Potential for improved molecular diagnosis of facioscapulohumeral dystrophy (FSHD) through D4Z4 array quantitation using Bionano Optical Maps

Jonathan Pevsner

Abstract

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most common hereditary form of muscle disease along with Duchenne muscular dystrophy and myotonic muscular dystrophy. It is an autosomal dominant disease affecting 1:15,000 people. Clinically, FSHD is typically characterized by a distinctive regional distribution of muscle involvement including weakness of the face, scapulae, foot dorsiflexors and hip girdles. FSHD1 accounts for 95% of FSHD cases and is associated with a contraction of a 3.3 kilobase D4Z4 macrosatellite repeat in the subtelomeric region of chromosome 4q35. In unaffected individuals this repeat occurs in 11 to 150 repeat units. FSHD1 patients typically have 1 to 10 repeats, specifically of a haplotype called the A allele. An almost identical repeat in 10q26.3 is unrelated to the disease but complicates genetic testing significantly. The D4Z4 contraction is thought to result in a open chromatin structure that permits expression of the double homeobox 4 (DUX4) gene, facilitated by a polyadenylation sequence specific to the A allele. DUX4 encodes a transcription factor, and its aberrant expression may cause deleterious effects on downstream targets in muscle. Genetic testing for FSHD, while sensitive and specific, is also complex, laborious, and specialized. We are using Saphyr technology of Bionano Genomics to determine the genomic architectures of the D4Z4 region of chromosomes 4q35 and 10q26.3 in FSHD. The Saphyr system uses NanoChannel arrays to linearize and image high molecular weight genomic DNA. The repeats on chromosomes 4q35 and 10q26.3 are readily distinguished, the repeat sizes at each locus are quantified, and the A and B alleles are distinguished. We have assembled D4Z4 genomic maps from m20 normal individuals and quantified as many as 272 D4Z4 repeats on a single allele. We contrast results from Saphyr with those of whole genome sequencing, and 10X Genomics sequencing at this locus.

FSHD case 1: pedigree and phenotype

Figure 1. Pedigree consistent with autosomal dominant inheritance of FSHD. Proband 50A was a 42 year old man who has had difficulty lifting his arms over head. On examination he had slight weakness of his facial muscles (orbicularis oculi and oris). He had scapular winging when attempting to abduct or extend his arms. He had moderate weakness in the elbow and knee flexors and extensors, hip flexors, and severe weakness in his foot dorsiflexors.

Figure 2. Reads were de novo assembled and then compared to human genome build GRCh38 (green bar). Above: IrysView software image of chromosomes 4q (top) and 10q (bottom). Genomic DNA was nicked with EcoR I and/or BglI or XapI.

Bionano Genomics Methodology

Bionano Genomics Saphyr system is based on an optical Next-Generation Mapping technique (NGM). In this study, those Bionano maps were used to detect structural variants (SVs).

- DNA > 100 kb is extracted from blood, labeled at specific restriction sites, and linearized through NanoChannel arrays.
- Molecular patterns are digitized and assembled de novo to create megabase-scale optical maps.
- For quality control SVs are filtered based on quality scores of the assembly, the alignments, and the variants.

FSDH case 2: identifying D4Z4 contractions on 4q

In a second case the number of repeats determined by Bionano (4qA, 6 repeats; 4qB, 58 repeats) closely matched the estimates from Iowa (4qA, 6 repeats; 4qB, 34 and >60 repeats).

Whole genome sequence approaches

We performed whole genome sequencing (1) using Illumina short-read technology at 12x average depth of coverage (n=3 controls) and (2) using the barcoding approach of 10X Genomics (63x average depth of coverage, 99.3% of SNPs phased, N50 phase block 2.4 Mb). We could not resolve the D4Z4 region with either method nor could we distinguish the 4q and 10q loci.

Conclusions

FSHD is a devastating muscular dystrophy that is usually caused by contractions of a highly repetitive genomic region (D4Z4) on chromosome 4q. Bionano Genomics Saphyr System is able to (1) determine the number of repeats on both the 4qA and 4qB alleles, and (2) distinguish the repeats on chromosomes 4q and 10q. The current gold standard assay is laborious and highly specialized, requiring PFGE and Southern blotting. The Bionano approach may be useful for diagnosis of FSHD and for studies of the underlying biological mechanisms.

Acknowledgments

I thank Dr. Kathryn Wagner for obtaining informed consent, providing patient samples, and for clinical descriptions. I thank Dr. Alex Hastie and colleagues at Bionano for data analysis. I am not supported by Bionano Genomics other than my attendance at this meeting.

References

2. Cao H et al. (2014) Gigascience. 3(1):34. PMID: 25071094