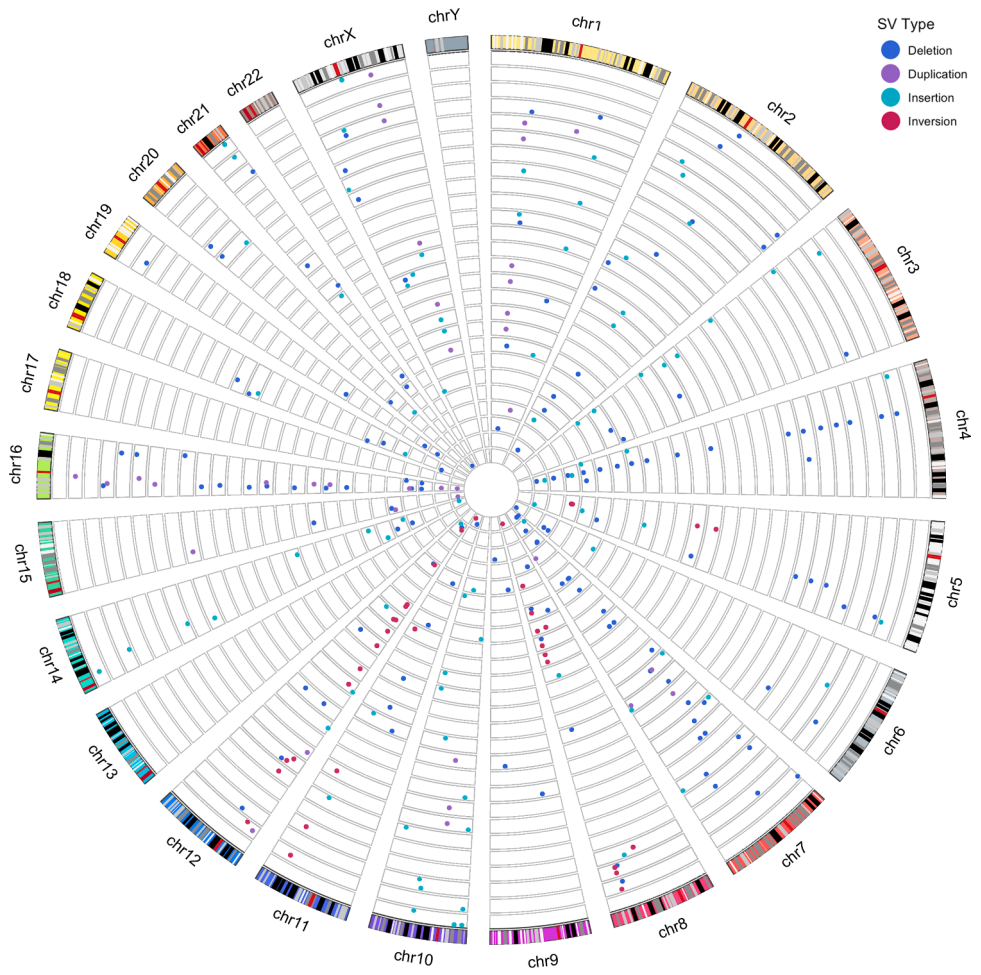

Fig. 3. Circos plot illustrating the locations of the 114 novel SVs existing in at least two non-sibling participants.

4. Discussion

In this study we used OGM to identify structural variations in an ASD cohort, underscoring the importance of large-scale genomic rearrangements in neurodevelopmental disorders and offering insights into SVs as potential molecular diagnostic and therapeutic targets for ASD. Our findings underscore the potential involvement of novel genes, including *MAU2* and *ZFHX3*, in ASD pathogenesis, potentially through disruptions in neurodevelopmental regulatory mechanisms. These discoveries not only expand our understanding of the genetic architecture underlying ASD but also highlight new avenues for exploring its etiology.

The validation of structural variations by Sanger sequencing reinforces the reliability of OGM as a robust method for detecting ASD-associated SVs. This accuracy is crucial, given the challenges in identifying and validating pathogenic variants due to the heterogeneity of ASD and the difficulties in recruiting well-characterized cohorts [56]. While our study yields perspectives on the involvement of SVs in ASD, it also emphasizes the necessity for further validation in larger and more diverse cohorts to ensure the generalizability of these findings. During participant recruitment, we encountered challenges due to the limited research on SVs in ASD and the prospective nature of OGM for this field. A notable obstacle was the lack of familiarity among parents with these concepts, which led to refusal of blood collections. In addition, due to social limitations in communication, blood draws became a stressful event for both participants and parents. This underscores the urgency of our research in facilitating participant recruitment and advancing the study of SVs in ASD.

The identification of recurrent SVs overlapping known gene regions suggests that structural variations may play a significant role in ASD by disrupting critical neurodevelopmental pathways. Specifically, the involvement of *MAU2* and *ZFHX3* genes points to potential mechanisms by which these variants could contribute to the disorder. Future research should aim to expand the scope of investigations through larger cohort studies, integration of whole-exome sequencing to evaluate SV impacts at the transcriptome level, and functional validation using animal models to characterize downstream biological effects. Given the male bias in ASD, sex-linked SVs warrant investigation as well. Among the 57 novel SVs we identified, 35 were found only in male participants, including the validated *MAU2*_SV. However, due to the difficulty in recruiting ASD participants, especially females, our cohort had limited female participants (N = 5), which was insufficient to draw accurate sex-based conclusions. Nevertheless, the SVs reported in this study can serve as evidence in future sex-linked studies in larger ASD cohorts. Family-based recruitment is also promising to study the inheritance of SVs related to ASD. For instance, in our study, we prioritized recruiting participants with ASD-diagnosed siblings, successfully enrolling 4 sibling pairs (8 individuals) out of 26 total participants. Notably, *ZFHX3*_SV was detected in a pair of sisters, suggesting its potential hereditary pattern. These efforts will be instrumental in establishing a more comprehensive understanding of how structural variations contribute to ASD pathology and in identifying novel diagnostic and therapeutic targets. By addressing the challenges associated with participants recruitment and sample heterogeneity, researchers can build upon this foundation to uncover new insights into the genetic and biological underpinnings of ASD.

Table 3
Characteristics of the 57 novel SVs identified in at least two non-sibling participants and overlapping genes.

Participants	Chr	Start	End	Type	SV size (bp)	Overlap Genes	Positive participant number
C03M/C08M/C11F/C13M/C14_2M/C20M	16	18909935	18924619	deletion	6136/6169/6152/6143/6135/6117	*SMG1*	6
C01_2M/C03M/C06M/C14_2M/C22M	11	108694034	108695145	inversion_paired	24250/24250/24250/24250/24250	*DDX10*	5
C01_2M/C03M/C06M/C14_2M/C22M	11	108713881.5	108723799	inversion_paired	24250/24250/24250/24250/24250	*DDX10*	5
C01_2M/C14_1M/C19M/C20M	19	19320522	19326334	deletion	954/934/940/934	*MAU2*	4
C09M/C12_1M/C15M/C19M	16	22478974.5	22491631	deletion	504/531/504/509	*SMG1P1*	4
C03M/C09M/C15M/C20M	1	65598245	65614302	deletion	542/546/577/538	*LEPR*	4
C01_2M/C17M/C18M/C21M	2	80005448	80012015	deletion	668/616/683/705	*CTNNA2*	4
C04_2F/C07M/C10F/C15M	1	184428737	184432814	insertion	674/702/710/684	*C1orf21*	4
C10F/C16_1M/C21M	18	6104531	6116210	deletion	2518/2421/2435	*L3MBTL4*	3
C13M/C16_1M/C17M	3	6698131	6707261	insertion	572/549/609	*GRM7-AS3*	3
C01_2M/C15M/C18M	4	7831067	7841016	insertion	612/599/569	*AFAP1*	3
C07M/C11F/C21M	7	8258610	8259674	insertion	981/911/972	*ICA1*	3
C07M/C18M/C22M	15	20628640	20689014	duplication_inverted	60374/60374/60374	*NBEAP1*	3
C03M/C05M/C19M	16	22433445	22438802	duplication_inverted	5357/5357/5357	*RRN3P3; SMG1P1*	3
C01_1F/C01_2M/C11F	21	35495648	35528547	insertion	6390/6373/6289	*LOC100506403*	3
C09M/C18M/C21M	12	40785645	40796735	deletion	607/510/596	*CNTN1*	3
C05M/C06M/C09M	10	46334173	46839937	duplication_inverted	505764/505764/505764	*AGAP14P; FAM25BP; ANXA8L1; LINC00842; HNRNPA1P33; NPY4R; GPRIN2; SYT15; SYT15-AS1; SHLD2P1; GLUD1P2; BMS1P1; AGAP13P*	3
C14_1M/C17M/C20M	7	67012574	67024171	deletion	2335/2245/2235	*TYW1*	3
C11F/C12_2M/C15M	11	67704392	67980926	deletion	251621/251547/251658	*FAM86C2P; LINC02754*	3
C04_1F/C04_2F/C20M	16	73383124	73391548	deletion	2153/2199/2161	*ZFHX3; LINC01568*	3
C01_1F/C02M/C10F	10	77913100	77921771	insertion	829/975/978	*DLG5*	3
C11F/C14_1M/C14_2M	X	101011018	101026650	insertion	1099/1148/1187	*TRMT2B*	3
C04_1F/C04_2F/C17M	7	107345835	107362573	deletion	5599/5575/5587	*COG5*	3
C05M/C12_1M/C12_2M	2	112065427	112069329	insertion	532/579/531	*TMEM87B*	3
C02M/C20M/C22M	6	124525899	124536552	deletion	530/679/546	*NKAIN2*	3
C08M/C11F	12	1704751	1705901	insertion	1015/1096	*ADIPOR2*	2
C04_2F/C09M	3	5167874	5171913	insertion	2798/2839	*ARL8B*	2
C01_2M/C06M	12	8230346	8238848.5	duplication	8502/8502	*FAM86FP; LINC02449*	2
C10F/C12_1M	2	10629719	10639034	deletion	816/744	*NOL10*	2
C01_1F/C07M	10	18258429	18263616	insertion	1541/1533	*CACNB2*	2
C14_2M/C15M	12	21623450	21641776	inversion_paired	22556/22556	*LDHB*	2
C04_1F/C12_1M	7	21917405	21922834.8	deletion	558/615	*CDCA7L*	2
C04_2F/C07M	7	22167581	22181734	deletion	651/574	*RAPGEF5*	2
C02M/C05M	14	30704057	30714012	insertion	567/551	*SCFD1*	2
C14_1M/C18M	15	33610425	33626110	deletion	4176/4185	*RYR3*	2
C07M/C18M	9	34065848	34088597	deletion	3356/3361	*DCAF12*	2
C03M/C12_1M	X	38621628	38629703	deletion	981/956	*TSPAN7*	2
C03M/C12_1M	X	38634354.5	38640891	insertion	1060/922	*TSPAN7*	2
C02M/C05M	12	40739132	40757795	deletion	2583/2619	*CNTN1*	2
C03M/C09M	2	42284920	42293451	insertion	503/506	*EML4*	2
C02M/C10F	21	44562105	44577003	deletion	575/526	*TSPEAR; KRTAP10–4*	2
C04_1F/C19M	2	54861627	54866802	insertion	4168/4153	*EML6*	2
C04_2F/C17M	20	56395918	56405599	insertion	783/917	*CSTF1*	2
C06M/C09M	1	62108157	62139884	insertion	2541/2585	*PATJ*	2
C19M/C21M	4	70025175	70032349	insertion	2186/2196	*HTN3*	2
C02M/C18M	10	73327612	73341231	insertion	3103/3033	*CFAP70*	2
C06M/C21M	7	75468275	75521168	deletion	41348/41352	*POM121C; SPDYE5; PMS2P3*	2
C07M/C19M	16	80637333	80667034	deletion	1799/1759	*CDYL2*	2
C14_2M/C18M	5	95582130.5	95587973	insertion	805/679	*ARSK*	2
C04_2F/C10F	11	97871977	97882191	insertion	515/520	*LINC02713*	2
C03M/C13M	8	99051914	99058259	insertion	2950/2914	*VPS13B*	2
C09M/C20M	9	105720930	105736728	deletion	3350/3408	*TMEM38B*	2

(*continued on next page*)

Table 3 (*continued*)

Participants	Chr	Start	End	Type	SV size (bp)	Overlap Genes	Positive participant number
C05M/C20M	6	124225522	124244602	deletion	1124/1071	*NKAIN2*	2
C10F/C19M	7	152127710	152251657.5	duplication	123947/123947	*KMT2C*	2
C04_2F/C10F	7	156626235	156634752	insertion	2329/2339	*RNF32-DT*	2
C03M/C13M	1	174924777	174940199	deletion	3377/3418	*RABGAP1L; RABGAP1L-AS1*	2
C03M/C15M	3	185215768	185246141	deletion	832/903	*EHHADH*	2

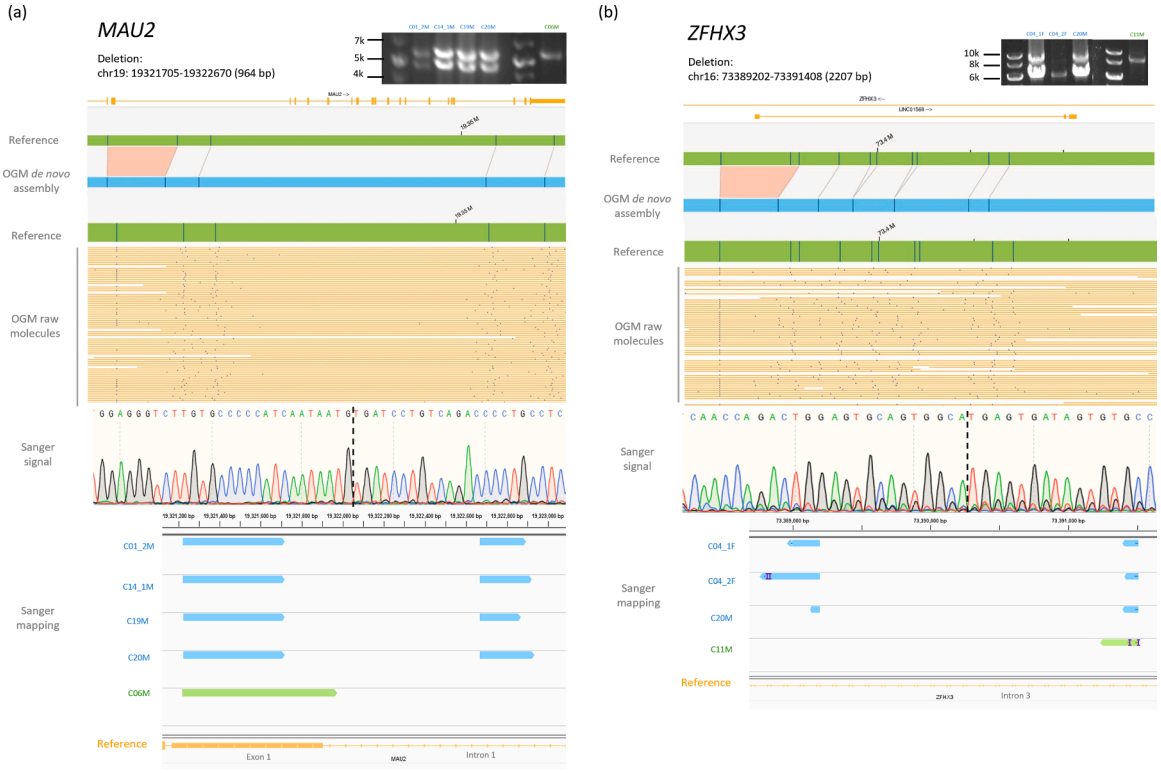

Fig. 4. Validation of SVs in *MAU2* and *ZFHX3* genes.(a) OGM detected an estimated 941-bp heterozygous deletion in *MAU2* (representative sample with deletion: C01_2M, without deletion: C06M). PCR and Sanger sequencing confirmed a 964-bp deletion. The Sanger signal channel shows the sequencing result of the deletion-containing gel band, with the breakpoint illustrated by the dotted line. (b) OGM detected an estimated 2,171-bp heterozygous deletion in *ZFHX3* (representative sample with deletion: C04_1F, without deletion: C11M). PCR and Sanger sequencing confirmed a 2,207-bp deletion. The Sanger signal channel shows the sequencing result of the deletion-containing gel band, with the breakpoint illustrated by the dotted line.

CRediT authorship contribution statement

Yunjia Zhang: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wai-Tong Chien:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Judy Yuet-Wa Chan:** Writing – original draft, Validation, Investigation, Writing – review & editing. **Tsz-Yan Cheung:** Writing – original draft, Validation, Investigation. **Zhiqian He:** Visualization, Software, Formal analysis. **Kai-Hang Yip:** Writing – original draft. **Dorothy Fung-Ying Chan:** Resources. **Josephine Shuk-Ching Chong:** Resources. **Mary Miu-Yee Waye:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Sek-Ying Chair:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Ting-Fung Chan:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Ethics declaration

This study was approved by The Joint Chinese University of Hong Kong - New Territories East Cluster Clinical Research Ethics Committee (Reference number: 2022.425) and registered at ISRCTN (ISRCTN79375633). Informed consent was obtained from the participants and their families, ensuring that all personal data would be anonymized, and the research findings will be published in international journals for research purposes.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Hong Kong Research Grants Council Area of Excellence Scheme (AoE/M-403/16), a donation from Mr. and Mrs. Sunny Yang, funding from the Alice Ho Miu Ling Nethersole Charity Foundation (Foundation), and the Innovation and Technology Commission, Hong Kong Special Administrative Region Government to the State Key Laboratory of Agrobiotechnology, CUHK. Any opinions, findings, conclusions, or recommendations expressed in this publication do not reflect the views of the Government of the Hong Kong Special

Administrative Region or the Innovation and Technology Commission. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2025.05.043.

Data availability

Data are available upon request. All variants identified in this study have been provided in the supplementary tables. All other genome sequencing data are subject to controlled access because they may compromise the privacy of research participants, especially for potentially vulnerable groups. These data may become available upon a data transfer agreement approved by the local ethics committee and can be obtained after contacting the corresponding authors (T.-F. Chan and S.-Y. Chair) upon request.

References

[1] Lord C, Brugha TS, Charman T, Cusack J, Dumas G, et al. Autism spectrum disorder. Nat Rev Dis Prim 2020;6(1):5.
[2] Gillberg C, Billstedt E. Autism and Asperger syndrome: coexistence with other clinical disorders. Acta Psychiatr Scand 2000;102(5):321–30.
[3] Autism spectrum disorder in under 19s: support and management. National Institute for Health and Care Excellence: Clinical Guidelines. Published online 2021. Accessed August 18, 2024. ⟨http://www.ncbi.nlm.nih.gov/books/NBK571989/⟩.
[4] Zeidan J, Fombonne E, Scorah J, Ibrahim A, Durkin MS, et al. Global prevalence of autism: a systematic review update. Autism Res 2022;15(5):778–90.
[5] Data and statistics on autism spectrum disorder. Centers for Disease Control and Prevention. Published online July 19, 2024. Accessed August 18, 2024. ⟨https://www.cdc.gov/autism/data-research/index.html⟩.
[6] Hirota T, King BH. Autism spectrum disorder: a review. JAMA 2023;329(2):157–68.
[7] Wong VCN, Hui SLH. Epidemiological study of autism spectrum disorder in China. J Child Neurol 2008;23(1):67–72.
[8] EDB to enhance support for students with autism spectrum disorders. The Government of the Hong Kong Special Administrative Region. Published online 2019. Accessed August 18, 2024. ⟨https://www.info.gov.hk/gia/general/201910/03/P2019100300291.htm?fontSize= 1⟩.
[9] Shen L, Liu X, Zhang H, Lin J, Feng C, et al. Biomarkers in autism spectrum disorders: current progress. Clin Chim Acta 2020;502:41–54.
[10] Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. Lancet 2018;392(10146):508–20.
[11] Medavarapu S, Marella LL, Sangem A, Kairam R. Where is the evidence? A narrative literature review of the treatment modalities for autism spectrum disorders. Cureus 2019;11(1):e3901.
[12] Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, et al. Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. Pediatrics 2011;128(3):e488–95.
[13] Tick B, Bolton P, Happé F, Rutter M, Rijsdijk F. Heritability of autism spectrum disorders: a meta-analysis of twin studies. J Child Psychol Psychiatry 2016;57(5):585–95.
[14] Geschwind DH, State MW. Gene hunting in autism spectrum disorder: on the path to precision medicine. Lancet Neurol 2015;14(11):1109–20.
[15] Qiu S, Qiu Y, Li Y, Cong X. Genetics of autism spectrum disorder: an umbrella review of systematic reviews and meta-analyses. Transl Psychiatry 2022;12(1):249.
[16] Krumm N, Turner TN, Baker C, Vives L, Mohajeri K, et al. Excess of rare, inherited truncating mutations in autism. Nat Genet 2015;47(6):582–8.
[17] Brandler WM, Antaki D, Gujral M, Noor A, Rosanio G, et al. Frequency and complexity of de novo structural mutation in autism. Am J Hum Genet 2016;98(4):667–79.
[18] 1000 Genomes Project ConsortiumAuton A, Brooks LD, Durbin RM, Garrison EP, et al. A global reference for human genetic variation. Nature 2015;526(7571):68–74.
[19] Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, et al. An integrated map of structural variation in 2,504 human genomes. Nature 2015;526(7571):75–81.
[20] Chiang C, Scott AJ, Davis JR, Tsang EK, Li X, et al. The impact of structural variation on human gene expression. Nat Genet 2017;49(5):692–9.
[21] Hurles ME, Dermitzakis ET, Tyler-Smith C. The functional impact of structural variation in humans. Trends Genet 2008;24(5):238–45.
[22] Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, et al. A structural variation reference for medical and population genetics. Nature 2020;581(7809):444–51.
[23] Alibutud R, Hansali S, Cao X, Zhou A, Mahaganapathy V, et al. Structural variations contribute to the genetic etiology of autism spectrum disorder and language impairments. Int J Mol Sci 2023;24(17):13248.
[24] Chen CH, Chen HI, Chien WH, Li LH, Wu YY, et al. High resolution analysis of rare copy number variants in patients with autism spectrum disorder from Taiwan. Sci Rep 2017;7(1):11919.
[25] Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron 2015;87(6):1215–33.
[26] Chair SY, Chow KM, Chan CWL, Chan JYW, Law BMH, et al. Structural variations identified in patients with autism spectrum disorder (ASD) in the Chinese population: a systematic review of case-control studies. Genes 2024;15(8):1082.
[27] Spielmann M, Lupiáñez DG, Mundlos S. Structural variation in the 3D genome. Nat Rev Genet 2018;19(7):453–67.
[28] D'haene E, Vergult S. Interpreting the impact of noncoding structural variation in neurodevelopmental disorders. Genet Med 2021;23(1):34–46.
[29] Brandler WM, Antaki D, Gujral M, Kleiber ML, Whitney J, et al. Paternally inherited cis-regulatory structural variants are associated with autism. Science 2018;360(6386):327–31.
[30] Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, et al. Most genetic risk for autism resides with common variation. Nat Genet 2014;46(8):881–5.
[31] Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. Nat Rev Genet 2011;12(5):363–76.
[32] Li W, Olivier M. Current analysis platforms and methods for detecting copy number variation. Physiol Genom 2013;45(1):1–16.
[33] Shaffer LG. A cytogeneticist's perspective on genomic microarrays. Hum Reprod Update 2004;10(3):221–6.
[34] Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, et al. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol 2019;20(1):117.
[35] Yang L. A practical guide for structural variation detection in the human genome. Curr Protoc Hum Genet 2020;107(1):e103.
[36] Dremsek P, Schwarz T, Weil B, Malashka A, Laccone F, et al. Optical genome mapping in routine human genetic diagnostics-its advantages and limitations. Genes 2021;12(12):1958.
[37] Mantere T, Neveling K, Pebrel-Richard C, Benoist M, van der Zande G, et al. Optical genome mapping enables constitutional chromosomal aberration detection. Am J Hum Genet 2021;108(8):1409–22.
[38] Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 2016;17(6):333–51.
[39] Yuan Y, Chung CYL, Chan TF. Advances in optical mapping for genomic research. Comput Struct Biotechnol J 2020;18:2051–62.
[40] Shelton JM, Coleman MC, Herndon N, Lu N, Lam ET, et al. Tools and pipelines for BioNano data: molecule assembly pipeline and FASTA super scaffolding tool. BMC Genom 2015;16:734.
[41] Savara J, Novosád T, Gajdoš P, Kriegová E. Comparison of structural variants detected by optical mapping with long-read next-generation sequencing. Bioinformatics 2021;37(20):3398–404.
[42] Ebert P, Audano PA, Zhu Q, Rodriguez-Martin B, Porubsky D, et al. Haplotype-resolved diverse human genomes and integrated analysis of structural variation. Science 2021;372(6537):eabf7117.
[43] Signs and symptoms of autism spectrum disorder. Centers for Disease Control and Prevention. Published online July 19, 2024. Accessed January 14, 2025. ⟨https://www.cdc.gov/autism/signs-symptoms/index.html⟩.
[44] Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34(18):3094–100.
[45] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25(16):2078–9.
[46] Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013;14(2):178–92.
[47] Parenti I, Diab F, Gil SR, Mulugeta E, Casa V, et al. MAU2 and NIPBL variants impair the heterodimerization of the cohesin loader subunits and cause Cornelia de Lange Syndrome. Cell Rep 2020;31(7):107647.
[48] Zhu Z, Wang X. Roles of cohesin in chromosome architecture and gene expression. Semin Cell Dev Biol 2019;90:187–93.
[49] Ascaso Á., Arnedo M., Puisac B., Latorre-Pellicer A., Del Rincón J., et al. Cornelia de Lange Spectrum. An Pediatr (Engl Ed). 2024;100(5):352-362.
[50] Rieger M, Moutton S, Verheyen S, Steindl K, Popp B, et al. Microdeletions at 19p13.11p12 in five individuals with neurodevelopmental delay. Eur J Med Genet 2023;66(1):104669.
[51] Pérez Baca MDR, Jacobs EZ, Vantomme L, Leblanc P, Bogaert E, et al. Haploinsufficiency of ZFHX3, encoding a key player in neuronal development, causes syndromic intellectual disability. Am J Hum Genet 2024;111(3):509–28.

[52] Al-Naama N, Mackeh R, Kino T. C2H2-type zinc finger proteins in brain development, neurodevelopmental, and other neuropsychiatric disorders: systematic literature-based analysis. Front Neurol 2020;11:32.
[53] He MF, Liu LH, Luo S, Wang J, Guo JJ, et al. ZFHX3 variants cause childhood partial epilepsy and infantile spasms with favourable outcomes. J Med Genet 2024; 61(7):652–60.
[54] Hashimoto R, Nakazawa T, Tsurusaki Y, Yasuda Y, Nagayasu K, et al. Whole-exome sequencing and neurite outgrowth analysis in autism spectrum disorder. J Hum Genet 2016;61(3):199–206.
[55] Peter B, Dinu V, Liu L, Huentelman M, Naymik M, et al. Exome sequencing of two siblings with sporadic autism spectrum disorder and severe speech sound disorder suggests pleiotropic and complex effects. Behav Genet 2019;49(4):399–414.
[56] Lotufo Denucci B, Silva de Lima L, Ferreira Lima Mota I, Rocha Madureira Azevedo J, Germino Veras L, et al. Current knowledge, challenges, new perspectives of the study, and treatments of Autism Spectrum Disorder. Reprod Toxicol 2021;106:82–93.