Genomic breakpoint analysis facilitates identification of X-chromosomal inversion among molecularly unsolved cases of Duchenne Muscular Dystrophy.

Naga Guruju, Vanessa Jump, Babi Ramesh Reddy Nallamilli, Ruby Liu, Sritha Moram, Christin Collins, Lora Bean, and Madhuri Hegde

1 Introduction

Duchenne (DMD) and Becker Muscular Dystrophy (BMD) are X-linked recessive disorders with a frequency of 1 in 5000 male infants caused by pathogenic variants in the dystrophin gene. Copy number variants (CNVs) such as intragenic deletions (65%) and duplications (5-10%) are the most frequent type of variants observed in DMD and BMD followed by approximately 30-35% single-nucleotide variants and complex rearrangements. In approximately 2-5% of patients, no causative variants in the *DMD* coding region are detected by conventional methods. Causative variants in these patients include deep intronic variants causing splicing defects or complex structural rearrangements that include balanced rearrangements such as inversions or translocations. Balanced inversions involving the *DMD* gene are rarely reported because determining the inversion breakpoints is complicated by the fact that they usually occur in the large *DMD* introns and cannot be detected due to the technical limitations of conventional molecular techniques such as MLPA and microarray. We used breakpoint junction sequence analysis to identify balanced inversion rearrangements affecting the *DMD* gene in three cases.

Note: all coordinates are in GRCh37.

2 Methods

We performed total RNA sequencing (RNAseq) on muscle biopsy and splice aberration events are evaluated using the Integrative Genomics Viewer v2.8.0.7. The Agilent Sureselect *DMD* targeted sequence capture method was used to enrich the entire *DMD* gene, which includes all exons and introns from the genomic DNA. DNA was analyzed by next-generation sequencing (NGS) on the Illumina MiSeq or NovaSeq™ 6000 with 2x150 paired-end reads. CNVs within the *DMD* gene were assessed using BioNano's NxClinical v6.1 software (El Segundo, CA). Split reads were visualized in IGV. Sanger sequencing was performed to confirm the breakpoint junctions wherever possible using case-specific primers across the breakpoints. In one case, Bionano optical genome mapping was used to confirm the results.

3 Results

In three males with clinical features of DMD and previous negative sequencing and del/dup genetic testing results, RNA sequencing analysis showed aberrant *DMD* expression. Analysis of NGS soft-clipped reads was used to discover the disease-causing complex rearrangements and investigate the breakpoints. A large balanced inversion affecting the *DMD* gene was found in all three cases. Sanger sequencing was used for confirmation of the breakpoints. In case 1, Bionano optical genome mapping was used to confirm the maternal results. Case 1 was found to harbor a 5.9 Mb paracentric inversion of chromosome Xp21.1-Xp21.3. The inverted region extends from Xp21.3 (chrX:26929141) through intron 2 in *DMD* gene (Xp21.1, 32873783). This inversion disrupts the *DMD* gene from exon 3 through the end of the gene and was maternally inherited. Case 2 was found to harbor a 9.1 Mb paracentric inversion of chromosome Xp21.1-Xp22.11. The inverted region extends from Xp22.11 (chrX:23028859) through intron 44 in the *DMD* gene (Xp21.1, chrX:32142959). This inversion disrupts the *DMD* gene from exon 45 through the end of the gene and found to be maternally inherited. Case 3 was found to harbor a 92.4 Megabase (Mb) pericentric inversion of chromosome Xp21.2 – Xq25. The inverted region extends from intron 67 in the *DMD* gene (Xp21.2, chrX:31214006) through Xq25 (chrX:123586570). This inversion disrupts the *DMD* gene from exon 68 through the end of the gene. Parental testing was not performed.

Normal chromosome X ideogram. The DMD gene is depicted in green.

Ideogram of Case 1: paracentric inversion disrupting the DMD gene (5.94 Mb from Xp21.3 to DMD intron 2).

Ideogram of Case 2: paracentric inversion disrupting the DMD gene (9.1 Mb from Xp22.11 to DMD intron 44).

Ideogram of Case 3: pericentric inversion disrupting the DMD gene (92.4 Mb from DMD intron 67 to Xq25).

Figure 1 - Ideogram of 3 cases of inversion disrupting the DMD gene

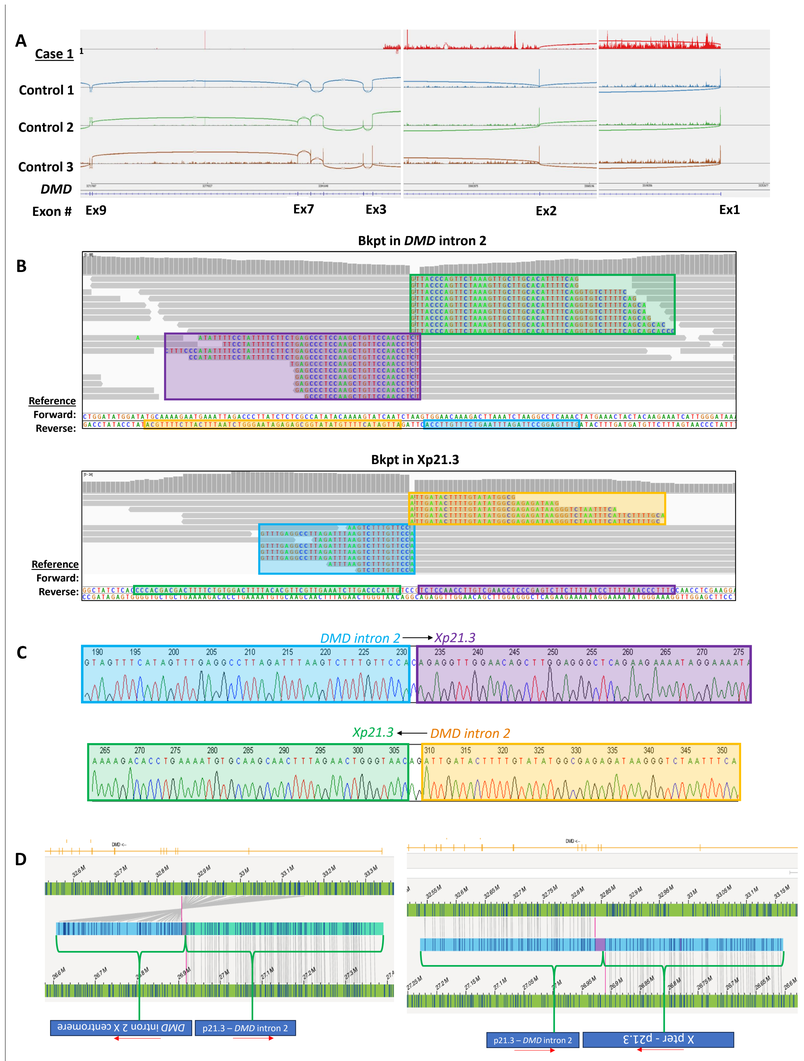
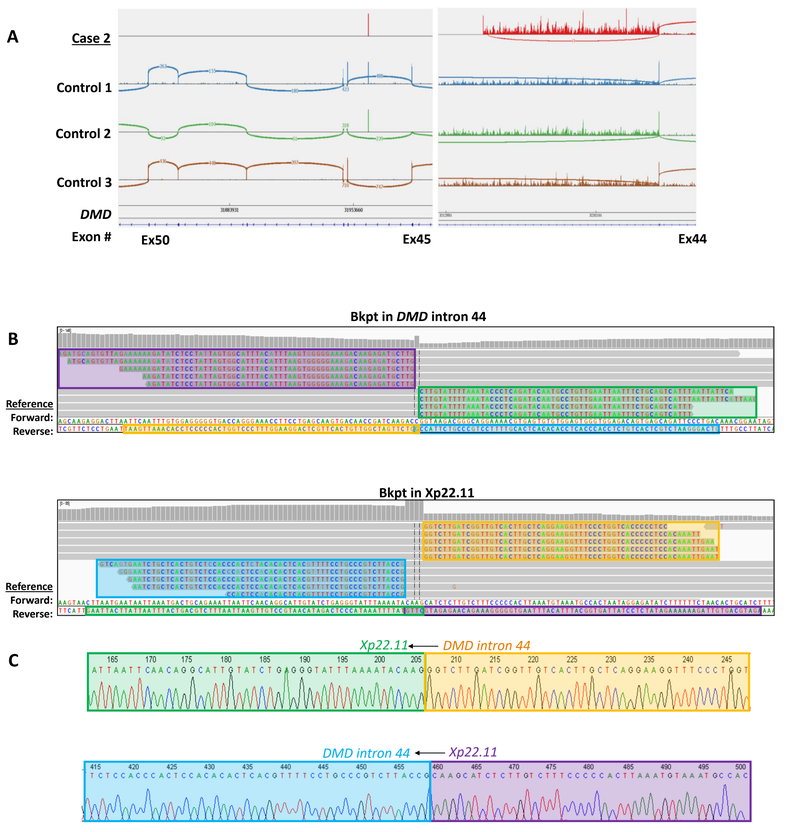
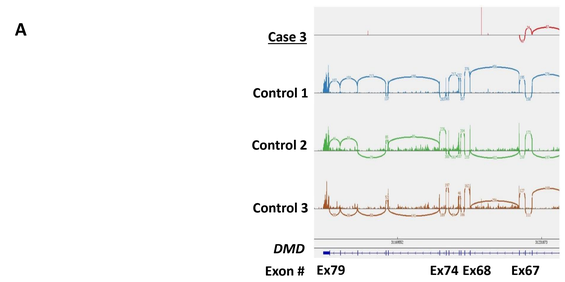


Figure 2 - Breakpoint analysis of case 1 harboring a paracentric inversion disrupting the DMD gene. The inversion is 5.94 Mb spanning from Xp21.3 to DMD intron 2. (bkpt=breakpoint)

A *RNA sequencing on muscle biopsy showing aberrant expression involving DMD exons 2-79 (exons 1-9 are shown).* ***B*** *IGV viewer of short-read whole DMD gene NGS sequencing showing the inversion bkpts in DMD intron 2 and Xp21.3. Mismatched bases from split reads (soft-clipped reads) are highlighted to indicate matching sequences. The split reads in DMD intron 2 match the reference sequence in Xp21.3, and the split reads in Xp21.3 match the reference sequence in DMD intron 2.* ***C*** *Sanger sequencing confirming the bkpts.* ***D*** *Bionano optical genome mapping confirms the paracentric inversion in the patient's mother.*





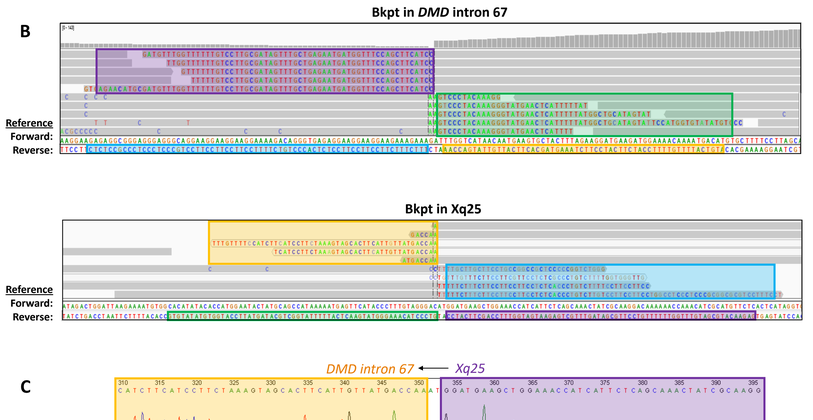


Fig. 3 Breakpoint analysis of case 2 harboring a paracentric inversion disrupting the DMD gene. The inversion is 9.1 Mb from Xp22.11 to DMD intron 44. (bkpt=breakpoint)

A *RNA sequencing on muscle biopsy showing aberrant expression involving DMD exons 44-79 (exons 44-50 are shown).* ***B*** *IGV viewer of short-read whole DMD gene NGS sequencing showing the inversion bkpts in DMD intron 44 and Xp22.11. Mismatched bases from split reads (soft-clipped reads) are highlighted to indicate matching sequences. The split reads in DMD intron 44 match the reference sequence in Xp22.11, and the split reads in Xp22.11 match the reference sequence in DMD intron 2.* ***C*** *Sanger sequencing confirming the bkpts.*

Fig. 4 Breakpoint analysis of case 3 harboring a paracentric inversion disrupting the DMD gene. The inversion is 92.4 Mb from DMD intron 67 to Xq25. (bkpt=breakpoint)

A *RNA sequencing on muscle biopsy showing aberrant expression involving DMD exons 67-79 (exons 67-79 are shown).* ***B*** *IGV viewer of short-read whole DMD gene NGS sequencing showing the inversion bkpts in DMD intron 67 and Xq25. Mismatched bases from split reads (soft-clipped reads) are highlighted to indicate matching sequences. The split reads in DMD intron 67 match the reference sequence in Xq25, and the split reads in Xq25 match the reference sequence in DMD intron 67.* ***C*** *Sanger sequencing confirming the bkpts.*

4 Conclusion

In conclusion, analysis of NGS soft-clipped reads facilitated the identification of large balanced inversions that could not be identified by conventional *DMD* assays. Our results emphasize the importance of utilizing breakpoint analysis of NGS data or long-read sequencing to identify complex rearrangements affecting the *DMD* gene, particularly in cases where dystrophinopathy was identified clinically and histologically. NGS split-read analysis results corroborate with the aberrant RNA expression observed. The identification of split reads in patients with dystrophinopathy enhances our understanding of the mechanisms underlying structural rearrangements, thereby facilitating the precise molecular diagnosis of DMD that is essential in clinical evaluation, variant classification, and therapeutics.