Identification and accurate sizing of D4Z4 repeat units in patients suspected of facioscapulohumeral muscular dystrophy (FSHD) using optical genome mapping

Naga Guruju, Vanessa Jump, Ramesh Nallamilli, Ruby Liu, Zhi wen Luo, Madhuri Hegde Revvity Omics, Pittsburgh, PA

BACKGROUND

- Facioscapulohumeral muscular dystrophy (FSHD) is characterized by progressive and often asymmetric muscle weakness of the face, scapular stabilizers, shoulders, arms, lower leg, and hip girdle. Around 20% of patients are wheelchair-bound, and some also present with sensory, cardiac or neurological manifestations. The age of onset is variable and ranges from infancy to adult.
- Two genetically distinct FSHD subtypes have been identified, FSHD1 and FSHD2. FSHD1 is observed in 95% of patients.
- The number of D4Z4 repeats on chromosome 4q generally varies from 11 to 150 copies in healthy individuals. Contraction to 1-10 repeats in the permissive allele 4qA causes FSHD1. The telomeric region on chromosome 10 (q26) is 99% identical to 4qA region and is not associated with FSHD.
- FSHD2 is clinically identical to FSHD1 but has a different genetic cause. In FSHD2, affected individuals have 11 or above repeats (normal range) on the permissive chromosome 4 haplotype (4qA) and a pathogenic variant in either *SMCHD1* (80% cases), *DNMT3B* or other genes causing hypomethylation of the D4Z4 region.

METHODS

Optical Genome Mapping:

- We performed optical genome mapping (OGM) using the Bionano Genomics Saphyr with subsequent analysis by Bionano Enfocus FSHD analysis software (Bionano, San Diego, CA) to identify FSHD haplotype and D4Z4 repeat number in patients suspected of FSHD.
- Molecules aligning the D4Z4 repeat regions on chromosome 4 using human genome reference build GRCh38 are distinguished from regions of high homology on chromosome 10.
- The permissive (4qA) and non-permissive alleles (4qB) were assigned using the dynamic-programming algorithm included in the Enfocus FSHD analysis pipeline.
- The D4Z4 repeat size is determined based upon the measurement of the interval distance between labels flanking the D4Z4 arrays.

***SMCHD1* gene sequencing by Next-Generation Sequencing (NGS):**

- A custom Agilent SureSelect enrichment kit was used to enrich the *SMCHD1* gene, followed by next-generation sequencing on an Illumina system with 100 base pair paired-end reads. The analyzed regions include the coding exons and 10bp of flanking intronic regions on both sides of each exon.
- Copy number variants (CNVs) were assessed using NxClinical software (BioNano, El Segundo, CA).

RESULTS

- We performed FSHD testing using optical mapping for 547 patients suspected of FSHD, including 301 males (55%) and 246 females (45%).
- 308/547 patients referred were positive for a D4Z4 contraction resulting in a diagnosis of FSHD1.
- Additional testing performed for 252/547 pts to identify variants in *SMCHD1* resulted in the identification of 10 patients positive for FSHD2 (pathogenic or likely pathogenic *SMCHD1* variant, see Table 2). In our FSHD2 cohort, the 4qA allele size ranged from 8 to 18 repeats. We observed 5 patients with a variant of unknown significance in *SMCHD1*.
- No follow-up FSHD2 testing was requested for 88/295 patients carrying at least one 4qA allele with normal repeats. The possibility of FSHD2 in 25/88 of these patients showing a 4qA allele between 11 and 20 repeats could not be ruled out.
- 2 cases showed biallelic contraction (A1/A10 and A6/A9). 4 patients showed homozygous contraction (A5/A5, A8/A8, and two patients with A7/A7).
- 9/308 (3%) of patients positive for 4qA contraction had mosaic 4q alleles with a contraction on at least one 4qA allele.

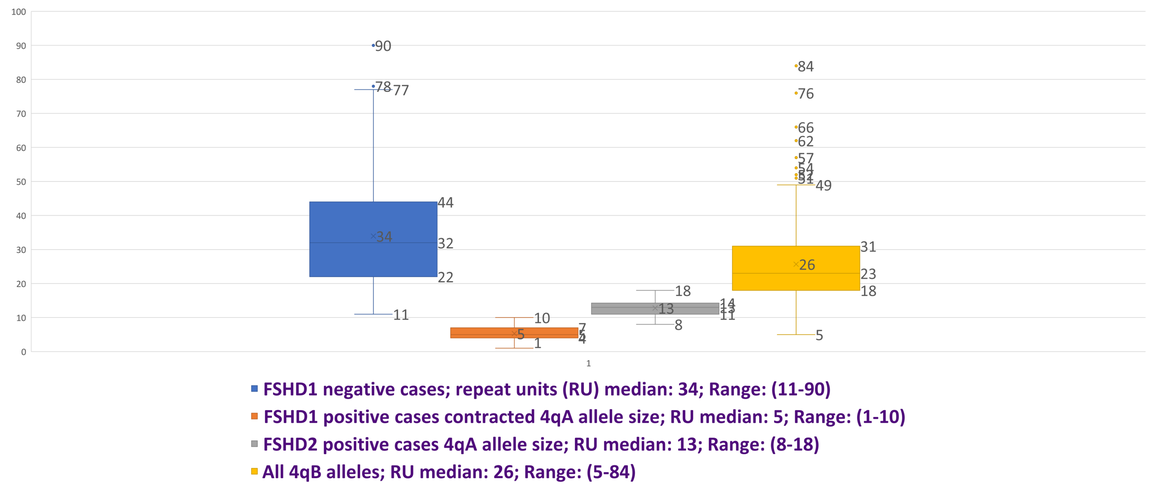

Figure 1. Comparison of D4Z4 Allele Sizes Among Patients Suspected of FSHD

Disease Association	Haplotype	D4Z4 Number of Repeats	SMCHD1	547 Total Patients tested (overall diagnostic yield = 58%)
FSHD Type 1	4qA (Permissive)	1-10	Not analyzed	308 patients (56%)
FSHD Type 2	4qA (Permissive)	8-18	Pathogenic /LP variant	9 patients (2%, see Table 2 below)
FSHD 1 and 2	4qA (Permissive)	1-10	Pathogenic variant	1 patient
Mosaic FSHD1	4qA (Permissive)	1-10	Not analyzed	9 patients
cis duplication of region proximal to D4Z4 repeats	4qA (Permissive)	1-10	Not analyzed	3 patients
Biallelic contraction	4qA (Permissive)	1-10	Not analyzed	2 patients
Homozygous contraction	4qA (Permissive)	1-10	Not analyzed	4 patients

Table 1. FSHD Patient Result Summary

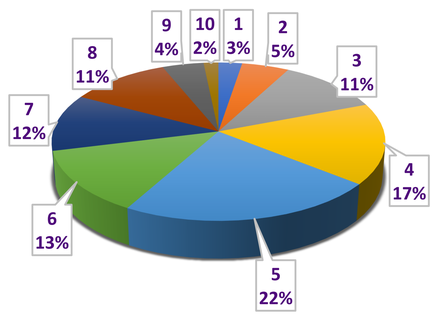

Figure 2. 4qA allele size distribution among 308 patients positive for FSHD1. The most frequent 4qA repeat size observed was 5 repeats (69 pts, 22%), followed by 4 repeats (51 pts, 17%), and 6 repeats (40 pts, 13%).

Variant	Classification	4q35 allele 1	4q35 allele 2
SMCHD1 c.2071_2075del	Pathogenic	A13	A19
SMCHD1 c.3276+4_3276+7del	Pathogenic	A13	B24
SMCHD1 c.4566G>A (p.Thr1522=)	Pathogenic	A15	A29
SMCHD1 c.1462C>A (p.Arg488Ser)	Pathogenic	A11	A35
SMCHD1 c.1186C>T (p.Gln396Ter)	Pathogenic*	A23	A42
SMCHD1 c.2176_2179del	Likely Pathogenic	A14	B66
SMCHD1 c.35_45dup	Likely Pathogenic	A18	B22
SMCHD1 c.3938C>G (p.Ser1313Ter)	Likely Pathogenic	A8	A55
SMCHD1 c.5286dup	Likely Pathogenic	A11	A19

Table 2. Variants detected in the *SMCHD1* gene in patients positive for FSHD2.

* Patient is asymptomatic, family studies showed 4qA/11 in combination with variant is associated with FSHD2.

CONCLUSION

- Southern blot is the gold standard method in the diagnosis of FSHD however, southern blot is a labor-intensive, time consuming, radioactive method that requires a large quantity of high-quality DNA.
- Our results demonstrate that whole genome optical mapping offers an alternative method for FSHD1 diagnosis that is less time consuming and yields a more accurate estimate of the number of D4Z4 units (+/- 1 repeat sizing) than the conventional southern blotting technique. In addition, OGM can identify mosaic cases, bi allelic contractions and homozygous contraction of D4Z4 allele.
- A combination of optical genome mapping to identify FSHD haplotype and D4Z4 repeat number, and NGS to identify sequence and copy number variants in the *SMCHD1* gene is a practical and cost-effective option for accurate diagnosis of FSHD type 1 and 2.