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

Jonathan Pevsner

Kennedy Krieger Institute

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 an 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 genome maps from n=20 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.

Standard D4Z4 assay for FSHD

Standard FSHD assays must distinguish the nearly identical 4q and 10q subtelomeric D4Z4 arrays. A “gold standard” method includes pulsed field gel electrophoresis of DNA cleaved distal to the D4Z4 region and at rare 4q-specific or 10q-specific restriction sites followed by Southern blotting with hybridization probe p13E-11 (ref. 1). We present data here from the CLIA lab at the University of Iowa Hospitals and Clinics, an assay based on restriction digests with *Eco*RI and/or *Bln*I or *Xap*I.

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.

FSHD case 1: pedigree and phenotype

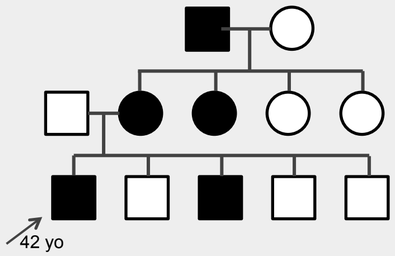


Figure 1. Pedigree consistent with autosomal dominant inheritance of FSHD. Proband 50A was a 42 year old man who has had difficulty lifting his arms over head. On examination he had slight weakness of his facial muscles (orbicularis oculi and oris). He had scapular winging when attempting to abduct or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

FSHD case 1: identifying D4Z4 contractions on 4q

Testing at U. of Iowa revealed 3.3 repeats (on 4qA) and 7.33 repeats (on 4qB). Bionano results were closely comparable with 4 repeats (4qA) and 8 repeats (4qB). This is shown graphically using IrysView software following whole genome *de novo* assembly (Figure 2, upper panel). This analysis identifies the D4Z4 repeat status on chromosome 4q A and B alleles, and further differentiates paralogs on the closely related 10q locus (Figure 2, lower panel).

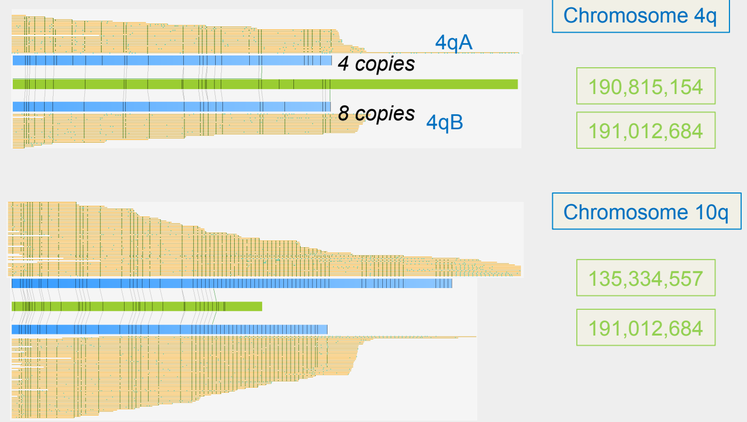


Figure 2. Reads were de novo assembled and then compared to human genome build GRCh38 (green bar). Above: IrysView software image of chromosomes 4q (top) and 10q (bottom). Genomic DNA was nicked with Nt.*Bsp*QI (top) or Nb.*Bss*SI (bottom; similar results were obtained using either enzyme).

FSHD 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).

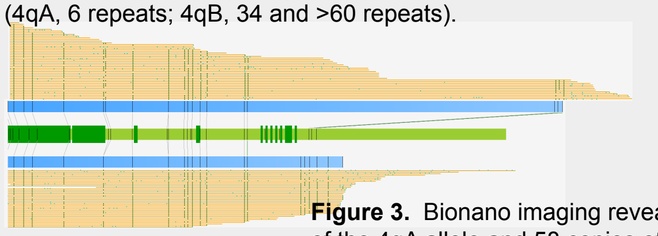
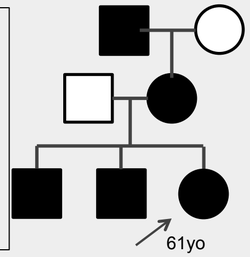


Figure 3. Bionano imaging revealed 6 copies of the 4qA allele and 58 copies of 4qB.

FSHD case 3: two 4qA alleles

Figure 4. Proband 64B is a 61 year old woman who had 20 years of neck and shoulder pain. She could not raise her arm over her head and her shoulders sloped downwards. On examination she had facial muscle weakness (o. oculi and oris) as well as scapular winging. She had good strength in her arms but weakness in her hip flexors and extensors, knee extensors and foot dorsiflexors.



The proband was part of a pedigree suggestive of autosomal dominant inheritance. Clinical testing indicted 7.33 and 10.66 repeats of the 4qA allele. Bionano results were closely consistent, showing 8 and 10 repeats.

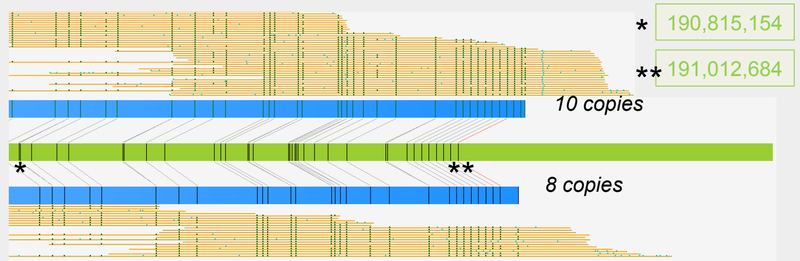


Figure 5. Bionano imaging revealed 10 copies of the 4qA allele and 8 copies of 4qB.

Whole genome sequence approaches

We performed whole genome sequencing (1) using Illumina short-read technology at 120x 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

1. Lemmers RJ et al. (2004) *Am J Hum Genet*. 75(6):1124-30. PMID: 15467981
2. Cao H et al. (2014) *Gigascience*. 3(1):34. PMID: 25671094