A multi-omic investigation of gene dysregulation in a complex chromothripsis-like translocation event using Oxford Nanopore sequencing

Sissel Juul[1], Carly Tyer[1], Rebecca Stubbs[1], Emma Mizrahi-Powell[2], Priyesh Rughani[1], Bryan Leland[1], Sean McKenzie[1], Scott Hickey[1], Gilad Evrony[2]

[1]Oxford Nanopore Technologies Inc, New York City, NY

[2]Center for Human Genetics and Genomics, New York University Grossman School of Medicine, New York City, NY

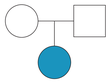
Oxford Nanopore Technologies Applications

Abstract

Rare genetic diseases caused by complex structural variation can be challenging to diagnose, and for some, it is not possible to resolve their structures by standard reference-based variant calling. Further, the downstream effects of these complex variants on genes may benefit from comprehensive multi-omic data. Here, we present a case from New York University's Undiagnosed Diseases Program (UDP) exemplifying this challenge. In a child with a severe developmental disability, we investigated a complex translocation whose structure had eluded all prior investigations. Using telomere-to-telomere (T2T) assembly with ultra-long reads, ligation-based sequencing, and Pore-C—all from Oxford Nanopore Technologies—alongside ONT-based cDNA sequencing and direct detection of 5mC/5hmC modifications, we identified and fully resolved a highly complex chromothripsis-like translocation between chromosomes 2 and 16. The event involved multiple breakpoints spanning several megabases. With this haplotype-resolved assembly, we then identified genes affected by the rearrangements. Additional transcriptomic and epigenomic datasets uncovered associated changes in gene expression and methylation. Overall, this study highlights the effectiveness of comprehensive multi-omic analyses in resolving challenging undiagnosed rare disease cases involving complex structural variation.

Background

A) Case Summary



5-year-old female with significant global developmental delays, hypotonia, and dysmorphic features

B) Previous tests and sequencing

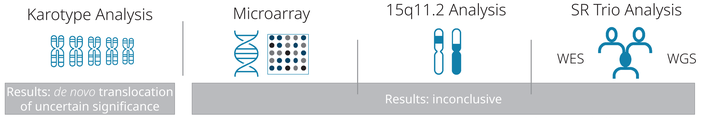


Fig. 1. Case summary and prior testing. A) Case summary including pedigree and key phenotypic features of affected individual. B) Prior genetic investigations were unable to fully resolve complex structural variation. The karyotype analysis recovered a *de novo* translocation of uncertain significance: 46,XX,t(2;16)(p21;q22). Re-analysis of the short read (SR) trio whole genome sequencing (WGS) and whole exome sequencing (WES) data at UDP found complex structural variant involving chromosomes 2 and 16, but the structure could not be resolved.

Methods

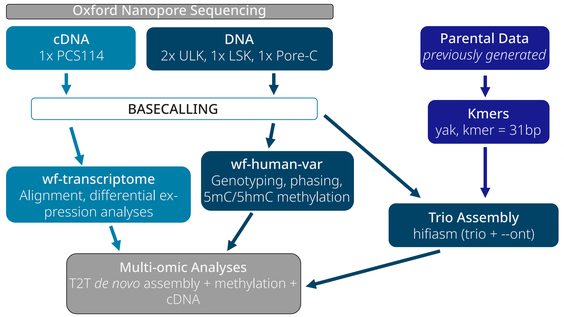


Fig. 2. Workflow for haplotype-resolved genome assembly and multi-omic analyses. Sequencing types: Ultra Long Kit (ULK), Ligation Sequencing Kit (LSK); Pore-C (chromatin conformation capture), cDNA-PCR Sequencing Kit (PCS114). The Oxford Nanopore workflows are available in EPI2ME (see QR code). All sequencing performed on the PromethION 24.

Nanopore-only T2T assembly of a complex translocation event reveals genic breakpoints

The resulting *de novo* telomere-to-telomere assembly provided 33/46 complete chromosomal contigs, including two contigs which confirmed the balanced t(2;16)(p;q) translocation event seen on karyotype and fully resolved 34 inter- and intra-chromosomal translocation breakpoints (Fig. 3; Fig. 4). This complex structural variant occurred on the paternally inherited haplotype and was likely caused by a germline chromothripsis-like event (Fig. 4B,C). Our analyses identified 17 candidate genes directly affected by structural breakpoints. One disrupted gene, *BCL11A*, has been identified as a lead candidate for further investigation based on the patient's phenotype (Fig. 1; Fig. 3C).

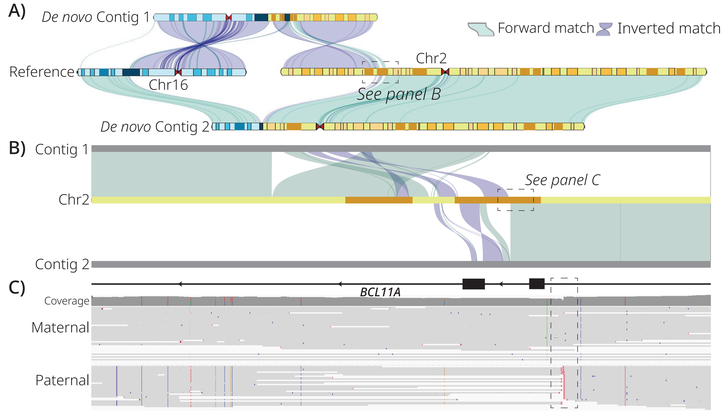


Fig. 3. *De novo* haplotype-resolved genome assembly reveals widespread inter- and intra-chromosomal rearrangements. A) and B) Synteny plot shows a large-scale translocation between chromosome arms of chromosome 2 (yellows) and chromosome 16 (blues), indicating a major structural variant. Additional smaller-scale rearrangements are visible throughout the genome, reflecting complex structural variation and a chromothripsis-like event. B) Zoom-in of chromosome 2 where *BCL11A* occurs. C) Read pileup at translocation breakpoint in *BCL11A*.

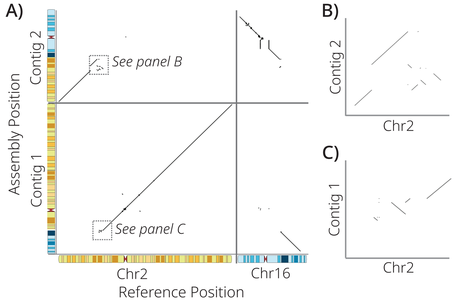


Fig. 4. Alignment of assembled chromosomes to the reference genome confirms multiple inter- and intra- chromosomal translocations and inversions. A) Translocation between chromosomes 2 and 16. The translocated segments appear as off-diagonal alignments. B) and C) Numerous additional smaller inversions and translocations are also observed, consistent with a chromothripsis-like event.

References

[1]Bauer, D. E. & Orkin, S. H. Hemoglobin switching's surprise: the versatile transcription factor *BCL11A* is a master repressor of fetal hemoglobin. Curr. Opin. Genet. Dev. 33, 62–70 (2015).

[2]Liu, N. et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell 173, 430-442.e17 (2018).

Transcriptome analysis shows increased expression of fetal gamma-globin genes

An increase in expression of fetal gamma-globin genes (*HBG1/2*), along with decreased expression in the adult beta-hemoglobin genes (*HBB, HBA1/2*; Fig. 5A,B), was observed by differential expression analysis. This finding supports the disruption of the *BCL11A* gene, resulting in an increase in fetal hemoglobin (HbF) . During development, the major hemoglobin expressed in red blood cells changes from HbF to adult hemoglobin (HbA). The level of HbF is genetically controlled and a critical modifier of the clinical severity of some diseases[1]. *BCL11A* is a transcriptional repressor that plays a key role in the developmental silencing of HbF[2].

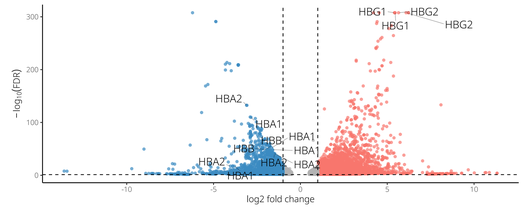


Fig. 5. cDNA sequencing reveals differential expression in *BCL11A*-associated genes, indicating a shift in hemoglobin regulation. Transcriptomic data show increased expression of fetal hemoglobin and decreased expression of adult hemoglobin genes.

Epigenetic analysis identifies differential methylation and chromatin regulatory structures near *BCL11A*

BCL11A shows hypomethylation about 70 kb from a translocation breakpoint, reflecting downstream epigenetic consequences of the translocation (Fig. 6).

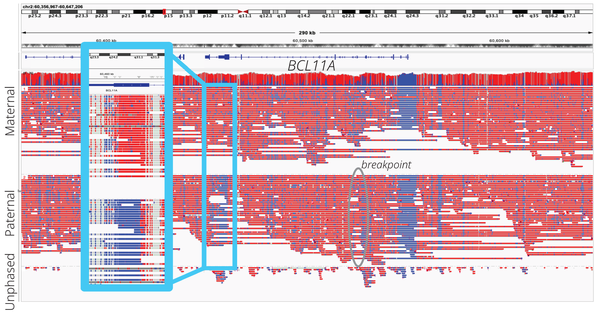


Fig. 6. Epigenetic changes near *BCL11A*. Read pileup shows differential methylation between haplotypes: paternal haplotype is hypomethylated (blue) and the maternal haplotype is hypermethylated (red) at exon 4 in *BCL11A* near the breakpoint. Inset highlights a zoomed-in view, emphasizing the allele-specific epigenetic regulation potentially influencing gene expression. This same region is hypermethylated in both haplotype in the parental genomes (not shown).

Conclusions

This study demonstrates the power of comprehensive multi-omic profiling to provide deep insights into the functional impact of genic disruptions to reveal novel rare disease mechanisms. Specifically, we identified the disruption of the *BCL11A* gene as the result of a complex, balanced chromosome arm translocation. Epigenetic and transcriptomic analyses confirmed there were increased expression and hypomethylation of *BCL11A* near the breakpoint. This study highlights the effectiveness of combining *de novo* T2T assembly, transcriptome analysis, and 5mC profiling for challenging undiagnosed rare disease cases with complex structural variations, such as the chromothripsis-like translocation studied here.