Molecular Cytogenetics

BRIEF REPORT Open Access

Revealing the impact of partial gene duplications in *ASH1L*: integration of optical genome mapping and RNA sequencing

Grégoire Blavier[1], François Lecoquierre[1], Anne-Marie Guerrot[1], Géraldine Joly Hélas[1], Stéphane Rondeau[2], Anne Boland[3], Jean-François Deleuze[3], Gaël Nicolas[1], Pascal Chambon[1] and Kévin Cassinari[1*]

Abstract

Introduction Partial gene duplications are structural variants that are challenging to interpret, particularly in the context of neurodevelopmental disorders. The *ASH1L* gene, associated with autism spectrum disorders and cognitive impairment, exemplifies the complexity of such variants. This study explores the integration of Optical Genome Mapping (OGM) with traditional cytogenetic techniques and RNA sequencing to enhance the characterization of *de novo* partial gene duplications.

Methods Initial detection of the duplication was performed using array comparative genomic hybridization (CGH) and exome sequencing, which were insufficient to resolve the detailed structure or predict functional impacts. OGM was employed to clarify the structural arrangement, while RNA sequencing assessed the expression profile of the *ASH1L* gene.

Results OGM identified a tandem arrangement of two duplications at 1q22. One duplication resulted in a 3-exon intragenic duplication with a predicted frameshift effect, which conventional methods had misinterpreted as a single event. RNA sequencing revealed no reduction in *ASH1L* mRNA levels despite the frameshift, suggesting the non-activation of the nonsense-mediated decay (NMD) system.

Discussion These findings challenge conventional views on the functional consequences of structural variants. The study demonstrates the capability of OGM to uncover complex genomic rearrangements that evade detection by traditional methods. Integrating advanced genomic tools enhances diagnostic precision and broadens our understanding of the pathogenicity of structural variants in developmental disorders.

Keywords Partial duplication, Optical mapping, CGH array, Exome, Autism, ASH1L, RNAseq

*Correspondence:
Kévin Cassinari
kevin.cassinari@chu-rouen.fr
[1]Department of Genetics and reference Center for Developmental Disorders, Univ Rouen Normandie, Normandie Univ, Inserm U1245 and CHU Rouen, Rouen 76000, France
[2]Department of Early Medico-Social Action (CAMSP), CHU de Rouen, Rouen, France
[3]Université Paris-Saclay, CEA, Centre National de Recherche en Génomique Humaine (CNRGH), Évry-Courcouronnes 91057, France

Introduction

Gains and losses of DNA segments, known as copy number variants (CNV) are a common cause of genetic diseases. Among them, duplications of a full gene (designed here as full duplications) or a portion of a given gene (designed here as partial duplications). Duplications that encompass the region of an entire gene do not disrupt the coding sequence of the transcript, but may result in increased transcription and protein expression. While they may increase gene expression, leading to some monogenic disorders [1], they do not inherently alter the resulting protein structure. Duplications may be in direct or inverted orientation relative to the original copy and the insertion of the duplicated sequence may be either tandem (directly beside the original sequence) or interspersed (at another location in the genome). Dependent on the size, orientation and context of the duplication there can be various consequences including gain or loss of expression due to disruption of the gene's coding sequence [2]. Among partial gene duplications, a distinction can be made between partial overlapping gene duplications including an intragenic breakpoint and the second one out of the gene sequence and partial intragenic gene duplications containing both breakpoints within the same gene. Also, duplications can adopt different arrangements, such as tandem duplications, which are defined by duplicated segments juxtaposition, or alternative configurations.

Specifically, a partial tandem intragenic duplication overlapping exons can lead to a duplication of coding DNA sequences.

Such intragenic tandem duplications can significantly influence the encoded protein, potentially resulting in a protein extension when the duplicated segment is in-frame, or leading to haploinsufficiency through a frameshift and a premature stop codon. Alternatively, interspersed duplications, where the duplicated segment is inserted elsewhere in the genome, may disrupt other genes or regulatory elements, resulting in additional pathogenic mechanisms. It is now recognized that intragenic duplications of this type are nearly as prevalent in the human genome as deletions of coding sequences or protein-truncating nucleotide variants [3]. Nevertheless, due to their variable effects, caution is advised in interpreting such variations according to international guidelines and complementary analytical approaches to properly assess their impact on gene expression [4]. Even when the complete structural configuration is resolved, tertiary regulatory mechanisms, such as splicing, methylation, or disruption of promoter and enhancer elements, may further influence the phenotypic outcome.

Partial tandem overlapping gene duplications represent a complex category of structural variation. Although tandem duplications generally have a lower likelihood of markedly altering the expression of the donor gene, their effects remain context-dependent.

However, traditional cytogenetic techniques, comparative genomic hybridization (arrayCGH) or exome sequencing often fail to fully elucidate the comprehensive impact of partial duplications on gene structure and expression.

Recent advances in long-fragment DNA analysis technologies have profoundly enhanced our comprehension of structurally complex genomic regions containing segmental duplications or repetitive elements, which complicates accurate interpretation of their functional impact. Techniques such as long-read sequencing and optical genome mapping (OGM) enable the characterization of complex structural rearrangements, particularly within highly repetitive or segmentally duplicated regions, providing substantial advantages over short-read approaches. This study highlights the use of OGM, based on ultra-high molecular weight DNA molecules, for the detailed interpretation of gene duplications.

OGM, developed by Bionano Genomics through the Saphyr system, employs high-resolution genomic analysis capable of detecting and characterizing variants ranging from simple to highly complex in terms of number and structure. This method allows the analysis of fluorescently-labeled, high molecular weight DNA, with sizes ranging 150 kb up to a few megabases. The specific CTTAAG sequence recognized for DNA labelling occurs approximately every 6 kb in the human genome with a specific pattern for each chromosomal region. After preparation, DNA samples undergo nanopore chip-based linearization for image capture using a laser-equipped scanner. Algorithms then *de novo* assemble these images to generate the unique labeled genomic pattern of a given sample. Structural variations (including CNV) detection relies on a comparison to the reference genome and a quantification of the mapped molecules.

In this study, we illustrate the effectiveness of OGM in characterizing partial duplications, such as this involving the *ASH1L* gene. This approach enabled us to resolve the structure of a complex duplication variant that could not be clarified using array-CGH or exome sequencing, thereby providing a more accurate interpretation of its potential pathogenicity. ultra-high molecular weight DNA techniques like OGM prove valuable in accurately characterizing and reclassifying partial gene duplications, particularly those involving genes with clinical significance like *ASH1L*.

Methods

Array-CGH

DNA samples from all individuals were isolated from whole blood using standard procedures. Array-CGH analysis was conducted using the Agilent SurePrint

4 × 180 K microarray kit following standard protocols (Agilent Technologies, Santa Clara, USA).

Exome sequencing

Exomes were captured using Agilent SureSelect Human All Exon kits V6 UTR (Agilent Technologies, Santa Clara, CA, USA). Final libraries were sequenced on an Illumina NovaSeq 6000 with paired ends of 150 bp reads (Illumina, San Diego, CA, USA) at the French National Center for Human Genomics Research (CNRGH, Evry, France). Reads were mapped to the GRCh37 build using BWA 0.7.5a (Li and Durbin, 2009). Duplicate reads were flagged using Picard Tools 1.101 (http://broadinstitute.github.io/picard/), and GATK was applied for short ins ertion and deletion (indel) realignment and base quality score recalibration. Copy number detection was performed using a CANOES-based workflow, as previously described [5].

Optical genome mapping

OGM, developed by Bionano Genomics through the Saphyr system, employs high-resolution genomic analysis capable of detecting and characterizing variants ranging from simple to highly complex in terms of number and structure. This method allows the analysis of fluorescently-labeled, high molecular weight DNA, with sizes ranging 150 kb up to a few megabases. The specific CTTAAG sequence recognized for DNA labelling occurs approximately every 6 kb in the human genome with a specific pattern for each chromosomal region. After preparation, DNA samples undergo nanopore chip-based linearization for image capture using a laser-equipped scanner. Algorithms then *de novo* assemble these images to generate the unique labeled genomic pattern of a given sample. Structural variations (including CNV) detection relies on a comparison to the reference genome and a quantification of the mapped molecules.

Ultra-high-molecular-weight (UHMW) DNA was isolated from whole peripheral blood volume based on leukocyte count (EDTA) using the Bionano Prep SP DNA Isolation Kit following the manufacturer's instructions (Bionano Genomics, San Diego, CA, USA). Briefly, cells were lysed using lysis-and-binding buffer (LBB), and gDNA was bound to a nanobind disk, washed, and eluted in the provided elution buffer. UHMW gDNA molecules were labeled using the Direct Label and Stain (DLS) DNA Labeling Kit (Bionano Genomics, San Diego, CA, USA) with Direct Label Enzyme (DLE-1) and DL-green fluorophores. After washing out excess DL-green fluorophores, the DNA backbone was counterstained overnight before quantification and visualization on a Saphyr instrument. Labeled UHMW gDNA was loaded onto a Saphyr chip for linearization and imaging. Structural variants were called using the Bionano Access 1.7 software with the *de novo* assembly pipeline.

RNA sequencing

RNA was extracted from Paxgene blood samples following the manufacturer's protocol. Reverse transcription was performed using random primers from the Agilent SureSelect cDNA Module (Agilent Technologies). Coding exons were captured using Agilent Magnis with SureSelect all exon V7 probes. Sequencing was performed on an Illumina NextSeq 500 at the Rouen sequencing facility, where eight samples were pooled and sequenced using High Output Kit v2.5 (150 Cycles, 400 M reads). Data were processed using the nf-core rnaseq pipeline, including read alignment with STAR on GRCh38, quantification with Salmon, and aberrant junction analysis with RSeQC. Aberrant splicing and mono-allelic expression at specific loci of interest were analyzed using IGV. For quantitative analysis, adjusted gene expression was compared using transcripts per million (TPM) metrics among five probands and three additional samples from the same sequencing batch, all of whom had developmental disorders and variants of unknown significance in unrelated genes.

Patients

The parents of the minor patient provided informed written consent for genetic analyses in a medical setting and for the participation of the REDIA study, focused on autism spectrum disorder genetics, following the approval by the French Ouest-I ethics committee.

Patient and results

Clinical description

A 15-year-old male presented to the genetic consultation for the etiological assessment of autism spectrum disorder. He was born eutrophic (weight of 3920 g (77th percentile), length of 51.5 cm (59th percentile), and head circumference of 36.5 cm (81st percentile)) at 38 weeks of gestation. The prenatal period was characterized by maternal hypertension, and abnormal fetal cardiac rate was detected during delivery. Neonatal examination was unremarkable, except for plagyocephaly and clavicle fracture.

During his early developmental stages, the patient achieved several milestones within the normal range: he attained a sitting posture at 8–9 months, began walking at 11 months, and first words between 12 and 16 months. However, a complex oral and written language disorder was later diagnosed and required speech therapy intervention. He also experienced sleep disorder since birth.

He attended a standard school, with support from a school life assistant and accommodations to address his educational needs. Behavioral assessments revealed

interaction difficulties, anxiety, attention deficit hyperactivity disorder (ADHD), after which he developed hyperphagia, motor and oral tics and other behavioral peculiarities, culminating in a formal diagnosis of autism spectrum disorder (ASD) after thorough evaluation Autism Diagnostic Observation Schedule (ADOS-2). These findings were consistent with the phenotype associated with ASH1L haploinsufficiency.

Upon physical examination, no distinctive facial dysmorphism was noted. Ophthalmological assessments indicated myopia and astigmatism. Cognitive evaluation using the *Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V)* showed subscores ranging from 71 to 83. Electroencephalography revealed polyspikes and waves but without clinical seizures.

He had a maternel half-sister, asymptomatic and an elder brother who was also diagnosed with mild autism spectrum disorder, in addition to intellectual disability (QIT = 48). He had also myopia and astigmatism and discreet morphological features: anteverted nostrils, attached lobes, triangular fingers. The neurodevelopemental phenotype appears different from the proband. Their mother had been diagnosed with ankylosing spondylitis, testing positive for HLA-B27. She had three miscarriages. Aside from the conditions mentioned, the parents were asymptomatic.

Genomics explorations

Array-CGH (Agilent 180 K) analysis detected a 199-kb interstitial heterozygous duplication on long arm of chromosome 1: arr[GRCh38] 1q22(155,253,465_155,452,974) x3. This duplication was confirmed by custom universal digital PCR [6], and parental analysis showed that this duplication occurred *de novo*, after confirmation of parenthood by microsatellites markers. The duplication was not detected in the brother. The duplication encompassed the entiretyof eight non-disease-relevant genes and partially overlapped the *ASH1L* gene (from exon 4 to the 3' extremity of the gene). Initially classified as a variant of uncertain significance (VUS), this patient was further assessed by exome sequencing to investigate an alternative cause. Exome sequencing did not detect any sequence variants that were classified as pathogenic or likely pathogenic in genes that would explain the clinical presentation; however, a single duplication at 1q22 seq[GRCh38] 1q22(155234977_155459950) was reported, consistent with the duplicated region reported by array-CGH (Fig. 1).

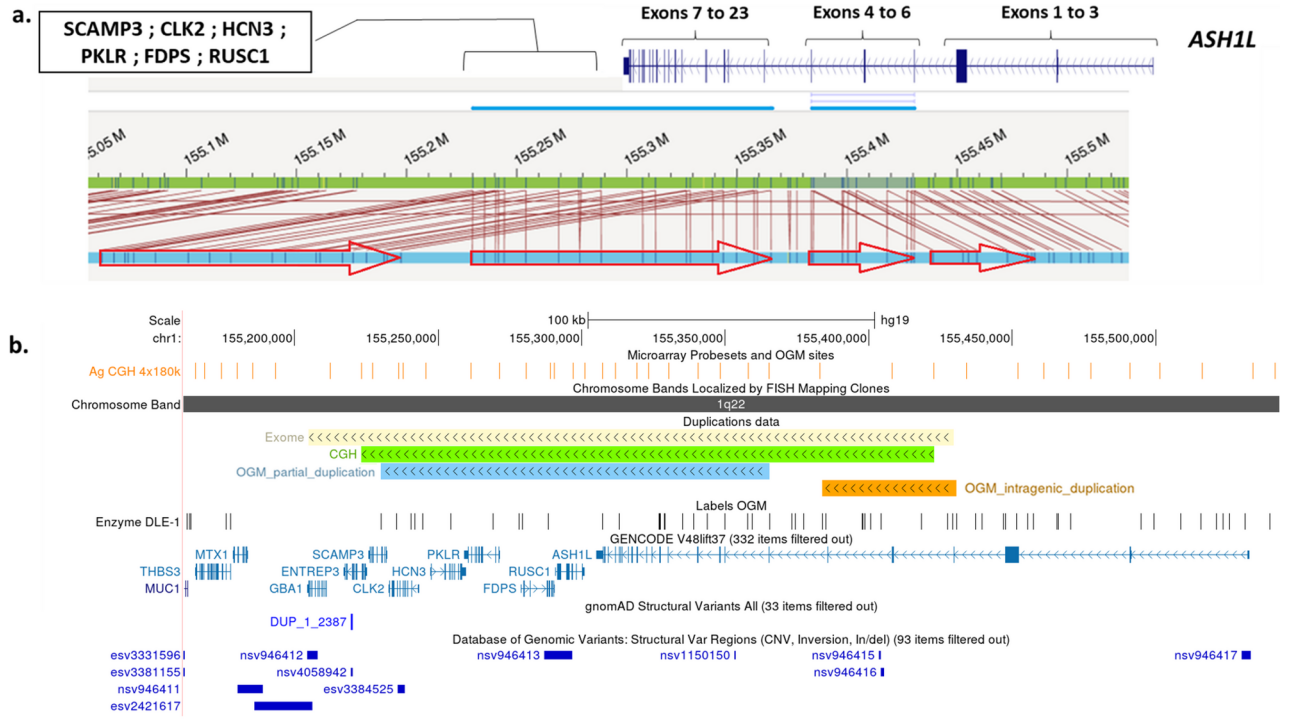

Fig. 1 Representation of the detected complex rearrangement **(a)** Optical Genome Mapping (OGM) view, focused on duplicated region. Green band represents reference genome's map, blue band patient's map. Blue lines from Bionano SV feature shows both adjacent duplications. Red arrows overlap duplicated labels on patient's map. **(b)** UCSC view of duplicated region. Panels represent, from top to bottom: Ag CGH 4 × 180k track shows CGH array probes, Duplications data track shows Exome, CGH and OGM detected events positions, Enzyme DLE-1 track shows OGM labels, UCSC Genes track shows genes name and positions including exons and introns position on each strand (blue arrows), GnomAD Structural Variants track (filtered on Duplications) and Database of Genomic Variants (DGV) Structural Variants (filtered on Duplications and Gain)

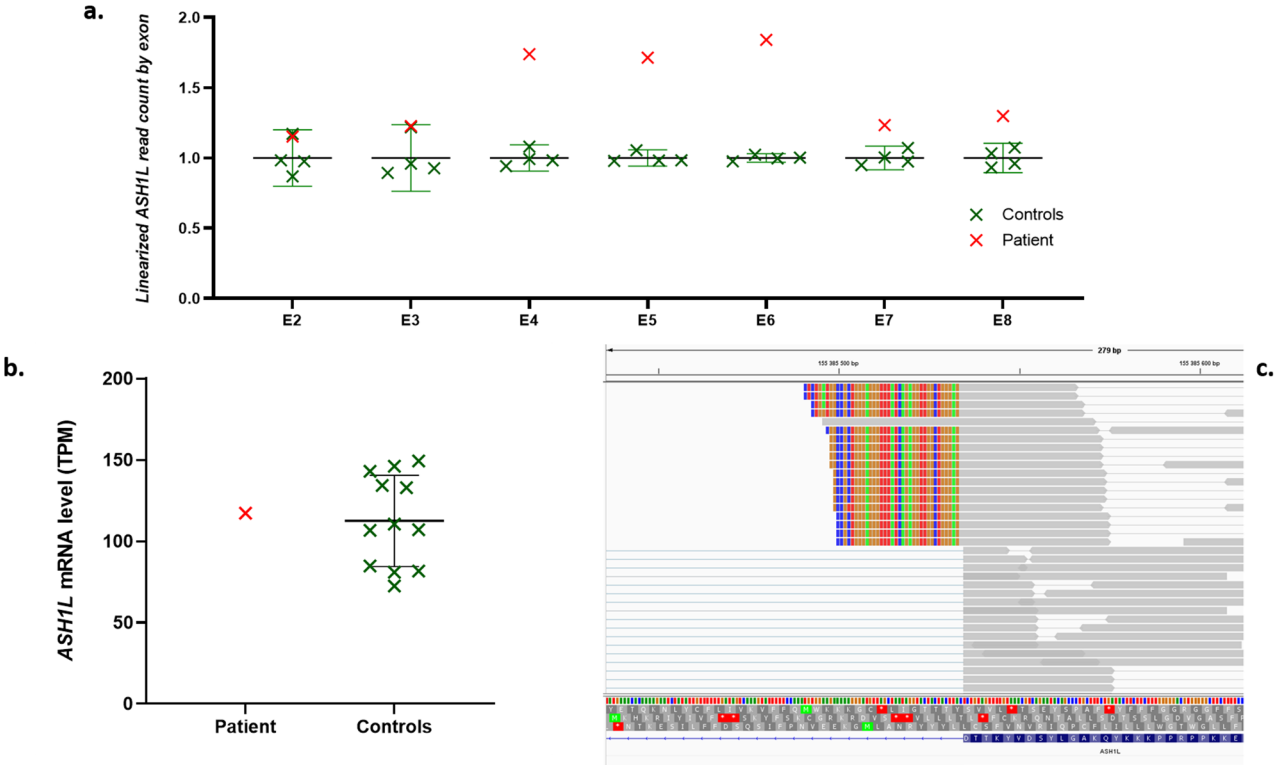

Fig. 2 RNAseq exploration **(a)** Linearized *ASH1L* mRNA read count, by exon, focused on exon 2 to 8. Red symbols represent patient's mRNA levels, green symbols for controls. **(b)** *ASH1L* mRNA level, expressed in transcript per million (TPM), and compared to 12 control sample, with an RNAseq performed in the same analytic conditions. **(c)** RNAseq reads view from Integrative Genomics Viewer (IGV), showing normal exon 6 reads and reads with aberrant exon 6 – exon 4 junctions (about 50% of each)

To further characterize this *de novo* CNV, optical genome mapping (OGM) was conducted and showed that this 199-kb *de novo* duplication was more complex than a simple duplication. OGM confirmed the presence of a large 135-kb tandem duplication on chr1q22, overlapping *ASH1L* exons 7 to 23 of as well as 6 genes mapping downstream ogm[GRCh38] 1q22(155,260,269_ 155,395,699)x3. If that duplication were the only event, it would result in the following sequence: 5'–3' full sequence of *ASH1L*, followed by 6 other genes, then a duplication with a breakpoint between exon 6 and exon 7 of *ASH1L* resulting in the sequence of exons 7 to 23 of *ASH1L* followed by these 6 genes. However, OGM also revealed a second duplication, with a size of 47-kb and adjacent to the first one. This second duplication was characterized as an intragenic tandem duplication overlapping exons 4, 5, and 6 of *ASH1L* ogm[GRCh38] 1q22(155,413,922_ 155,460,768)x3 (Fig. 1a). Thus, the 5'–3' full sequence of *ASH1L* is interrupted by an intragenic duplication of exons 4 to 6. The duplication results in an extra copy of exons 4 to 6 of ASH1L, causing a shift in the reading frame and leading to a premature stop codon. (Figure S1,Figure S2) and thus likely to a loss of function (Fig. 1b). With this refined understanding of the duplication structure, we classified this variation as likely pathogenic, due to *ASH1L* loss of function.

Transcriptomic assessment

To validate the loss-of-function hypothesis, transcriptomic RNAseq analysis was conducted on a blood sample. First of all, analysis of the BAM files highlighted reads displaying an exon 6 - exon 4 junction (see Fig. 2.C), confirming the expression of the aberrant transcript of the *ASH1L* gene. Furthermore, a 2 fold increase in mRNA levels from exons 4 to 6 was detected, demonstrating the expression of this intragenic duplication (see Fig. 2.A). Unexpectedly, a normal level of *ASH1L* mRNA global expression was observed, suggesting that the NMD (nonsense-mediated decay) system was not activated. However, limitation arises because all reads mapping to *ASH1L* are counted towards TPM, including duplicated exons' TPM, which means that gene quantity might appear normal, while some exons are expressed twice. If these data are not adjusted for, the expression of the other exons could be underestimated. This observation was confirmed by quantifying *ASH1L* gene transcripts per million (TPM), which fell within the same range as the RNAseq data from 12 control samples sequenced under similar conditions (see Fig. 2.B).

Discussion

In this observation, we illustrate how precise genomic characterization using optical genome mapping (OGM) and transcriptomic analysis with RNAseq led to the reclassification of a *de novo* partial duplication involving the *ASH1L* gene, known for its association with neurodevelopmental disorders such as autism spectrum disorders and sleep disorders [7] (OMIM#617796), which aligns with the patient's phenotype. The inheritance pattern is autosomal dominant, with a subsequent mechanism of a loss of function [8, 9]. In our case, an intragenic tandem duplication resulted in a frameshift within *ASH1L* by duplicating exons 4, 5, and 6, strongly suggesting *ASH1L* loss of function.

Clinical features observed in our patient are consistent with those described for *ASH1L* loss of function variant carriers [10]. The use of Bionano OGM technology provided critical insights into the duplication structure initially identified by aCGH and exome sequencing. Initially thought to involve a single duplication with one breakpoint affecting only the gene's terminal region, OGM analysis revealed that there were actually two duplications, one with two breakpoints. This technology enabled a more precise characterization of this rearrangement due to a better coverage of this region, highlighting the frameshift nature of the small intragenic duplication where duplicated exons are out of phase.

Subsequently, RNAseq analysis substantiated the pathogenicity of the variant, indicating a likely truncated protein due to the frameshift, despite the absence of NMD activation. RNAseq data indeed indicated absence of decrease of *ASH1L* mRNA levels and, conversely, an increase in mRNA levels spanning exons 4 to 6, along with aberrant reads displaying an exon 4 – exon 6 junction. These findings suggest the expression of the aberrant transcript and that the nonsense-mediated decay (NMD) system was not activated, potentially due to tissue-specific differences in NMD expression. Alternative hypotheses for NMD escape include alternative splicing events that exclude exons susceptible to NMD or mechanisms that prevent Exon Junction Complex (EJC) assembly [11, 12].

This study underscores the importance of detailed characterization of *de novo* structural rearrangements, especially when their functional impact is not fully understood. OGM represents a valuable and mature approach for the structural characterization of complex genomic rearrangements, providing high-resolution insights that complement other genomic technologies. The use of Bionano OGM technology provided precise insights into the structure of the duplication detected by array-CGH and exome sequencing. The duplication was initially identified as a single event involving the terminal region of the gene and a single breakpoint, and was therefore classified as a VUS due to the absence of a clearly defined pathogenic mechanism. OGM analysis later revealed that there were actually two adjacent duplications, one of which had two breakpoints. It is likely that short read genome sequencing, and, more likely, long read genome sequencing, would also have identified these breakpoints.

Similar recent reports have shown that complex structural variants, including duplications, can also be resolved using OGM in combination with complementary approaches, such as RNAseq for transcript level interpretation [13], or clinical NGS and aCGH for copy number detection [14]. Such integrative strategies improve the clinical interpretation of duplications that would otherwise remain ambigous.

Despite these advantages, OGM, like all molecular techniques that infer genome structure from DNA molecules, faces limitations in regions with very high repeat content, such as centromeres or loci enriched in segmental duplications. These genomic regions remain challenging for most molecular cytogenetic and sequencing approaches, including array-CGH, SNP arrays, and both short- and long-read sequencing [15, 16].

The observation reported here underscores the importance of finely characterizing *de novo* structural rearrangements whose gene impact is not yet established. OGM emerges as a valuable tool given its maturity and accessibility, with interpretation tools currently matured, serving as a good interim solution until the full deployment of long-read sequencing technologies.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13039-025-00740-5.

Supplementary Material 1: Figure S1. Reverse UCSC genome browser view of exons 4 and 6 of ASH1L. Figure S2. Simulation using Alamut software by creating a variant on the first codon of exon 7, c.6009C>G, to identify the first changed amino acid

Acknowledgements

The REDIA study was supported by a grant from the French Ministry of Health (PHRC-I 18-38). This work did benefit from support of the France Génomique National infrastructure, funded as part of the «Investissements d'Avenir» program managed by the Agence Nationale pour la Recherche (contract ANR-10-INBS-09). Collaboration CEA-DRF-Jacob-CNRGH—CHU de Rouen.

Author contributions

Grégoire Blavier <gregoire.blavier@chu-rouen.fr>Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – origina François Lecoquierre <francois.lecoquierre@chu-rouen.fr>Supervision, Validation, draft Writing – review & editing Anne-Marie Guerrot <anne-marie.guerrot@chu-rouen.fr>Investigation Géraldine Joly Hélas <geraldine.joly-helas@chu-rouen.fr>Investigation Stéphane Rondeau <stephane.rondeau@chu-rouen.fr>Investigation Anne Boland <boland@cng.fr>Resources Jean-François Deleuze <deleuze@cng.fr>Resources Gaël Nicolas <gael.nicolas@chu-rouen.fr>Supervision, Validation, draft Writing – review & editing Pascal Chambon <pascal.chambon@chu-rouen.fr>Supervision, Validation, draft Writing – review & editing Kevin Cassinari : Conceptualization, Methodology,

Validation, Formal analysis, Investigation, Writing – original, draft Writing – review & editing, Supervision.

Funding
The REDIA study was supported by a grant from the French Ministry of Health (PHRC-I 18–38).

Data availability
Data available upon request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 27 June 2025 / Accepted: 13 November 2025
Published online: 02 December 2025

References

1. Collins RL, et al. A cross-disorder dosage sensitivity map of the human genome. Cell. 2022;185:3041–e305525.
2. Newman S, Hermetz KE, Weckselblatt B, Rudd MK. Next-generation sequencing of duplication CNVs reveals that most are tandem and some create fusion genes at breakpoints. Am J Hum Genet. 2015;96:208–20.
3. Collins RL, et al. A structural variation reference for medical and population genetics. Nature. 2020;581:444–51.
4. Brandt T, et al. Adapting ACMG/AMP sequence variant classification guidelines for single-gene copy number variants. Genet Med. 2020;22:336–44.
5. Quenez O, et al. Detection of copy-number variations from NGS data using read depth information: a diagnostic performance evaluation. Eur J Hum Genet. 2021;29:99–109.
6. Cassinari K, et al. A simple, universal, and cost-efficient digital PCR method for the targeted analysis of copy number variations. Clin Chem. 2019;65:1153–60.
7. Shen W, Krautscheid P, Rutz AM, Bayrak-Toydemir P, Dugan SL. *De Novo* loss-of-function variants of *ASH1L* are associated with an emergent neurodevelopmental disorder. Eur J Med Genet. 2019;62:55–60.
8. Krumm N, O'Roak BJ, Shendure J, Eichler EE. A de novo convergence of autism genetics and molecular neuroscience. Trends Neurosci. 2014;37:95–105.
9. Stessman HAF, et al. Disruption of POGZ is associated with intellectual disability and autism spectrum disorders. Am J Hum Genet. 2016;98:541–52.
10. Cordova I, et al. Expansion of the genotypic and phenotypic spectrum of ASH1L-related syndromic neurodevelopmental disorder. Genes. 2024;15:423.
11. Nguyen LS, Wilkinson MF, Gecz J. Nonsense-mediated mRNA decay: inter-individual variability and human disease. Neurosci Biobehav Rev. 2014;46:175–86.
12. Dyle MC, Kolakada D, Cortazar MA, Jagannathan S. How to get away with nonsense: mechanisms and consequences of escape from nonsense-mediated RNA decay. Wiley Interdiscip Rev RNA. 2020;11:e1560.
13. Xiao B, et al. Combining optical genome mapping and RNA-seq for structural variants detection and interpretation in unsolved neurodevelopmental disorders. Genome Med. 2024;16:113.
14. Dharmadhikari AV, et al. Optical genome mapping improves clinical interpretation of constitutional copy-number gains and reduces their VUS burden. Genet Med. 2025;27:101452.
15. Xu Y, et al. Optical genome mapping for chromosomal aberrations detection-false-negative results and contributing factors. Diagnostics. 2024;14:165.
16. Mostovoy Y, et al. Resolution of ring chromosomes, Robertsonian translocations, and complex structural variants from long-read sequencing and telomere-to-telomere assembly. Am J Hum Genet. 2024;111:2693–706.

Publisher's note