Long-read sequencing of a neurodevelopmental disorder patient reveals complex rearrangement involving the *ARID1B* gene

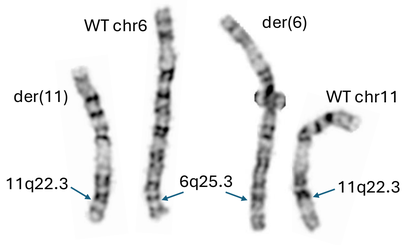
Tam P. Sneddon[1,2], Mai Xiong[1], Karen E. Weck[1,2], Bradford C. Powell[2], Kimberly Foss[2], Erin L. Heinzen[2], Ana Berglind[2], Muge Gucsavas-Calikoglu[3], Stephanie Peck[4], Yael Shiloh-Malawsky[4], Zheng (Jane) Fan[4], Senyene E. Hunter[4]

[1] Department of Pathology and Laboratory Medicine; [2] Department of Genetics; [3] Department of Pediatrics, Division of Genetics and Metabolism; [4] Department of Neurology; University of North Carolina, Chapel Hill, NC, USA

Case Presentation

The aim of the University of North Carolina's (UNC) Genetic Determinants of Neurological and Developmental Disorders (GDNDD) IRB-approved study is to identify novel genetic variants via genome sequencing. One GDNDD participant was a 5-year-old male with a history of febrile seizures, epileptic encephalopathy, and expressive speech delay. Prior clinical genetic testing in the participant included a karyotype with a t(6;11)(q25.3;q22.3) apparently balanced translocation (see Figure 1), a chromosomal microarray with a 1.48 Mb interstitial deletion within 4p13 deemed not significant, a normal fragile X assay, and no significant findings on epilepsy and neurotransmitter disorders gene panels.

Figure 1: Partial karyogram showing the t(6;11)(q25.3;q22.3) translocation



Abbreviations: chr, chromosome; der, derivative; WT, wild type

Genome Sequencing at UNC

Short-read sequence (SRS) was generated at the UNC high throughput sequencing facility (HTSF), and the data analyzed by the UNC clinical genomic analysis (GENYSIS) core facility[1] to identify single nucleotide, copy number, and structural variants (using Delly2[2] and annotated using AnnotSV[3]), but the case was determined to be negative.

Subsequently, long-read sequence (LRS) was generated at the HTSF, and structural variants were called using the Sniffles2 algorithm[4] and annotated using AnnotSV[3]. This revealed that the t(6;11) translocation disrupted the chromosome 6 *ARID1B* gene within NM_001374828.1 intron 2 of 19 and the chromosome 11 *GUCY1A2* gene within NM_000855.3 intron 4 of 7. A karyotypically cryptic concomitant 7.7 Mb inversion occurred at the chromosome 11 derivative breakpoint and disrupts the *CNTN5* gene within NM_014361.4 intron 1 of 24 (see Figures 2 and 3). Parentage was confirmed by DNA fingerprinting, and all three breakpoints were confirmed to be *de novo* by Sanger sequencing of research samples from the proband and both parents in the UNC clinical molecular genetics laboratory.

Research targeted methylation analysis for an epigenetic signature associated with causative variants in the *ARID1A, ARID1B, SMARCA2, SMARCA4,* and *SMARCB1* genes was sent to Greenwood Genetics and reported as abnormal and consistent with a BAFopathy 1 disorder[5].

ARID1B-Related Disorder

Figure 2: IGV[6] alignment of the chimeric long reads at the three breakpoints of the t(6;11) translocation and inversion. The upper alignment panel shows the chimeric reads. Three chimeric reads are highlighted. Blue read: 5' end maps to the 5' *GUCY1A2* locus while the 3' end maps to the 5' *ARID1B* locus. Red read: 5' ends maps to the 3' *CNTN5* locus while the 3' ends maps to the 3' *ARID1B* locus. Green read: 5' end maps to the 5' *CNTN5* locus while the 3' maps to the 3' *GUCY1A2* locus (see Figure 3 for interpretation and illustration of these breakpoints). The lower alignment panel shows the reads aligned to the non-rearranged loci in gray. Mismatches to the reference are highlighted by red, green, blue, and orange bars. The location of the breakpoints are indicated by the vertical red lines. Reference genome assembly: GRCh38.

The *ARID1B* gene encodes a component of the BRG1/BRM-associated factor (BAF) chromatin remodeling complex, disruption of which is linked to several neurodevelopmental syndromes, commonly referred to as BAFopathies. Loss-of-function heterozygous *de novo* variants in *ARID1B* are associated with the *ARID1B*-Related Disorder (*ARID1B*-RD) that constitutes a clinical continuum, from classic Coffin-Siris syndrome to intellectual disability with or without non-specific dysmorphic features[7-8]. Other findings seen in individuals with *ARID1B*-RD include feeding difficulties, ophthalmologic abnormalities, hearing impairment, seizures, and autistic features.

Upon review of the GDNDD participant's medical record, in addition to seizures and cognitive and speech delays, the participant had distinct facial features (relatively large head circumference, mild midface hyperplasia with upturned nose, and prominent ears) consistent with a diagnosis of *ARID1B*-RD. The *ARID1B* translocation and inversion was confirmed by Sanger sequencing of a clinical sample and reported in the participant's medical record as a likely pathogenic variant with the following ISCN[9] and HGVS[10] nomenclature:

seq[GRCh38] der(6)t(6;11)(q25.3;q22.3),der(11)inv(11)(q22.1q22.3)t(6;11)
NC_000006.12:g.156851028del;156851029_qterdelins[NC_000011.10:g.106847582_qter]
NC_000011.10:g.99115155del;99115156_106847581inv;106847582_qterdelins[NC_000006.12:g.156851029_qter]

Summary

To our knowledge this patient's *ARID1B* translocation has not been reported in the literature. However, at least four translocations involving the *ARID1B* gene, and a chromosome other than chromosome 11, have been reported in individuals with *ARID1B*-RD. All four individuals were reported to have ID, DD, absent or delayed speech, and distinctive facial features[11-14].

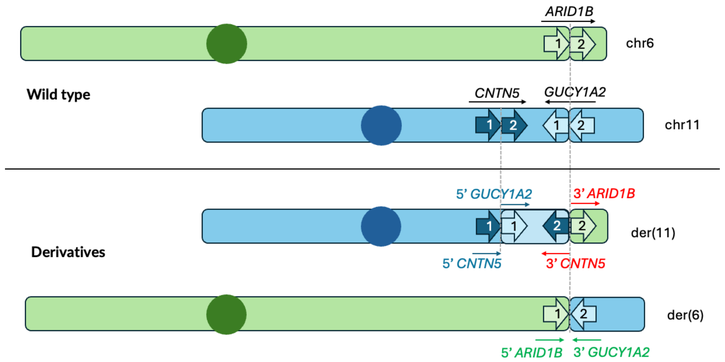


Figure 3: Illustration of the t(6;11) translocation and inversion identified by long-read sequencing. The blue, red, and green arrows correspond to the colors of the three example chimeric reads in Figure 2.

Conclusions

This case highlights the benefits of LRS to detect balanced translocations, which are often missed by SRS, and to further delineate the genomic architecture underlying such complex rearrangements that likely account for some of the missing diagnostic yield from clinical genetic testing.

References

1. UNC Clinical Genomic Analysis (GENYSIS) Core Facility: https://www.med.unc.edu/genysis/
2. Rausch T, *et al.* Bioinformatics. 2012 Sep 15;28(18):i333-i339. PMID: 22962449.
3. Geoffroy V, *et al.* Bioinformatics. 2018 Oct 15;34(20):3572-3574. PMID: 29669011.
4. Smolka M, *et al.* Nat Biotechnol. 2024 Jan 2:10.1038/s41587-023-02024-y. PMID: 38168980.
5. Greenwood Genetics EpiSign: https://ggc.org/episign
6. Integrative Genomics Viewer (IGV): https://software.broadinstitute.org/software/igv/
7. Online Mendelian Inheritance in Man (OMIM): https://www.omim.org/entry/135900
8. Vergano SA, van der Sluijs PJ, Santen G. *ARID1B*-Related Disorder. In: Adam MP, Feldman J, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2024.
9. Hastings R, *et al.* Cytogenet Genome Res. 2023;163(1-2):1-4. PMID: 37517400.
10. den Dunnen JT, *et al.* Hum Mutat. 2016 Jun;37(6):564-9. PMID: 26931183.
11. Backx L, *et al.* Cytogenet Genome Res. 2011;132(3):135-43. PMID: 21042007.
12. Halgren C, *et al.* Clin Genet. 2012 Sep;82(3):248-55. PMID: 21801163.
13. Sim JC, *et al.* Orphanet J Rare Dis. 2014 Mar 27;9:43. PMID: 24674232.
14. Malli T, *et al.* Am J Med Genet A. 2014 Dec;164A(12):3126-31. PMID: 25250687.

This study was funded by the American Epilepsy Society (AES) Sergievsky Award and subsidized by the UNC Program for Precision Medicine in Health Care (PPMH)